

THE HANDBOOK OF
GROUNDWATER
ENGINEERING

Second Edition

Edited by
Jacques W. Delleur



CRC Press
Taylor & Francis Group

THE HANDBOOK OF
GROUNDWATER
ENGINEERING

Second Edition

THE HANDBOOK OF
GROUNDWATER
ENGINEERING

S e c o n d E d i t i o n

E d i t e d b y
Jacques W. Delleur



CRC Press

Taylor & Francis Group

Boca Raton London New York

CRC Press is an imprint of the
Taylor & Francis Group, an informa business

CRC Press
Taylor & Francis Group
6000 Broken Sound Parkway NW, Suite 300
Boca Raton, FL 33487-2742

© 2007 by Taylor & Francis Group, LLC
CRC Press is an imprint of Taylor & Francis Group, an Informa business

No claim to original U.S. Government works
Printed in the United States of America on acid-free paper
10 9 8 7 6 5 4 3 2 1

International Standard Book Number-10: 0-8493-4316-X (Hardcover)
International Standard Book Number-13: 978-0-8493-4316-2 (Hardcover)

This book contains information obtained from authentic and highly regarded sources. Reprinted material is quoted with permission, and sources are indicated. A wide variety of references are listed. Reasonable efforts have been made to publish reliable data and information, but the author and the publisher cannot assume responsibility for the validity of all materials or for the consequences of their use.

No part of this book may be reprinted, reproduced, transmitted, or utilized in any form by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying, microfilming, and recording, or in any information storage or retrieval system, without written permission from the publishers.

For permission to photocopy or use material electronically from this work, please access www.copyright.com (<http://www.copyright.com/>) or contact the Copyright Clearance Center, Inc. (CCC) 222 Rosewood Drive, Danvers, MA 01923, 978-750-8400. CCC is a not-for-profit organization that provides licenses and registration for a variety of users. For organizations that have been granted a photocopy license by the CCC, a separate system of payment has been arranged.

Trademark Notice: Product or corporate names may be trademarks or registered trademarks, and are used only for identification and explanation without intent to infringe.

Library of Congress Cataloging-in-Publication Data

Delleur, J. W. (Jacques Willy)
The Handbook of groundwater engineering / Jacques W. Delleur. -- 2nd ed.
p. cm.
Prev. ed. entered under title; Delleur presented as editor in chief.
Includes bibliographical references and index.
ISBN-13: 978-0-8493-4316-2 (alk. paper)
ISBN-10: 0-8493-4316-X (alk. paper)
1. Groundwater flow. 2. Groundwater--Pollution. 3. Groundwater--Management. I. Title.

TC176.H35 2006
628.1'14--dc22

2006021247

Visit the Taylor & Francis Web site at
<http://www.taylorandfrancis.com>

and the CRC Press Web site at
<http://www.crcpress.com>

Introduction

This reference text is an expanded and updated version of the text by the same name published eight years earlier. As such, this text is a compilation of many classical and more modern approaches to analyzing and modeling subsurface flow and transport.

Since Darcy's pioneering work well over a century ago, our qualitative and quantitative understanding of the flow of water and transport of chemicals in the subsurface has soared to great heights. Much of our current understanding of flow processes has been developed over the past 40 years, and our understanding of transport processes has developed mainly over the last 25 years. Substantial recent advances have been driven largely by two factors: (a) the development of inexpensive high-speed computing, and (b) the international focus on environmental quality and global change.

Much of the groundwater work performed earlier in this century revolved around very simplistic conceptualizations of the subsurface, simplifications that were so drastic as to allow analytical solutions to flow problems and the subsequent development of type-curves. Consequently, well installation and design, and groundwater maintenance revolved around such idealizations as homogeneity, isotropy, and infinite domains, which as we all know, are completely unrealistic assumptions.

With the advent of computing in the 1960s, through its modern refinements in software and hardware, and derivation of new sophisticated numerical algorithms, we have made a quantum leap forward in our ability to predict the behavior of our groundwater resources. This, coupled with the tremendous amount of research dollars, tied mainly to the environment, has led to a flood of research manuscripts, new engineering tools, and commercially available groundwater flow and transport software.

Unfortunately, while we have made tremendous advances in our ability to model the subsurface, our ability to cost-effectively measure material coefficients that go into the models has not kept pace. Thus, modern models are full of uncertainty in input data, and hence, their output is uncertain. This uncertainty has led to the introduction of stochastic tools to the field of subsurface hydrology.

This reference text is a compilation of many classical and more modern approaches to understanding and modeling subsurface flow and transport. Chapters 1 through 4 provide an introduction to the subject. In Chapter 1, the basic functions and terminology requisite to the study of flow and transport are presented. This chapter also provides a historical background of the field. Subsequently, steady one- and two-dimensional flow and transport (Chapter 2) and deterministic two- and three-dimensional flow (Chapter 3) are discussed. Chapter 2 gives an overview of the occurrence of groundwater, while Chapters 3 and 4 present basic principles.

Flow and transport in geologic media are characterized by tremendous uncertainty in the geological environment. One can safely say that in no other field of engineering are we faced with more complexity and uncertainty than in groundwater flow and transport engineering. Uncertainty is dealt with in a number

of different ways in Chapters 5, 8, 16, 18, and 19. Stochasticity in unsaturated flow and transport is dealt with in Chapters 5 and 8, while Chapters 18 and 19 handle the stochastic saturated transport problem. Interpolation via geostatistics is studied in Chapter 16. Deterministic unsaturated flow and transport are handled in Chapters 6, 7, and 22, while Chapters 9, 12, 13, 20, 21, and 23 handle the saturated problem. Chapters 9, 20, and 21 deal with preferential flow and fractures. Chapter 13 extends the classical isothermal model to the nonisothermal case and Chapter 12 focuses on seawater intrusion. While most chapters involve modeling to one degree or another, Chapters 23, 24, and 26 specifically focus on issues related to modeling, such as verification and case studies. Chapters 14, 15, and 17 focus on characterizing aquifers and contaminants.

Remediation or containment of contaminants in the subsurface has become critically important in recent years. This involves observation, injection, and pumping well design and these topics are covered in Chapters 10 and 11. Chapter 27 discusses sustaining our potable water. Landfill design and monitoring are discussed in Chapters 33, 34, 35, and 37, while Chapters 36 and 37 focus on remediation strategies. In remediation and containment strategies for contaminated sites and for water rights in general, legal issues continually crop up and these are discussed in Chapter 32.

Most recently, climate change (Chapter 28) has come to the forefront of science. In the context of groundwater, this has given rise to the field of echo-hydrology (Chapter 29) and the use of GIS (Chapter 30).

This revised text has been in the works for a little over two years. The authors are considered to be among the best in their fields. All chapters have been peer-reviewed.

John H. Cushman
Department of Earth & Atmospheric Sciences
Department of Mathematics
Purdue University

Preface

A substantial amount of new knowledge in the field of groundwater engineering has been accumulated over the last decade. In particular, many new developments have taken place in the analysis of contaminant transport in the saturated and unsaturated zones. New techniques for the remediation of contaminated groundwater have been implemented. New approaches in computer modeling of the flow and movement of contaminants have appeared. The problem of modeling the heterogeneity of the geological formations and their effects on the flow and transport continue to be a challenge and have been the subject of many papers. This information is dispersed in many professional journals, textbooks, and reports. This second edition of the *Handbook of Groundwater Engineering* attempts to synthesize these facts and to provide this information in an easily accessible form. It contains a blend of professional practice and scientific information.

The typical readers of the handbook are expected to be geohydrologists and engineers possessing bachelors or masters degrees in applied sciences or in engineering, possibly with some years of experience.

The first chapter sets the stage by giving an introduction to the history of water management before the 19th century and an overview of the progress in groundwater hydrology until the 1980s. Chapters 2 and 3 provide an introduction to the geological setting and the analysis of simple flow and contaminant transport problems. Chapters 10 and 11 are concerned with the production of groundwater and hydraulics of wells. The last seven chapters (31 to 37) provide practical information concerned with pollution remediation, landfills, and the associated legal and regulatory aspects. The remaining 27 chapters deal with the analytical tools for the description, modeling of flow, and transport phenomena in the saturated and in the unsaturated zones.

This second edition includes 11 new chapters: History of Groundwater Hydrology (1), Infiltration and Run-on under Spatially Variable Hydrologic Properties (8), Sea Water Intrusion into Coastal Aquifers (12), Groundwater and Heat Flow (13), Groundwater Model Validation (24), Scale Issues (25), Accounting for Aquifer Heterogeneity in Solute Transport Modeling: A Case Study (26), Determining Sustainable Groundwater Development (27), The Impact of Global Change on Groundwater (28), Ecohydrology (29), and Evapotranspirative Cover Systems for Waste Containment (34). Three chapters have been completely rewritten: Infiltration (7), Nonreactive Contaminant Transport (18) and Contaminant Transport in the Unsaturated Zone—Theory and Modeling (22). Most of the remaining 23 chapters received substantial updating as needed.

I am grateful to have an excellent group of authors who are leaders in their respective specialties. Their names and affiliations appear in the headings of the individual chapters. The chapters were reviewed in a manner similar to papers in scientific and professional journals. I am grateful to the reviewers. They made a special effort to ascertain that the information presented is sound. Some reviewers read two chapters.

Editor

Jacques W. Delleur received his doctor of engineering science degree at Columbia University in New York City in 1955, his master's degree in civil engineering at Rensselaer Polytechnic Institute in Troy, New York in 1950, and his civil and mining engineering degree at the Universidad Nacional de Colombia (National University of Colombia) in 1949. In 1955, he joined Purdue University where he currently is Professor Emeritus of hydraulic engineering and was Head of the hydraulics and systems engineering area in the School of Civil Engineering. Dr. Delleur taught intermediate and advanced graduate courses in subsurface hydrology, surface hydrology, statistical hydrology, and hydraulics. He founded the graduate program in hydrology and hydraulics in the School of Civil Engineering at Purdue. He is author or coauthor of two books on hydrologic time series analysis. He is author or coauthor of more than sixty papers in refereed journals, seventy papers in conference proceedings, and sixty technical reports. These cover the areas of subsurface hydrology, hydrologic modeling, stochastic hydrology, urban hydrology, and hydraulics. The most recent publications related to groundwater coauthored by J.W. Delleur are concerned with the flow and transport of dissolved substances and how they are affected by geologic heterogeneity.

Dr. Delleur's research has been supported by the U.S. Department of the Interior, the National Science Foundation, the U.S. Department of Transportation, and the U.S. Department of Agriculture. He has served as an advisor to the U.S. Geological Survey, is a member of the international board of advisors of the American Society of Civil Engineers (ASCE) *Journal of Hydrologic Engineering*, is an associate editor of the *Journal of Hydraulic Engineering* (ASCE), and is a member of the scientific council of the *Revue des Sciences de l'Eau/Journal of Water Science*. He served as a reviewer for the National Science Foundation, for the Natural Sciences and Engineering Research Council of Canada, and for the scientific journals *Water Resources Research*, *Journal of Hydrology*, *Journal of the American Water Resources Association*, and the *Journal of Hydraulic Engineering*. He is a Fellow of the Indiana Academy of Sciences; received the 1961 Freeman Fellow Award of the ASCE; in 1983, received the NSF/CNRS US-France Senior Scientist Exchange Award; in 1992, received the Charles Harold Bechert Award of the Indiana Water Resources Association for significant contribution to the water resources profession in Indiana; and in 2002, received the Ven Te Chow Award from the ASCE for lifetime achievement in the field of hydrologic engineering. While on sabbatical leave, Dr. Delleur did research in hydrology at the French National Hydraulics Laboratory (1968–1969 and 1976–1977), at the University of Grenoble, France (1961–1962 and 1983–1984), and at the Vrije Universiteit Brussel (Free University of Brussels, Belgium (1991). He has been a guest lecturer

at the École Polytechnique Fédérale de Lausanne (Federal Polytechnics School, of Lausanne), Switzerland, at the Free University of Brussels, Belgium, at Imperial College in London, at the University of Tokushima, Japan, at the Indian Institute of Technology in Kanpur, India, at the Mahommadia School of Engineering in Rabat, Morocco, at the Taiwan National University in Taipei, Taiwan, and at the Universidad de los Andes (University of the Andes) in Bogotá, Colombia.

Contributors

Rafael Angulo-Jaramillo

Laboratoire d'Etude des Transferts
en Hydrologie et Environment
(LTHE)
Grenoble, France

Shmuel Assouline

Institute of Soil, Water and
Environmental Sciences
Volcani Center
Bet Dagan, Israel

Paul M. Barlow

U.S. Geological Survey
Northborough, Massachusetts

Jacob Bear

Faculty of Civil & Environmental
Engineering
Technion-Israel Institute of
Technology
Haifa, Israel

Michael H. Berman

GeoSyntec Consultants
Santa Barbara, California

Ernest R. Blatchley, III

School of Civil Engineering
Purdue University
West Lafayette, Indiana

Hans Boonstra

Alterra-ILRI
Wageningen, The Netherlands

Paul J. Botek

GeoSyntec Consultants
Santa Barbara, California

John D. Bredehoeft

The Hydrodynamics Group
Sausalito, California

Barry R. Christopher

Independent Consultant
Morrow, Georgia

C. Corradini

Istituto di Idraulica
University of Perugia
Perugia, Italy

Dennis L. Corwin

U.S. Department of Agriculture
Agricultural Research Service
U.S. Salinity Laboratory
Riverside, California

John H. Cushman

Department of Earth and
Atmospheric Science and
Department of Mathematics
Purdue University

Edoardo Daly

Department of Civil and
Environmental Engineering &
Nicholas School of
the Environment and
Earth Sciences
Duke University
Durham, North Carolina

Devaraj de Condappa

Institut de Recherche pour le
Développement
Cotonou, Bénin

F. De Smedt

Department of Hydrology and
Hydraulic Engineering
Vrije Universiteit
Brussels, Belgium

Jacobus J. de Vries

Faculty of Earth and Life Sciences
Vrije Universiteit
Amsterdam, The Netherlands

Samuel Debionne

Hydrowide
Grenoble, France

Jacques W. Delleur

School of Civil Engineering
Purdue University
West Lafayette, Indiana

Deborah Elcock Argonne National Laboratory Washington, D.C.	Susan Hubbard Lawrence Berkeley National Laboratory Berkeley, California	Kyoung Jae Lim Agricultural and Biological Engineering Purdue University West Lafayette, Indiana
Bernard Engel Agricultural and Biological Engineering Purdue University West Lafayette, Indiana	David W. Hyndman Department of Geological Sciences Michigan State University East Lansing, Michigan	P. A. Lapcevic British Columbia Ministry of the Environment Nanaimo, British Columbia Canada
T. R. Ginn Civil and Environmental Engineering University of California at Davis Davis, California	Chad T. Jafvert School of Civil Engineering Purdue University West Lafayette, Indiana	Darrell I. Leap Department of Earth and Atmospheric Sciences Purdue University West Lafayette, Indiana
R. S. Govindaraju Purdue University West Lafayette, Indiana	Mikko I. Jyrkama Department of Civil Engineering University of Waterloo Waterloo, Ontario, Canada	Keith Loague Department of Geological and Environmental Sciences Stanford University Stanford, California
Milton E. Harr School of Civil Engineering Purdue University West Lafayette, Indiana	Gabriel Katul Nicholas School of the Environment and Earth Sciences & Department of Civil and Environmental Engineering Duke University Durham, North Carolina	John S. McCartney Department of Civil, Architectural and Environmental Engineering University of Texas at Austin Austin, Texas
Ahmed E. Hassan Desert Research Institute Las Vegas, Nevada and Irrigation and Hydraulics Department Cairo University Giza, Egypt	Beth A. Keister URS Corporation Minneapolis, Minnesota	R. Morbidelli Instituto di Idraulica University of Perugia Perugia, Italy
Randel Haverkamp Laboratoire d'Etude des Transferts en Hydrologie et Environnement Université de Grenoble Grenoble, France	Peter K. Kitanidis Department of Civil and Environmental Engineering Stanford University Stanford, California	N. Nahar School of Civil Engineering Purdue University West Lafayette, Indiana
Jan W. Hopmans Department of Land, Air and Water Resources University of California at Davis Davis, California	Natalie Kleinfelter Department of Earth and Atmospheric Science and Department of Mathematics Purdue University	Kumar C. S. Navulur Digital Globe Longmont, Colorado
Michael F. Houlihan GeoSyntec Consultants Columbia, Maryland	Leonard F. Konikow Water Resources Division US Geological Survey Reston, Virginia	Loring F. Nies School of Civil Engineering Purdue University West Lafayette, Indiana

K. S. Novakowski Department of Civil Engineering Queen's University Kingston, Ontario, Canada	Jiří Šimůnek Department of Environmental Sciences University of California at Riverside Riverside, California	Pierre Viallet Hydrowide Grenoble, France
Jean-Yves Parlange Department of Agricultural and Biological Engineering Cornell University Ithaca, New York	Richard Soppe Alterra-ILRI Wageningen, The Netherlands	Clifford I. Voss U.S. Geological Survey Reston, Virginia
Amilcare Porporato Department of Civil and Environmental Engineering and Nicholas School of the Environment and Earth Sciences Duke University Durham, North Carolina	Tammo S. Steenhuis Department of Biological and Environmental Engineering Cornell University Ithaca, New York	Rony Wallach Department of Soil and Water Sciences Hebrew University of Jerusalem Jerusalem, Israel
Markus G. Puder Argonne National Laboratory Washington, D.C.	E. A. Sudicky Department of Earth Sciences University of Waterloo Waterloo, Ontario, Canada	William B. White Department of Geosciences Pennsylvania State University State College, Pennsylvania
Nancy L. Raneck Argonne National Laboratory Washington, D.C.	Jon F. Sykes Department of Civil Engineering University of Waterloo Waterloo, Ontario, Canada	Chunmiao Zheng Department of Geological Sciences University of Alabama Tuscaloosa, Alabama
Thomas E. Reilly U.S. Geological Survey Reston, Virginia	John E. Thompson School of Civil Engineering Purdue University West Lafayette, Indiana	Quanlin Zhou Earth Sciences Division Lawrence Berkeley National Laboratory Berkeley, California
Pedro C. Repetto Vice President URS Corporation Denver, Colorado	Martinus Th. van Genuchten George E. Brown, Jr. Salinity Laboratory U.S. Department of Agriculture Agricultural Research Service Riverside, California	Jorge G. Zornberg Department of Civil, Architectural and Environmental Engineering University of Texas at Austin Austin, Texas
Yoram Rubin University of California Department of Civil and Environmental Engineering Berkeley, California	John A. Veil Argonne National Laboratory Washington, D.C.	

Reviewers

Shmuel Assouline

Institute of Soil, Water, and Environmental Sciences
Agricultural Research Organisation
Volcani Center
Bet Dagan, Israel

Craig Benson

Department of Civil and Environmental Engineering
University of Wisconsin-Madison
Madison, Wisconsin

Ryan Berg

Ryan R. Berg & Associates, Inc.
Woodbury, Minnesota

Alexandre R. Cabral

Department of Civil Engineering
University of Sherbrooke
Sherbrooke, Québec, Canada

Jinsong Chen

Lawrence Berkeley National Laboratory
Berkeley, California

Prabhakar Clement

Department of Civil Engineering
Auburn University
Auburn, Alabama

Brent Clothier

HortResearch
Palmerston North, New Zealand

Jane Frankenberger

School of Agricultural and Biological Engineering
Purdue University
West Lafayette, Indiana

R.S. Govindaraju

School of Civil Engineering
Purdue University
West Lafayette, Indiana

Andrew Guswa

Picker Engineering Program
Smith College
Northampton, Massachusetts

Philip T. Harte

U.S. Geological Survey
Suncook, New Hampshire

Edwin Harvey

School of Natural Resources
University of Nebraska
Lincoln, Nebraska

Jan W. Hopmans

Department of Land, Air, and Water Resources
University of California
Davis, California

Karl Keel

URS Corporation
Minneapolis, Minnesota

Natalie Kleinfelter

Department of Mathematics
Purdue University
West Lafayette, Indiana

Christian Langevin

U.S. Geological Survey
Reston, Virginia

Angel Martin

U.S. Geological Survey
Urbana, Illinois

Jonathan B. Martin

Department of Geological Sciences
University of Florida
Gainesville, Florida

Michael May

URS Corporation
Denver, Colorado

David J. Mulla

Department of Soil, Water, and Climate
University of Minnesota
St. Paul, Minnesota

Kumar C.S. Navulur

Digital Globe
Longmont, Colorado

Mike Paque

Ground Water Protection Council
Oklahoma City, Oklahoma

Flynn Picardal

School of Public and Environmental Affairs
Indiana University
Bloomington, Indiana

Thomas E. Reilly
U.S. Geological Survey
Reston, Virginia

Vijay Singh
Department of Civil and
Environmental Engineering
Louisiana State University
Baton Rouge Louisiana

Charlie Speirs
GeoSyntec Consultants
Kennesaw, Georgia

Daniel M. Tartakovsky
Department of Mechanical and
Aerospace Engineering
University of California,
San Diego
La Jolla, California

Garth van der Kamp
National Water Research Institute
Saskatoon, Saskatchewan, Canada

James Wang
Department of Civil and
Environmental Engineering
Northeastern University
Boston, Massachusetts

Brian Wood
Department of Environmental
Engineering
Oregon State University
Corvallis, Oregon

Wouter Zijl
Geological Survey of the
Netherlands
Netherlands Organisation for
Applied Scientific Research
(TNO)
Utrecht, The Netherlands

Brendan A. Zinn
U.S. Geological Survey
Reston, Virginia

Acknowledgments

I am grateful to this excellent group of authors, who are leaders in their respective specialties. Seventy-one authors from eight countries contributed their knowledge to this handbook. Their names and affiliations appear in the headings of the individual chapters and in the List of Contributors.

The chapters were reviewed in a manner similar to papers in scientific and professional journals. I am also grateful to 29 reviewers from five countries. They made a special effort to make certain that the information presented is sound. Some reviewers read two chapters. Their names appear in the List of Reviewers.

I wish to acknowledge the support I received from the School of Civil Engineering at Purdue University, in particular the encouragement of Dr. Fred Mannering, former head of the School, and Dr. Katherine Banks, current interim head. I also wish to acknowledge the encouragements of my colleagues Dr. John H. Cushman of the departments of earth and atmospheric sciences and mathematics and Dr. G.S. Govindaraju of the hydraulics/hydrology group. I also wish to acknowledge the valuable assistance of the CRC Press/Taylor & Francis Books staff, in particular Mr. Matthew Lamoreaux, editor, civil and environmental engineering and Ms. Jessica Vakili, project coordinator.

Jacques W. Delleur
Editor

Contents

1	History of Groundwater Hydrology <i>Jacobus J. de Vries</i>	1-1
2	Geological Occurrence of Groundwater <i>Darrell I. Leap</i>	2-1
3	Elementary Groundwater Flow and Transport Processes <i>Jacques W. Delleur</i>	3-1
4	Two- and Three-Dimensional Flow of Groundwater <i>F. De Smedt</i>	4-1
5	Groundwater and Seepage: Accounting for Variability <i>Milton E. Harr</i>	5-1
6	Soil Properties and Moisture Movement in the Unsaturated Zone <i>Randel Haverkamp, Samuel Debionne, Pierre Viallet, Rafael Angulo-Jaramillo, and Devaraj de Condappa</i>	6-1
7	Infiltration <i>Jan W. Hopmans, Jean-Yves Parlange, and Shmuel Assouline</i>	7-1
8	Infiltration and Run-On under Spatially Variable Hydrologic Properties <i>R.S. Govindaraju, N. Nahar, C. Corradini, and R. Morbidelli</i>	8-1
9	Modeling the Movement of Water and Solute through Preferential Flow Paths and Fractures <i>Rony Wallach, Tammo S. Steenhuis, and Jean-Yves Parlange</i>	9-1
10	Well Hydraulics and Aquifer Tests <i>Hans Boonstra and Richard Soppe</i>	10-1
11	Well Design and Construction <i>Hans Boonstra and Richard Soppe</i>	11-1

12	Sea Water Intrusion into Coastal Aquifers <i>Jacob Bear and Quanlin Zhou</i>	12-1
13	Groundwater and Heat Flow <i>Darrell I. Leap</i>	13-1
14	Hydrogeological Characterization Using Geophysical Methods <i>Susan S. Hubbard and Yoram Rubin</i>	14-1
15	Geophysical and Tracer Characterization Methods <i>David W. Hyndman</i>	15-1
16	Geostatistics: Interpolation and Inverse Problems <i>Peter K. Kitanidis</i>	16-1
17	Groundwater Contaminants <i>Ernest R. Blatchley, III and John E. Thompson</i>	17-1
18	Nonreactive Contaminant Transport in the Saturated Zone <i>Natalie Kleinfelter and John H. Cushman</i>	18-1
19	Reactive Contaminant Transport in the Saturated Zone: Review of Some Upscaling Approaches <i>J.H. Cushman and T.R. Ginn</i>	19-1
20	Groundwater Flow and Solute Transport in Fractured Media <i>K.S. Novakowski and E.A. Sudicky</i>	20-1
21	Groundwater Flow in Karstic Aquifers <i>William B. White</i>	21-1
22	Contaminant Transport in the Unsaturated Zone Theory and Modeling <i>Jiří Šimůnek and Martinus Th. van Genuchten</i>	22-1
23	Groundwater Modeling <i>Leonard F. Konikow, Thomas E. Reilly, Paul M. Barlow, and Clifford I. Voss</i>	23-1
24	Groundwater Model Validation <i>Ahmed E. Hassan</i>	24-1
25	Scale Issues <i>Keith Loague and Dennis L. Corwin</i>	25-1
26	Accounting for Aquifer Heterogeneity in Solute Transport Modeling: A Case Study from the Macrodispersion Experiment (MADE) Site in Columbus, Mississippi <i>Chunmiao Zheng</i>	26-1
27	Determining Sustainable Groundwater Development <i>John D. Bredehoeft</i>	27-1
28	The Impact of Climate Change on Groundwater <i>Mikko I. Jyrkama and Jon F. Sykes</i>	28-1

29	Ecohydrology: Water, Carbon, and Nutrient Cycling in the Soil–Plant Atmosphere Continuum <i>Edoardo Daly, Gabriel Katul, and Amilcare Porporato</i>	29-1
30	The Role of Geographical Information Systems in Groundwater Engineering <i>Bernard Engel, Kyoung Jae Lim, and Kumar C.S. Navulur</i>	30-1
31	Biodegradation <i>Loring F. Nies and Chad T. Jafvert</i>	31-1
32	Legal Framework for Groundwater Protection in the United States <i>John A. Veil, Deborah Elcock, Nancy L. Ranek, and Markus G. Puder</i>	32-1
33	Landfills <i>Beth A. Keister and Pedro C. Repetto</i>	33-1
34	Evapotranspirative Cover Systems for Waste Containment <i>Jorge G. Zornberg and John S. McCartney</i>	34-1
35	Groundwater Monitoring <i>Michael F. Houlihan and Paul J. Botek</i>	35-1
36	Remediation of Contaminated Groundwater <i>Michael F. Houlihan and Michael H. Berman</i>	36-1
37	Geosynthetics <i>Jorge G. Zornberg and Barry R. Christopher</i>	37-1
	Index	I-1

1

History of Groundwater Hydrology

1.1	Introduction.....	1-1
1.2	Water Management and Water Science before the 19th Century	1-2
	Water Management in Ancient Societies • Techniques and Concepts in the Hellenic-Roman Civilization • Evolution of Water Science during the Renaissance • Foundation of Hydraulics and Hydrology	
1.3	Developments in Subsurface Hydrology	1-6
	Emergence of Basic Understanding of Subsurface Water Behavior • Aquifer Exploration • Advanced Developments in Groundwater Hydraulics • Integrated Analysis of Regional Groundwater Systems • Evolution of Hydrochemistry and Isotope Hydrology	
1.4	Main Trends since the Early 1980s.....	1-28
	Acknowledgments.....	1-29
	References	1-29
	Further Information	1-39

Jacobus J. de Vries
Vrije Universiteit Amsterdam

1.1 Introduction

Although water is a renewable resource and we use little more than 10% of the total precipitation surplus for public water-supply, irrigation, and industrial processes, its availability is restricted through an uneven distribution, both in time and space. In this respect, there is no essential difference between ancient times and the present day; society has always experienced problems with water: too little, too much, too variable, too polluted. Over more than 6000 years mankind has tried to manage these water problems: by intervening in its natural courses through redistribution, storage, and regulation, to accommodate their requirements for irrigation, drainage, flood protection, drinking water, sanitation, and power generation.

Management and exploitation resulted in systematic knowledge of behavior of water under more or less controlled conditions. Gradually, from the Renaissance onwards, this empirical knowledge merged

with scientific developments into a basic understanding of the hydrological cycle and the mechanics of flowing water.

Scientific developments in subsurface hydrology emerged in the early 19th century in western Europe, particularly in connection with the search for unpolluted drinking water for the growing cities, and in relation to land drainage and dewatering of excavations for building or for mining. The center of progress subsequently shifted to the United States at the beginning of the 20th century, notably because of the importance of groundwater for the development of the semi-arid mid-western areas. The focal point of developments remained with North America, but merged with advances elsewhere from the mid-20th century, due to the expansion of international communication and dissemination of knowledge. Several parallel, but more or less isolated developments, had taken place in different countries prior to the mid-20th century. The results often remained unnoticed internationally because communication in applied science and engineering was often restricted to national journals.

The aim of this chapter is (1) to present a brief eclectic introduction to the history of water management and hydraulics before the 19th century, based on secondary literature, and (2) to give a more detailed overview of the advancement of scientific knowledge of groundwater hydrology until the 1980s, based on primary sources. The present overview is necessarily restricted to the mainstream of developments that has become known through widely accessible publications, but a few lesser known discoveries from the Netherlands are also included, illustrating regional achievements in connection with specific hydrogeologic conditions and needs. Many key contributions have been referenced, either because of their ground breaking nature, or to exemplify current developments; though the account does not aspire to being totally comprehensive. For further explanation of the theoretical concepts in this chapter, the reader is particularly referred to Chapter 2, Chapter 3, and Chapter 4 of this volume.

1.2 Water Management and Water Science before the 19th Century

1.2.1 Water Management in Ancient Societies

The oldest civilizations emerged predominantly in relation to the use of river water for irrigating crops under (semi-)arid climatic conditions. Most probably this process began with small-scale irrigation units, regulated by autonomous communities and gradually extended over larger areas with expanding feeder, distribution, and drainage systems. Control of these schemes included technical aspects such as channel and levee construction, as well as administration related to maintenance, taxation, and legislation. The technicians involved in these waterworks made use of rational empirical knowledge and can be considered as professional hydraulic engineers, who mastered the skills of writing and the application of basic mathematics.

All these so-called “hydraulic civilisations,” which developed from the 4th millennium B.C. (notably along the Nile, the Euphrates and Tigris, the Indus, and the Hwang-Ho as well as the Inca civilization in the downstream areas of the rivers at the western slopes of the Andes mountains), showed essentially the same evolution with a strong bureaucratic and stable government, concentrated in urban centers, and characterized by excess production of staple food. This food surplus allowed the development and diversification of non-directly productive professions like priests, architects, surveyors, bookkeepers, astronomers, and warriors. The large population allowed for expansion by trade, craft, and combat, and led to unification of regional power to large kingdoms or empires (cf. Smith, 1976; Back, 1981).

The oldest large-scale irrigation and drainage works began about 4000 B.C. with the Sumerians, between the Euphrates and Tigris. Lifting devices were until Hellenic times (ca. 500 B.C.) restricted to the chain of pots, drawn over a pulley by man or animal, and by the *shaduf* — a lowering and rising bucket on a counter-balanced pole. The most remarkable hydraulic structures that originated in pre-Hellenic times are the *qanats*; underground galleries fed by a spring, or one or more mother wells, dug in an aquifer and connected to the surface by vertical servicing and ventilation shafts at about 50-m intervals. They

originated in the Middle East early in the 1st millennium B.C., and spread rapidly over North Africa, India, and China. Some are tens of kilometers long, collect their water from alluvial fans or springs at the foot of a hill-slope and carry it, without being subjected to evaporation, to the dry plains. Tens of thousands of qanat systems are still in use, particularly in Iran where they drain the large alluvial cones that fill the valleys of the arid highlands. Some reach depths of more than 50 m (cf. Issar, 1990, p. 157).

Accounts of dug wells are found in ancient literature and records from all over the arid regions of Asia and the Mediterranean. Depths seldom exceed 50 m, but their diameters are sometimes such as to allow spiralling paths for donkeys. Core drilling is known from around 3000 B.C. in Egypt and has also been practised for thousands of years in China. Depths of more than 1000 m are reported from the latter country, where churn drilling with bamboo casing was used. These small diameter wells were, except under artesian conditions, not suitable for groundwater extraction in the absence of pumping techniques; they seem to have been used mainly in the mining of minerals, brines, and gas. Artesian wells, among others, are known from oases in the western desert of Egypt where, from 2000 B.C., many shafts and wells reached depths of 100 m (Tolman, 1937, p. 12).

One of the oldest wells discovered in the western world was found in 1991 near the German village of Erkelenz, close to the German–Dutch border. It is a square dug well with a depth of 15 m in a wooden casing of oak. Dendro-chronological and C-14 dating of the wooden framework revealed its age to be about 5300 B.C.; it is the oldest wooden structure known to date (TNO-NITG, 2004). Well-drilling in Western Europe began in the 11th century after the discovery of flowing wells in Flanders by Carthusian monks from a convent near Lille. Several wells were subsequently drilled to depths of almost 100 m by percussion techniques and were used to drive a mill. The wells tapped water from fractured chalk, which outcrops and is recharged in higher plateaus of the Province of Artois. These flowing wells were called artesian wells after the name of the region (Tolman, 1937, p. 316).

1.2.2 Techniques and Concepts in the Hellenic-Roman Civilization

The Roman Civilization became famous for its development of hydraulic construction technology. Most of their ideas, however, were of an innovative character and adapted from well-established technologies from surrounding societies. Architecture and the use of water resistant concrete (cement of clay and lime, made water tight by volcanic ash) were taken mainly from the Greeks, whereas their famous semi-circular arch in bridge-building was of Etruscan origin.

The most remarkable Roman achievements were their water-supply systems, with aqueducts bringing water from outside the cities into complex distribution networks within the cities, similar to the modern public water supply schemes. The Romans made use of long arcades or siphons to bridge deep valleys, constructed tanks and water towers to feed the city's distribution system, and used lead pipes and copper valves for pressurised flow. Rome was probably the first city to enjoy a closed sewer, as early as 500 B.C. Water-raising machines were clearly adopted from the Persians and the Greeks. The *Archimedean screw* (dating from about 250 B.C.) was widely applied in Roman times for irrigation, land drainage, and dewatering of mines. Larger quantities of water could be lifted by the *noria* (also called *Persian wheel*), a large wheel with pots or flow-boards, driven by the stream and thus lifting water under its own power. In case of a well, water could be raised by a chain-of-pots, powered by man or animal via a horizontal, right-angled gear with toothed wheels (cf. Smith, 1976).

Hellenic science was certainly inherited from the older civilizations in the Middle East, but unique was their initiation of intellectual thinking and inquiry, purely as a game. Several philosophers developed speculative theories about the origin and interconnection of elements of the hydrological cycle, particularly the origin of springs and river flow. Their knowledge, however, remained fragmentary because of a lack of adequate information and interest in experiments and systematic observations. The concepts of the Greeks were dominated by the idea that rivers and springs were recharged by way of subterranean channels from the sea. The driving forces were sought in processes such as pressure differences exerted by wind or by rock weight, and suction through vacuum. Desalination was attributed to filtration and distillation. There were also theories about subterranean evaporation and subsequent condensation as well as condensation

from atmospheric water; Aristotle (385–322 B.C.), in particular, propagated this so-called *condensation theory*. Closer to the concept of the hydrological cycle were the theories proposed by Xenophanes of Colophon (about 570–475 B.C.), who associated evaporation from the oceans with clouds, rain, and the origin of springs and streams (Brutsaert, 2005, p. 561). In general, however, rainfall alone was considered quite insufficient to feed rivers and its ability to replenish groundwater and springs was doubted. This is understandable in an arid environment where rainwater seldom seems to penetrate more than a few decimeters before it evaporates. Seneca (3 B.C.–A.D. 65) declared: “As a diligent digger among my vines I can affirm my observation that no rain is ever so heavy as to wet the ground to a depth of more than ten feet” (Biswas, 1970, p. 97; Brutsaert, 2005, p. 567).

A still unanswered interesting question is whether the Romans possessed the conceptual insight to combine flow velocity and cross-sectional area to arrive at the flow quantity per time unit, and if they had any knowledge of pipe and open channel hydraulics to connect hydraulic gradient, friction, and cross-sectional area to the water flux. The accounts of the famous Roman water-supply engineer Frontinus (A.D. 1st century), do not reveal a comprehensive understanding of these hydraulic interactions. Although he seemed to be aware of the continuity principle, he nowhere mentioned the use of flow velocities, but based his flow considerations on cross-sectional area. He possibly used standard channel slope and roughness, so that the cross-sectional area was indeed indicative for the water fluxes, although he seemed to be vaguely aware of the interrelationship between velocity and roughness. Nevertheless, there is an explicit indication that Hero of Alexandria (around 100 B.C.) recognized that flow quantity depends on area and velocity (cf. Biswas, 1970; Dooge, 2004).

Roman concepts of the hydrological cycle did not differ much from speculations by the Greek philosophers, as can be read in Marcus Vitruvius' famous books on Roman architecture and technology (ca. 27 B.C.). However, Vitruvius did mention existing ideas about rainfall being the origin of springs, at least at the foot of mountains (cf. Biswas, 1970, p. 89).

1.2.3 Evolution of Water Science during the Renaissance

Developments in water technology in Medieval Europe were mainly focused on the improvement of power generation by water mills to accommodate the emerging industry, as well as on the reclamation of fertile, wet, alluvial soils. An important innovation in the Low Countries in the 15th century was development of the wind-powered mill to lift water for land drainage, and for reclamation of land from lakes and sea embayments. These ingenious technological developments were not based on theoretical insight, but did encourage the adaptation of the scientific traditions of antiquity, which entered Europe notably via contact with the Moslem and Byzantine civilizations in the Early Renaissance.

Scientific thinking related to the dynamics of water began in the Early Renaissance, notably with the genius Leonardo da Vinci (1452–1519), who was probably the first to arrive at the determination of stream discharge by multiplying the flow velocity by the stream's wetted cross-sectional area. He was surely the first to clearly formulate the law of continuity, with his statement: “Each part of the river gives in equal time the same quantity, whatever the slope, width, depth or roughness.”

The dawn of modern scientific thinking in hydraulics in the 17th century is exemplified by Benedetto Castelli (1578–1643), who studied under Galileo Galilei (1564–1642) in Padua and was Professor of mathematics in Rome. He presented a theoretical analysis of flowing water, and extended the principle of continuity for nonuniform flow. Domenico Guglielmini (1655–1710) subsequently connected a constant flow velocity on a sloping surface with a balance between the gravitational force and resistance. This would eventually lead, in the 18th century, to a general formulation of the relation between friction, slope, and velocity (Dooge, 2004).

Quantitative hydrology began with the lawyer Pierre Perrault (1608–1680) and the scientist Edmé Marriotte (1620–1684), who definitely proved rainfall to be more than enough to support the flow of springs and rivers by comparing the results of their discharge measurements in the Seine basin with the amount of precipitation on the catchment. Perrault, however, believed that surface runoff was the source

of river flow and that springs were mainly fed by streams. He probably considered infiltration to be just sufficient to keep groundwater in balance with evaporation.

A better understanding of the process of infiltration and subsequent discharge of groundwater by springs and rivers was earlier proposed by craftsman (inventor of enamelled pottery) and amateur naturalist, Bernard Palissy (1510–1590), on the basis of sound observations and deductions. He did not master Greek and Latin, and was therefore not hindered by the ideas of the old philosophers; his scientific attitude is evident from his statement: “I have no other books than Heaven and Earth, which are open to all.” Palissy had also assumed rainfall to be sufficient to feed the rivers, in a century that great scientists like Johannes Kepler (1571–1630) and René Descartes (1596–1650) still adhered to the concepts of the Greek philosophers.

The picture of the hydrological cycle was finally completed with the evaporation experiments by the astronomer Edmond Halley (1656–1742), who assessed the water balance of the Mediterranean sea by up scaling evaporation measurements in cylinders. He concluded the evaporation on the oceans to be more than enough to generate the necessary rainfall to feed the rivers (Brutsaert, 2005, p. 583).

1.2.4 Foundation of Hydraulics and Hydrology

The 18th century saw the various hydraulic concepts gradually merging toward a comprehensive theory of hydrodynamics based on Newtonian mechanics. Henry de Pitot (1695–1771), a French hydraulic engineer, studied the distribution velocity within streams and was not satisfied with the use of floats. He therefore developed his well-known L-shaped tube to determine the rise of the water head above stream surface as a measure for the velocity, making use of the principle that the velocity head under ideal conditions must equal the pressure head. The same principle was the subject of a more theoretical treatise on fluid flow published in 1738 by the Swiss physicist and mathematician, Daniel Bernoulli (1700–1782). His equation of the pressure–velocity relation in permanent flow without friction was subsequently extended by a potential energy term, and is now known in that form as the Bernoulli equation for conservation of energy. From this equilibrium principle, Daniel Bernoulli and his father, Johann Bernoulli, both arrived in 1738 at the basic hydraulics equation: $v^2 = 2gh$, where v is velocity, g is acceleration due to gravity, and h is hydraulic head.

Based on the same principles, Antoine Chézy (1718–1798), a professor at the École des Ponts et Chaussées, developed his well-known stream flow equation. He argued that velocity due to gravity is only uniform if the acceleration due to gravity equals the resistance caused by the wetted perimeter. In this approach, that is still widely used, the resistance due to friction along the stream bed, as well as the internal friction, has to be determined empirically. Many attempts were made to obtain resistance coefficients based on velocity distribution and channel geometry, including the Darcy–Bazin experiments on open channel flow and pipe flow, and the stream channel experiments by the Irishman Robert Manning (1816–1897).

Henry Philibert Gaspard Darcy (1803–1858) was a mechanical and hydraulic engineer who was, among others, concerned with water-supply systems. His treatise on the successful design and construction of Dijon’s public water supply, published in 1856, contains the results of his famous experiments on groundwater percolation through sand filters. From this study he derived his well-known linear law of groundwater flux and hydraulic gradient, which became the foundation of quantitative groundwater hydrology. This law of groundwater flow through porous media shows similarity with the Hagen–Poisseuille equation for laminar viscous flow through small-diameter pipes, developed earlier by the French physician Jean Louis Poisseuille (1799–1869) and the German hydraulic engineer Gotthilf Heinrich Ludwig Hagen (1797–1884). Darcy, who had a sound theoretical background, was certainly aware of this analogy (cf. Biswas, 1970).

Other important contributions to water science at the beginning of the 19th century were discovery of the composition of the water molecule, H_2O , by Jöns Jacob Berzelius (1779–1848), and the fundamental law of evaporation by diffusive and turbulent transport of vapor, formulated in 1804 by the founder of atomic chemistry, John Dalton (1766–1844).

1.3 Developments in Subsurface Hydrology

1.3.1 Emergence of Basic Understanding of Subsurface Water Behavior

1.3.1.1 Groundwater Dynamics

Interest in subsurface water origin and movement has a long history, notably in arid areas where it was often the only source of drinking water. In humid areas curiosity arose especially in connection with the intriguing, not to say mysterious, phenomena of springs and artesian wells, which were usually supposed to be connected to powerful subterranean water arteries. A more scientific approach to groundwater hydrology was possible only after the fundamentals of geology were established at the end of the 18th century. These principles provided the essential framework of the properties and structure of the rock strata, including the role of fractures, joints, solution cavities and weathering horizons, and the perception of water-bearing and water-confining layers. A well-known pioneer in connecting rock properties with groundwater behavior was the Englishman William Smith (1769–1839), who, in the early 19th century, applied his knowledge to improve the supply of groundwater for the town of Scarborough (Biswas, 1970, p. 306). However, it was discovered recently (Stephens and Ankeny, 2004) that Smith was preceded by Joseph Elkington (1739–1806), a farmer who developed his knowledge by drainage work on his swampy lands. He showed remarkable understanding of the control of geologic structure and topography on groundwater circulation, including the location and discharge of springs, artesian pressure, and perched aquifers. The successful and widespread application of his knowledge on horizontal as well as vertical drainage with auger holes for pressure relief, in the second half of the 18th century, made him (according to Stephens and Ankeny), perhaps the first professional consulting hydrogeologist.

Interest in groundwater as a source of drinking water emerged in the western world at the beginning of the 19th century particularly because of the increasing pollution of surface water caused by the fast growing population and industrial activities in many cities in Europe and the United States. Free flowing artesian wells were notably considered as a potential source of drinking water, particularly after a number of successful drillings in France. The French scientist and politician François Arago came to the conclusion, on the basis of geologic considerations, that artesian water could be encountered in the water-bearing Albian Greensands of Cretaceous age; these he believed to extend under Paris below Tertiary confining clay layers. On his instigation drilling started in 1831 at the Grenelle location and reached the Greensands in 1841 at a depth of 548 m. This was much deeper and much later than expected by Arago, and taking into consideration his rather simple percussion drilling tools and man-powered thread-wheels, one can imagine the immense technical and mental difficulties that Arago and his engineer Murot faced. The artesian pressure spouted water to a height of 33 m above surface and initially produced 4000 m³/d, which was a substantial boost to the Paris water supply (cf. Smith, 1976, p. 108).

The explanation of artesian pressure produced by an elevated groundwater table in aquifer outcrops had already been proposed by the Italian Bernardino Ramazzini (1633–1714), in his 1691 treatise on the artesian wells of Modena; this contained ideas which he, in turn, seemed to have obtained from the Arabian philosopher Al Biruni (973–1048). Ramazzini already understood the meaning of a hydraulic gradient between an artesian well and the pressure source in the case of flowing groundwater (cf. Biswas, 1970, p. 190). A comprehensive treatise on artesian wells was produced by the French engineer Garnier (1822). Early concepts relating to artesian water supposed rigid rock conditions, but Arago also mentioned (according to Versluys, 1931) existing ideas of artesian pressure generated by downward bending confining layers, which were deformed by overburden loading. However, he finally adhered to the explanation of hydraulic pressure propagation from an elevated outcrop.

A quantitative approach to groundwater hydrology began with the previously mentioned work of the French engineer Henry Darcy (Figure 1.1). In the framework of his ingenious work for the Dijon central drinking water supply system, he carried out percolation experiments using purification filters as depicted in Figure 1.2 (see also Chapter 3, Figure 3.5 and Figure 3.6, this volume). From these investigations, he derived his well-known law for the linear relation between hydraulic gradient and groundwater flux, which he published as an appendix to his monumental treatise on the water-supply scheme (Darcy, 1856). This



FIGURE 1.1 Henry Philibert Gaspard Darcy (1803–1858). Portrait by Perrodin, Collection of the Bibliothèque Municipale de Dijon.

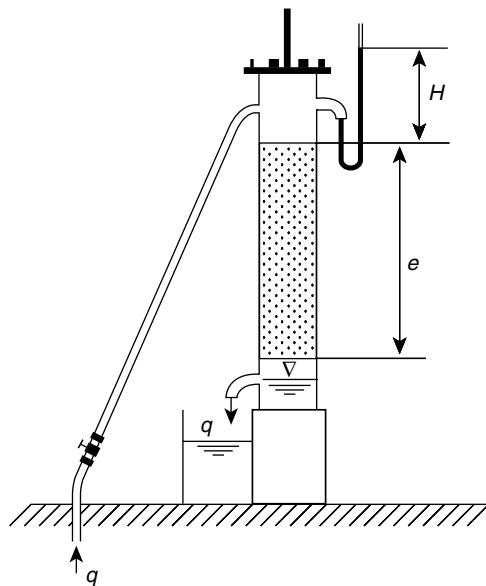


FIGURE 1.2 Principle of Darcy's percolation experiment with outflow under atmospheric (=zero) pressure. The results are formulated in Equation 1.1 for the case that the lower boundary of the column is chosen as reference elevation level, so that elevation head and pressure head are both zero at the outflow surface.

work not only revealed Darcy's technical skills, but also accounted for his thorough understanding of the role of topography and geology for the groundwater conditions and their meaning for his water-supply scheme (cf. Brown et al., 2003).

For percolation through a cylinder with outflow under atmospheric pressure, he derived his well-known formula:

$$q = Ks(H + e)/e \quad (1.1)$$

where q = percolation flux [$L^3 T^{-1}$], K = permeability factor [LT^{-1}], s = surface area of the sand filled cylinder [L^2], e = length of the sand column [L], and H = hydraulic pressure at the upper boundary of the sand column [L]. $(H + e)$ = hydraulic head and $(H + e)/e$ = hydraulic gradient (Figure 1.2).

Darcy also derived the non-steady flow equation for the condition of a falling pressure head:

$$q_t = q_0 \exp(-Kt/e) \quad (1.2)$$

where t = time [T] and q_0 = percolation flux [$L^3 T^{-1}$] at $t = 0$.

This non-steady groundwater flow equation was later applied and further expanded by Edmond Maillet (1905) to analyze the dry weather discharge recession of rivers and springs. Darcy certainly realised that his experiments were carried out under ideal conditions with rather homogeneous porous media, which may differ substantially from actual field conditions. However, he had previously studied the changing discharge of artesian wells as a function of their changing water pressure; these findings complied with his linear law and thus suggested its universal validity for groundwater flow through porous media (Brown, 2002).

Many investigators around the turn of the century repeated Darcy's experiments and confirmed their applicability for conditions of non-turbulent flow (*inter alia*: Seelheim, 1880; Hazen, 1893; King, 1899; Slichter, 1905). All these studies included the establishment of semi-empirical relationships between the mechanical composition of sediment and its permeability, in an attempt to find a physical expression for the Darcian permeability factor on the basis of the Hagen–Poisseuille theory for viscous flow in pipes. This resulted in a characterization of permeability as a quadratic function of the pore size or its related grain size. Well-known was Hazen's *effective grain-size* D_{10} , which represents the size which is finer than 90% of the grains, and his study on the influence of the particle distribution uniformity, represented by the D_{10}/D_{60} ratio. More elaborate experiments on permeability and sediment properties include those of Zunker (1923), who introduced the concept of the *specific surface*. Related formulae were proposed by Terzaghi (1925), Kozeny (1927), and Carman (1937, 1956), leading to the well-known Kozeny–Carman equation. Graton and Fraser (1935) included the influence of the arrangement of the particles on porosity.

Darcy's law allowed groundwater flow problems to be formulated in mathematical models, and to solve these for given boundary conditions. In this way, the French engineer Jules Dupuit (1804–1866) derived his famous solution for steady horizontal flow to a well at the center of a circular island in an unconfined aquifer (Dupuit, 1863, Figure 1.3):

$$\pi K(h_R^2 - h_r^2) = Q \ln(R/r) \quad (1.3)$$

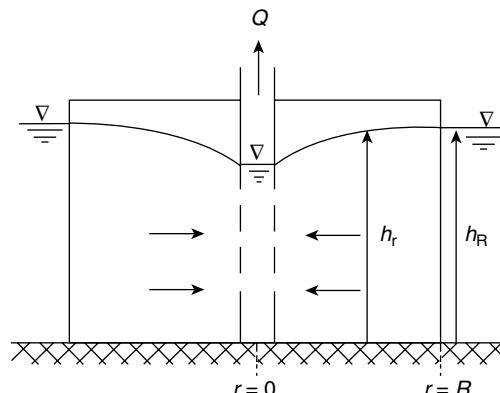


FIGURE 1.3 Horizontal, radial-symmetric groundwater flow to a well in the center of a circular island with radius R ; h = hydraulic head relative to the impervious base ("Dupuit assumptions").

where h_R and h_r are the hydraulic head relative to the base of the aquifer, at distance R and r from the well, respectively; R is the island's radius; Q is the groundwater extraction, replenished from the constant-head boundary at distance R . Dupuit's simplification of horizontal flow under a sloping water table is now generally known as the "Dupuit Assumption."

The German Adolph Thiem (1836–1908) developed (probably independent of Dupuit) identical formulae for unconfined and confined well flow, which he extensively applied in projects on groundwater extraction for public water supply for several large cities in Europe (Thiem, 1870). His son Günther Thiem (1906) was probably the first to explicitly apply the well flow formula in test-pumping experiments, with observation wells to determine aquifer properties. He systematically evaluated the results of his field observations under different boundary conditions, and particularly considered the relevance and physical meaning of the concept of the "radius of influence R " for the usual conditions where the well is not in the center of a circular island.

It is evident that to use Darcy's formula for engineering problems that were dominated by horizontal, vertical, or radial-symmetric flow did not pose too many difficulties, but application to more complicated, two-dimensional (2D) field conditions was hampered by a lack of insight to the actual flow field and the associated hydraulic head distribution. This is exemplified by discussion in the Netherlands in the 1890s in connection with the use of groundwater from the coastal dune area for drinking water. In 1852, Amsterdam began groundwater extraction from the dunes using a system of drainage canals. A scientific-technical committee which investigated the efficiency of the system, applied in their report the parallel flow version of Equation (1.3) for discharge to a channel fed by uniform infiltration. They assumed horizontal flow with the bottom of the canal forming the lower boundary of the flow system, and arrived at a parabolic drainage formula, which is also known from earlier work by the Danish engineer Ludvig Colding (cf. Section 1.3.4.3). In the committee's notation:

$$qx^2/kcb^2 + y^2/b^2 = 1 \quad (1.4)$$

where q = discharge = recharge [LT^{-1}]; k = quotient of hydraulic conductivity and porosity c [LT^{-1}]; y = groundwater level (hydraulic head) relative to the bottom of the channel [L]; $y = b$ at the water divide, where $x = 0$.

The assumption that the drainage base must constitute the lower flow boundary in an aquifer that extended further downward was in accordance with their idea that groundwater flow, under water table conditions, could not move in an upward direction (cf. De Vries, 2004). This doctrine was derived from the authoritative book by the German water-supply engineer, professor Otto Lueger (1883).

However, subsequent measurements of drainage water yield in the canals by the Director of the Amsterdam waterworks, engineer Johan M.K. Pennink (Figure 1.4), resulted in three times the amount that was calculated from Equation 1.4 for horizontal flow. This gave him the idea that the thickness of that part of the subsurface participating in the groundwater discharge was possibly underestimated in the then current theory. He therefore postulated radial converging and upward bending flow lines in a vertical section perpendicular to the channels. Pennink splendidly demonstrated the flow pattern below the channels from both field evidence and experiments with parallel-plate models for viscous flow (Pennink, 1904, 1905). The laboratory models consisted of spaces bounded by parallel glass plates, initially filled with liquids and in later experiments filled with sand. His field experiments were carried out with rows of piezometers to measure the hydraulic head at different distances from the canals and at various depths. From these observations he constructed the spatial pattern of hydraulic head contours and associated flow lines (Figure 1.5), which proved his concept of radial flow. Pennink thus clearly understood that the driving force at a point in the flow field is formed by the gradient of the summed elevation and pressure heads. This application of the concept of hydraulic gradient under field conditions in a 2D flow field was certainly not common knowledge. Pennink in fact solved graphically, more or less intuitively, the continuity equation.

Pennink's experiment also clearly explained the increase in hydraulic head with depth under areas of discharge. This phenomenon, as well as strong up welling of water in excavations at the foot of



FIGURE 1.4 Johan M.K. Pennink (1853–1936). Photograph Amsterdam Water-Supply.

the coastal dunes, had previously been misinterpreted as an indication for the occurrence of artesian water veins beneath the dunes; the artesian water hypothesis was incorrectly used as an alternative for the right explanation of the more than 150 m thick fresh water pocket below the dune area by way of the *Ghijben–Herzberg principle*. This principle of equilibrium between a fresh water pockets floating on surrounding salt groundwater was previously formulated by a captain in the Dutch Army Corps of Engineers, Willem Badon Ghijben (1845–1907). His ideas (Drabbe and Badon Ghijben, 1889) did not receive proper recognition until the German engineer Herzberg (1901) independently reached the same conclusions, after discovering fresh water extending to a depth of 60 m under the North Sea island of Norderney. It should be noted that the idea of fresh groundwater floating on salt water was predicated in the United States by Joseph Du Commun [1818], a French instructor of the West Point Military Academy (cf. Carlson, 1963).

Pennink furthermore produced the rather precise figure of 350 mm/year for the natural recharge of the dunes, and persistently warned for over-exploitation and salinization. He also designed a scheme for artificially recharging the dunes with river water, which was only executed half a century later following serious salinization problems (cf. De Vries, 2004).

The French physicist Valentin Joseph Boussinesq (1842–1929) and the Austrian hydraulic engineer Philipp Forchheimer (1852–1933) were the first to explicitly recognize, from theoretical considerations, that an analytical solution of a groundwater flow problem cannot be solely based on Darcy's law of motion, but should also obey the continuity principle, the combination yielding the Laplace equation. Boussinesq (1877) also noticed that horizontal one-dimensional (1D) transient groundwater flow is mathematically analogous to the transient diffusion equation for heat flow as proposed by Fourier in the early 19th century. This paved the way for applying the mathematical solutions for heat flow problems to equivalent groundwater problems. In 1886, Forchheimer solved the 2D Laplace continuity equation for steady-state, 2D groundwater flow in a homogeneous and isotropic medium using the conformal mapping technique, a method that was previously applied by Holzmüller (1882) for heat flow problems. He arrived at the basic formula:

$$\delta^2 h / \delta x^2 + \delta^2 h / \delta y^2 = 0 \quad (1.5)$$

which he combined with Darcy's flow formula to generalize Dupuit's well flow formula for confined horizontal flow conditions.

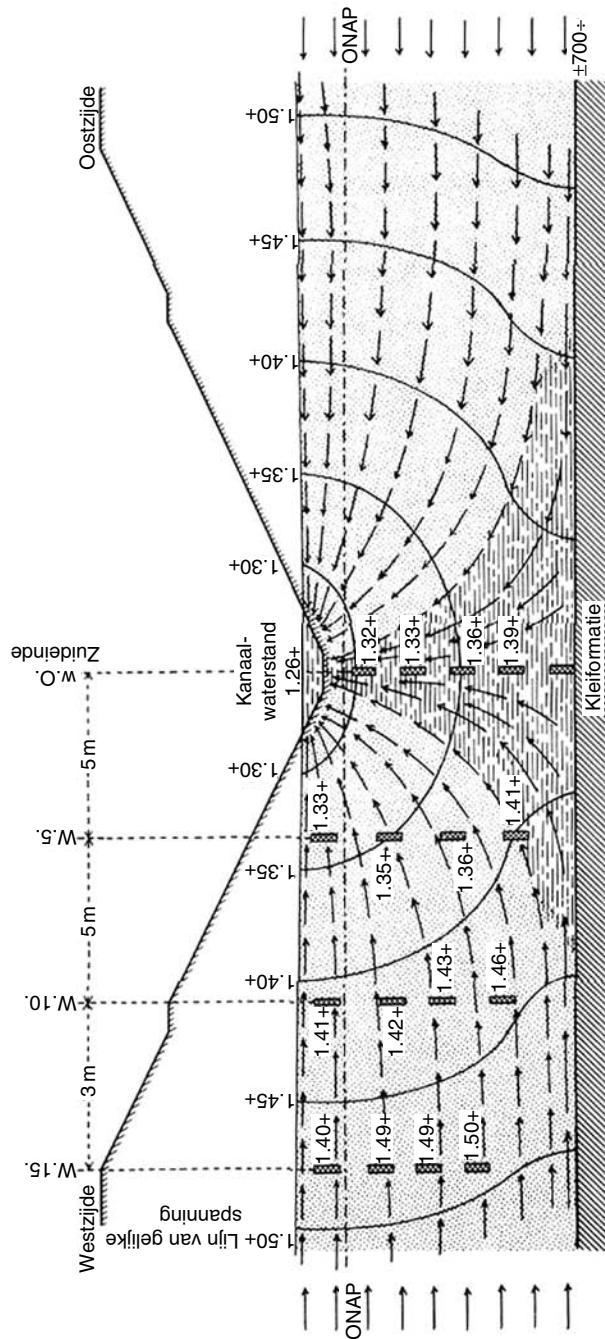


FIGURE 1.5 Flow net around a drainage canal based on hydraulic head observations, according to Pennink, 1905.

Forchheimer (1898) subsequently extended the general differential equation to the 2D unconfined “Dupuit free-surface” problem, for which he arrived at the formula that is now commonly known as the “Dupuit–Forchheimer equation”:

$$\delta^2(h^2)/\delta x^2 + \delta^2(h^2)/\delta y^2 = 0 \quad (1.6)$$

The transient version of Forchheimer’s Laplace equation appeared in the work of Boussinesq (1904). The professor of mathematics and collaborator with the US Geological Survey (USGS), Charles S. Slichter (1864–1946) arrived at the same results for steady-state flow and extended Forchheimer’s approach by including a vertical flow component (Slichter, 1899). Remarkably, he did not mention Forchheimer, although he was obviously familiar with German authors as he referred to Thiem and Lueger. Another associate with the USGS and professor of soil physics, Franklin H. King (1848–1911), in a companion paper, gave a general description of groundwater flow under the influence of a topography-controlled groundwater table and presented a sketch for the flow pattern near a stream which was similar to Pennink’s case (cf. Figure 1.5). King clearly explained that the flow of water beneath a channel or depression is forced upward under hydraulic pressure, driven by the inflow, and accumulation of water percolating downward through the surrounding higher land. In contrast to Pennink, however, King did not verify his ideas by field experiments. Conversely, Pennink was obviously not familiar with the theoretical approach, or at least did not know how to connect the results of his experiments with this theoretical knowledge; his work was mainly based on intuition and observation.

Slichter (1902) applied King’s sketch to discuss the flow pattern in a framework of general theory. He emphasized that it would be misleading to compare groundwater flow with pipe flow and stream flow because, as he stated, the frictional resistance in groundwater is not transmitted by the fluid layers. It is interesting, in this connection, that he considered the influence of an undulating impervious base for the flow pattern in an aquifer: “The contention of some German hydrographers (he then refers to Lueger; deV.) that there can be no motion in a region like ASB (referring to a figure with a concavity ASB in the impervious base of the aquifer; deV.) must be entirely abandoned. Water must circulate in all parts of the enlargements in the porous medium, for the same reason that heat would be conducted over similar enlargements in a conductor.” He further specified: “If it were not for the ever present controlling influence of gravity the motion would be entirely analogous to the flow of heat or electricity in a conductive medium.”

1.3.1.2 Geological Aspects

An extensive general study on the relation between geological structure and the occurrence and flow of groundwater was published in three volumes by Gabriel Daubrée (1887). Another outstanding French contribution was the study on the interaction of aquifers and rivers by Maillet (1905), who analyzed stream-flow recession curves to evaluate and forecast low flow. In this he made use of the principle of linear reservoirs (as represented by Equation 1.2) and the theoretical considerations of Boussinesq relating to non-steady flow, including the aspect of nonlinearity in shallow aquifers.

The misconception of stagnating water below the drainage base (as particularly propagated by Lueger, cf. Section 1.3.1.1) was also reflected in the sub division by Grund (1903) for karst areas, in which he made a distinction between moving karst water above sea level and stagnant groundwater below sea level. Grund considered karst aquifers as continuous systems, much like groundwater bodies in other porous media. This concept was challenged by Katzer (1909) who observed complex behavior, including diffuse flow through a porous medium as well as cave-river circulation. Katzer’s ideas were largely based on the results of extensive exploration by the French speleologist Martel (1894). Classical studies of the interaction between karst morphology and karst water circulation and dissolution, including the origin and evolution of cavities, and swallow holes, were carried out by the Yugoslavian Jovan Cvijić (1893, 1918). His excellent work is still the basis of modern karst hydrology and karst morphology.

The role of CO₂ in limestone dissolution was already understood at the end of the 18th century, but before the 1930s it was generally assumed that CO₂-driven dissolution could only be active in the zone near

the water table. It was notably the famous American geomorphologist Davis (1930) who found convincing arguments for active dissolution below the water table in cavity structures. Unsaturated conditions with respect to CaCO_3 , even at greater depths below the water table, was subsequently convincingly proved (e.g., Back, 1963), and Alfred Bögli (1964) demonstrated the process of mixing corrosion as a powerful means of continuous dissolution (cf. Section 1.3.5.1). Discussion on the possibility of a diffuse groundwater flow with a continuous water table (as propagated, among others, by Grund) vs. the idea of a restricted circulation by independent cave-river networks, continued for quite some time, notably in Europe (cf. Bögli, 1980, Chapter 6). A well-known protagonist of the latter view was Lehmann (1932). However, extensive research in the United States in the 1960s, in the deep carbonate rocks in the South Nevada mountains and intermountain valley systems, convincingly illustrated the ability of karst processes to create continuous aquifer development under complex structures (cf. Bredehoeft et al., 1982). Zötl's (1974) comprehensive text *Karsthydrologie* finally put an end to the debate in Europe with sound arguments in favor of the occurrence of large-scale connected karst water bodies with a coherent piezometric surface.

Nevertheless, karst hydrological systems can be extremely complex in structure and dynamics which often cause very specific behavior; for example, the occurrence of intermittent springs with their rhythmic periodicity. Katzer (1909) was the first to explain this remarkable phenomenon by a system of interconnected storage cavities and siphons.

1.3.1.3 Soil Water Physics

At the turn of the century a lack of insight to the nature and flow of water in the unsaturated zone hampered understanding of the process of infiltration and groundwater recharge. Particularly the observation that often after rainfall only the upper soil seemed to have become wetted, and that even after heavy rainfall no subsurface inflow was observed in depressions above the water table, made even serious scientists doubt that rainfall was the source of groundwater. Time and again the so-called *condensation theory*, as advocated by some philosophers of antiquity (cf. Section 1.2.2), emerged as the alternative source of groundwater production. This hypothesis was further supported by the observation that in confined aquifers the groundwater level sometimes rose before rainfall; a phenomenon that was later attributed to a decline of atmospheric pressure. Moreover, the results of poor lysimeter experiments often suggested evaporation on a drainage basin scale to be larger than the rainfall, so that an additional source of water was required to satisfy a balance. The controversy surrounding this subject was not only of scientific importance; emphasis on condensation as the single source of groundwater also threatened public health. For example, the influential antagonist of the rainwater-infiltration theory in Europe, the German geologist Otto Volger (1877), argued that concern for groundwater contamination by infiltration from polluted soil was a fear of phantoms, which would lead to unnecessary costs with the expensive construction of extraction wells for public water supply outside urban areas(!)

The dawn of understanding of the physics of the unsaturated zone began in the United States with studies by King and Slichter on soil water retention in relation to capillarity in porous material, but a breakthrough was the introduction in 1907 of the concept of capillary potential by Edgar Buckingham (1867–1940), then a physicist with the US Department of Agriculture. He recognized that, apart from gravity, the driving force of water flow in unsaturated soils is the difference in capillary suction, which in turn depends in homogeneous material on a difference in water content. He was also aware of the dependence of capillary conductivity on pore geometry and moisture content. Buckingham's theory can be formulated for the basic, 1D vertical steady-state flow case, by:

$$\nu = K_c(dh/dz \pm 1) \quad (1.7)$$

where ν = soil water flux [LT^{-1}]; K_c = capillary conductivity [LT^{-1}]; h = soil water head (capillary suction head) [L]; z = elevation above datum plane [L]; the gravity component = +1 for infiltration and -1 for capillary rise. The continuity between the flow in the unsaturated and the saturated zone is

evident from the equivalence of Equation 1.1 and Equation 1.7. However, because of the dependency of K_c on the moisture content, the flow in the unsaturated zone is generally strongly nonlinear.

The Australians W.H. Green (1868–1932) and G.A. Ampt (1887–1953) subsequently developed a theory of infiltration based on the Hagen–Poisseuille equation. They linearized Equation 1.7 by assuming vertical steady-state penetration by gravity of a wetted front with a sharp interface with the unsaturated underlying material (which means in Equation 1.7 that $dh/dz = 0$ at the interface). Their solution is still widely used (Green and Ampt, 1911). Interest in capillary processes and soil evaporation in the United States was particularly connected with the practise of dry farming.

Willard Gardner (1883–1964) incorporated Buckingham's capillary potential into a more complete flow theory (Gardner, 1920). This was put in a more definitive form by his student Lorenzo A. Richards (1931), who combined the flow equation with the continuity equation into the general nonlinear flow equation for water in unsaturated soils, now generally known as the "Richards equation." Another significant contribution by Gardner was the development of a porous-cup tensiometer. This instrument allowed systematic study of the soil water retention characteristics, for which Schofield (1935) in the UK introduced the concept of the *pF curve*, after Veihmeyer and Hendrickson (1931) had introduced the related *field capacity* concept, and Haines (1930) had discussed the *hysteresis effect*.

These developments established a basic understanding and mathematical expression of flow processes in the unsaturated zone, although application remained restricted to laboratory and plot-scales because of the strongly spatial heterogeneity within the unsaturated zone, and the complicated transient and nonlinear nature of the flow processes.

1.3.2 Aquifer Exploration

In the period around World War I the origin of groundwater, and the physical basis for the process of infiltration as well as a sound theory for groundwater flow, were well-established. The solution of a groundwater flow problem in idealized and schematized situations was essentially reduced to the mathematical solution of the general groundwater flow equation for specified boundary and initial conditions. Adequate dissemination of the theoretical knowledge in Europe was notably effected by Forchheimer's (1914) textbook *Hydraulik*, which contained a chapter on the theory of groundwater flow. Other more general German language books on hydrology were those of Konrad Keilhack (1912) and Prinz (1919). In the United States the knowledge to date was summarized in the volumes by Adolph F. Meyer (1917) and Daniel W. Mead (1919). In France, Daubrée's work was succeeded by Martel's (1921) thorough text on the geological aspects of groundwater occurrence. These early texts combined descriptive characterization of hydrogeologic conditions with basic mathematical treatment of groundwater flow, and stream- and groundwater hydrograph analyses. International exchange of knowledge was largely stimulated, notably in Europe, by the founding in 1922 of the International Association of Scientific Hydrology (IASH) as subsection of the International Union of Geodesy and Geophysics. Until the establishment of specialized hydrology journals in the 1960s, the IASH conferences and their proceedings remained the most important means of information exchange.

Groundwater hydrology developed as an empirical science by way of interaction between conceptualization and mathematical solution of flow cases on one side, and assessment and observations of the physical characteristics and behavior of aquifers and groundwater basins on the other. This approach was particularly related to more or less controlled engineering conditions such as groundwater extraction by wells for water supply, or dewatering of excavations for building, or for mining purposes and land drainage. This association of hydrogeology with societal-driven problems was particularly successful in the Netherlands, where shallow water tables demand adequate groundwater control, and where the hydrogeological framework is relatively simple (cf. De Vries, 2004). The subsurface consists of unconsolidated, relatively homogeneous Quaternary sediments with artificially controlled groundwater and surface water levels, and well-defined boundary conditions. The western part of the country can be characterized as a coastal lowland, consisting of leaky aquifers formed by a subsurface of more than 150 m

of medium-coarse Pleistocene fluvial sand, which is confined by a semi-pervious layer of up to 15 m of Holocene clay and peat with shallow water tables and abundant surface water.

A first theoretical analysis of groundwater flow to extraction wells and to deep circular polders in these leaky aquifers was presented by Kooper (1914), an officer in the Army Corps of Engineers and employed by the National Bureau for Drinking Water Supply (RID); the Bureau was established in 1913 to support the development of water-supply schemes, notably in rural areas. Kooper arrived at a mathematical expression using Bessel-functions, which proved to be quite adequate for groundwater exploration and management. The solution for this steady-state flow case was subsequently elaborated and made more accessible by Gerrit J. de Glee (1897–1975) in his 1930 Ph.D. thesis; his well flow formula for leaky aquifers is generally referred to as the “De Glee formula,” which is equivalent to that published in the United States by Jacob (1946). Similar formulae were developed in Germany by Schultze (1924), and were extensively applied in the drainage of excavations for hydraulic engineering construction in the coastal lowland of the northern part of that country.

Rather different hydrogeologic conditions and needs for groundwater were encountered in the United States (cf. Rosenshein et al., 1976; Bredehoeft et al., 1982). Interest in groundwater emerged in the last decades of the 19th century in connection with the opening of the semiarid midwestern plains. Water was demanded by the settlers for household use, stock breeding, irrigation and railroads, and the occurrence of groundwater with artesian properties was, in the absence of a suitable power source for mechanical pumps, a favorable condition for development. After the founding of the USGS in 1879, many eminent geologists received the task to investigate the large artesian basins, where the geologic structure has an important control over the movement of groundwater and the character of the aquifers and confining layers. Thomas C. Chamberlin (1885) produced the first studies on artesian basins, based on the assumption that the pressure-producing mechanism is the hydrostatic weight at the outcrop area, the influence of which extends down-dip. Another classic study was the exploration of the Dakota sandstone by Nelson H. Darton (1909). More difficult to visualize in those early days were the flow processes in the phreatic aquifers where hydraulic head is a subdued replica of the topography, and where the complex 2D regional flow patterns between recharge and discharge areas were difficult to understand (cf. Section 1.3.1.1).

A systematic approach to the regional hydrogeology of the United States and the underlying principles emerged under the eminent leadership of Oscar E. Meinzer (1876–1948), who became chief of the USGS Groundwater Branch in 1911 and who successfully integrated geological information with groundwater hydraulics. Exploration was concurrently advanced by the introduction of hydraulic rotary drilling and the deep-well turbine pump, which enabled to lift water from greater depths. Meinzer began with a systematic inventory of all regional information into a complete description of the various hydrogeological provinces and terrains, including a characterization of the porosity and permeability of different rock types. This information was laid down in a textbook that presented an outline of groundwater hydrology with clear, concise definitions; for example, of concepts such as *specific yield* and *specific retention* (Meinzer, 1923a, 1923b).

Meinzer’s contribution to scientific concepts is particularly related to his classical work on the origin of artesian pressure and its connection with rock compressibility. Studies and analyses of the Dakota sandstone artesian aquifer became classic in this respect. Many facts, concepts, and arguments were put together and evaluated to arrive at a sound understanding. A key argument was based on the observation that groundwater production in excess of replenishment resulted in a decline of both yield and artesian pressure in the vicinity of the well, but did not cause a lowering of the pressure at the recharge area. Since the confined aquifer remained full of water all the time, this could only be explained by the delivery of water from an elastic change in the aquifer’s volume by compression, through reduction in the buoyant force (Meinzer and Hard, 1925; Meinzer, 1928). Other arguments in favor of the elastic nature of aquifers were derived from observations on tidal and earthquake effects on well levels, water level reaction to an increase in overburden load caused by passing trains, slow expansion and recovery of depression cones around wells, and land subsidence near wells. Subsequent analysis of the different reactions of rigid and compressible aquifers to changes in atmospheric pressure helped to obtain insight on the nature of

compressibility (cf. Meinzer, 1942). Meinzer can, because of his unifying work, be considered as the architect of modern hydrogeology.

According to Tolman (1937, Chapter XI), discussion arose on the question whether artesian pressure could be produced solely by the load of the confining overburden, or if a combination with hydraulic pressure from the intake was a prerequisite. Meinzer was in favor of the latter view, but others argued that the Dakota sandstone was not continuous from the outcrop and showed an irregular pressure distribution (e.g., Russel, 1928). Discussions followed on the question of whether the yield from storage was from the aquifer itself, or from the confining layers or from both. Further insight to the processes of compressibility and consolidation were advanced on the basis of the classic work of (Czech-born) Karl Terzaghi (1883–1963) on soil mechanics, with his concept of *effective stress* as the total lithostatic pressure minus the hydraulic (buoyancy) pressure (Terzaghi, 1925, 1929).

Terzaghi's studies on land subsidence caused by groundwater extraction and subsequent compression and compaction of the aquifer and connected water-saturated clays, including the process of consolidation, were further advanced by Skempton (1944) and Terzaghi and Peck (1948). Analysis of the geological process of deformation in a three-dimensional (3D) stress field was pioneered by Biot (1941) and Geertsema (1957), and further extended by many others (e.g., Bredehoeft and Hanshaw, 1968). Myron L. Fuller of the USGS was probably the first in the United States to speculate on the relation between groundwater extraction and land subsidence (cf. Helm, 1982). Studies on such phenomena have been carried out in the United States for decades under the leadership of Joseph F. Poland, who established a research center in Sacramento, California (e.g., Poland, 1961; Poland and Davis, 1969). Another process of land subsidence took place in the Netherlands where drainage of coastal marshland has resulted in a lowering of the surface of more than 5 m during the last 800 years.

1.3.3 Advanced Developments in Groundwater Hydraulics

1.3.3.1 Flow Analysis and Aquifer Tests

Prior to the mid-1930s pumping tests were predominantly based on the Thiem-formula for steady-state radial flow from a circular fixed head boundary at the so-called distance of influence (cf. Figure 1.3). In response to needs of engineering practice, many attempts were made to assess the distance of influence and to evaluate other effects like hydraulic boundaries, hydraulic gradients previous to pumping, anisotropy, influence of diffuse infiltration, and so on. The most extensive evaluation of the limitations of this method was performed by Leland K. Wenzel from a large pumping and recovery test on Grand Island (Nebraska), using an analysis of the drawdown in 80 observation wells (Wenzel, 1932, 1936).

A breakthrough in the development of well hydraulics were the analyses by Charles V. Theis (1900–1987) of the nonsteady state conditions during groundwater extraction and recovery. Theis (1935) derived his formulae for drawdown and recovery of the water level in the vicinity of a pumped well through a delivery of water from storage, by making use of an analogy with heat conduction and capacity. Theis assumed an instantaneous release of water at the falling water table, and introduced the related concept of *storage coefficient*. He also adequately simulated the recovery phase by assuming a recharge of the well at the same constant rate as the extraction rate. Theis was assisted in these studies by his friend the mathematician C.I. Lubin who, however, refused co-authorship. An earlier attempt to solve the same problem was made by the German Weber (1928). His solution was based on a constant drawdown and a shifting radius of influence, whereas Theis considered an infinite distance of influence. Norman S. Boulton, in the UK, also predated Theis with a transient solution in the early 1930s, but according to Downing et al. (2004) his paper was rejected for publication. Transient flow in confined aquifers with change in storage from compressible fluids was further considered in research connected to oil exploration by Hurst (1934) and Muskat (1934).

Charles E. Jacob (1914–1970) subsequently derived non-equilibrium theory on the basis of pure hydrological concepts, and arrived at a formula for the confined storativity based on the physics of compressibility of both rock and water. He in fact produced the first comprehensive formulation of transient confined flow (Jacob, 1940, 1950). Jacob also analyzed the effect of a change in atmospheric

pressure on the water well level (*barometric efficiency*) in terms of compressibility of the aquifer. Subsequent mathematical solutions for unsteady-state radial flow in a leaky aquifer, without release from storage from the confining layer, were given by Hantush and Jacob (1955). These formulae were expanded by (Iraq-born) Mahadi Hantush (1921–1984) for conditions where water is released from storage in the confining layer (Hantush, 1960, 1964).

An engineer with the Amsterdam waterworks, L. Huisman, and mathematician J. Kemperman, extended the “Kooper–De Glee formula” for an aquifer between two semi-confining layers (Huisman and Kemperman, 1951). The solution of this problem was needed to adequately analyze multi layer pumping tests in the Amsterdam dune catchment. The theory was tested by data from a large-scale pumping test for design of a deep-well drainage scheme for a tunnel excavation below the shipping canal between Amsterdam and the North Sea. Subsequently, a solution for transient well flow in a leaky double-layered aquifer was derived by Hantush (1967).

Prediction of the decline in shallow groundwater levels by groundwater extraction or by land drainage has always been an important question in the Netherlands in connection with land subsidence, reduction of crop yield, up-welling of salt water, and damage to wooden foundation piles in the old cities. Large-scale exploitation of a 250 m thick, complex stratified aquifer system in the Netherlands Rhine-Meuse delta initiated in the 1980s rigorous mathematical analyses of steady and transient well flow in multiple-leaky aquifer systems by, notably, G.A. Bruggeman, C.J. Hemker, and C. Maas. They applied combined analytical-numerical techniques in a model that has some similarity with the double porosity concept in hard rock aquifers (see below), and verified the results by extensive field experiments. In the United States, Dutch-born Otto Strack developed his multilayer analytical element model. Recently Hemker and M. Bakker (Dutch Ph.D. graduate of Strack) extended the multilayer approach to include different horizontal anisotropic conductivities in one or more layers. 3D flow pattern analysis in such aquifers revealed that stream lines often create spirals, a bundle of which is termed “groundwater whirl”.

Advances in groundwater hydraulics in the first half of the 20th century, and notably in the period 1930–1960, were dominated by analysis of groundwater extraction by wells, though other subjects like flow in the vicinity of canals and seepage through and below dikes and dams were also considered, particularly in the Netherlands coastal lowlands. For example, by using a complex variable technique C.G.J. Vreedenburgh (De Vos, 1929) was probably the first to derive an exact mathematical solution for flow through a dam under free water table conditions. Moreover, he solved the problem of flow through an anisotropic medium (Vreedenburgh, 1935), and simulated 2D groundwater flow in an electrolyte tank on the basis of analogy between Darcy's law and Ohm's law (Vreedenburgh and Stevens, 1936). At that time Vreedenburgh was a professor at the Technical University of Bandung (former Dutch East Indies), and although his studies basically originated from the problem of seepage through dams and levees in irrigated fields, this work was also applicable to Dutch polder conditions. The problem of flow pattern analysis in an homogeneous but anisotropic medium was later on extensively discussed in the United States by (Dutch-born) M. Maasland in the framework of land drainage studies (Maasland, 1957).

Further mention should be made of Steggewentz' (1933) Ph.D. thesis on the propagation of ocean tides in coastal aquifers, which included an early solution for delayed vertical flow near the water table in response to water pressure oscillations. Steggewentz assumed that this flow resistance at the capillary fringe explained the relatively fast propagation of water pressure in phreatic aquifers in the coastal dune area. It was only in the 1950s that this concept appeared in the international literature as the *delayed yield* principle, in connection with the early phase of transient well flow (Boulton, 1954). Boulton's semi-empirical approach with exponentially decreasing leakage flux with time was further mathematically and physically extended in the 1970s, notably by S.P. Neumann and T.D. Streltsova. Several Dutch contributions to the analysis of ocean tides in leaky and compressible aquifers, including the problem of the finite response time, were subsequently produced (cf. Raats, 1969; Van der Kamp, 1973). Considerations in the United States of periodic groundwater flow and bank storage induced by stage oscillations in adjacent open water are notably known from Ferris (1951), and Cooper and Rorabaugh (1963).

More general analyses of the transient, nonlinear interaction of the saturated–unsaturated zone near the water table became within reach after soil scientists (notably Klute, 1952) introduced the numerical

finite difference approach of the Richards equation. In particular, the problem of pore-space drainage due to a rapidly falling water table was studied with respect to the meaning and process-dependence of the *specific yield* concept (e.g., Dos Santos Junior and Youngs, 1969). The soil physicist Don Kirkham (1967) examined in this connection the position of the moving free surface (flow line boundary) in the Dupuit–Forchheimer flow case. This problem is dominated by the paradox of a changing storage, without possible interaction of the saturated–unsaturated zones by a front that crosses the flow-line boundary.

The requirement for information on small-scale spatial variability of hydrogeologic properties came particularly into perspective with the need for a distributed flow approach in heterogeneous aquifers, in connection with numerical modeling of combined water and solute transport. These developments were stimulated and accelerated by the environmental pollution problems which surfaced especially in the early 1970s, after publication in 1972 of the Club of Rome report *The Limits to Growth*. There was a developing awareness that contamination by leaking waste disposal, oil spills, and by agriculture activities, together with increasing extraction of groundwater, had continuously deteriorated groundwater quality. Groundwater hydrologists evidently had to shift their focus from water quantity to water quality, and needed to include spatial heterogeneity in their flow analyses (cf. Section 1.3.5). Pioneering geostatistical studies of the transport parameters were carried out, among others, by Freeze (1975) in Canada, Gelhar (1977) in the United States, and Delhomme (1979) in France.

Many other studies have subsequently been carried out in the framework of reservoir engineering, particularly in connection with the disposal of hazardous and radioactive wastes. These studies include flow through fractured media and are closely related to regional stress and strain-field analysis in structural geology. A first approach to a general theory of inhomogeneous fractured media was the *double-porosity* concept, initially developed by the Russians Barenblatt et al. (1960). This concept is based on a division of the subsurface into a storage-providing matrix block and a transport-facilitating fracture system. Boulton and Streletsova (1977) pioneered well flow in a fissured medium in this way, after earlier work had been carried out on flow in single-fracture systems (Gringarten and Witherspoon, 1972). Important in this respect are the continuing advances in exploration technology, including surface and subsurface geophysics and remote sensing techniques, which have contributed to a better determination of the distribution of hydrogeological structures and flow patterns. In addition, inverse numerical computer modeling became a tool in the 1970s to study subsurface structures and associated distributed parameters.

Other aspects of groundwater in relation to geological structure and geo technical problems are associated with the geothermal gradient and temperature distributions in regional groundwater systems. Most accounts of groundwater temperatures are, however, concerned with thermal springs and date back to the 19th century. Early analysis concerning the simultaneous flow of water and heat are from the 1940s by Henri Schoeller in France (cf. Schoeller, 1962, Chapter VI). More recent, well-known studies from the United States are those by Stallman (1963), and Bredehoeft and Papadopoulos (1965). In a classic paper, Hubbert and Rubey (1959) analyzed the influence of groundwater pressure on fault movement, and notably explained the overthrust faulting on the basis of Terzaghi's theory of effective stress. Bredehoeft (1967) considered earth-tide induced groundwater level oscillations on the basis of the elastic vertical deformation of the earth crust due to the moving tidal bulge in the solid earth. Evidently, geology and hydrology strongly interact with processes on a geological time scale; examples are thermo-chemical convection, and water and geochemical-related diagenesis and consolidation processes.

1.3.3.2 The Problem of the Fresh–Salt Water Interface

Salinization of the coastal dune area in the Netherlands, caused by over exploitation of groundwater for drinking water, proceeded until artificial recharge with river water after World War II began to reverse the process. Different exploitation strategies had been exercised in the preceding period to prevent salinization; these were mainly based on a choice between either deep extraction from a semi-confined part of the subsurface or shallow extraction from the upper phreatic aquifer. Mazure (1943) presented an analytical solution for predicting the evolution of the fresh water pockets for these different scenarios. His approach was further elaborated by Todd and Huisman (1959) for the Amsterdam dune water catchment, making use of the first main frame electronic computer at the Amsterdam Mathematical Center for the laborious

numerical calculations. More advanced analyses, including the flow of salt water below the fresh water pocket toward the deep polders behind the dunes, as well as the influence of artificial recharge, were subsequently carried out when the computing power increased.

Further theoretical analyses of transient flow were initiated to predict groundwater evolution in the coastal dunes caused by the proposed artificial infiltration from broad channels. This work was carried out by the engineer with the Amsterdam Waterworks, J.H. Edelman, and published in his 1947 Ph.D. thesis (cf. Edelman, 1972). His work included analysis of (i) propagation into the adjacent aquifer of instantaneous water level changes in channels, (ii) transient well flow (an alternative solution to arrive at the Theis formula, based on Weber, 1928), and (iii) numerical calculations using a hexagonal finite difference network for regional groundwater flow, including density-driven flow. The finite difference approach was derived from mechanical engineering (Richardson, 1911), and Edelman used the relaxation method of Southwell (1940) for the iteration and optimization procedures.

In the course of time, many analytical and numerical studies of groundwater flow involving fresh and salt water followed in the Netherlands, initially supported by electrical resistance and capacitance-network analogons, and subsequently by computer simulation models. An example of the first category was the 5000 grid-point Electric Network Analogon for Groundwater Flow (ELNAG). This was developed in the 1960s to forecast steady-as well as transient fresh and salt groundwater flow in connection with a large-scale water management and flood protection project (the so-called Delta Works) in the south-western estuarine area of the Netherlands. This scheme, carried out from the 1950s to the 1980s, was a response to the devastating flood that struck the Netherlands in 1953. Other studies concentrated on reconstructing and predicting the long-term redistribution of fresh and salt water occurrences, caused by 1000 years of land reclamation and associated changing groundwater level topography (CHO-TNO, 1980; De Vries, 1981).

Many studies on the fresh-saltwater interface were carried out in the United States, particularly in the extensively studied coastal area of Florida. Luszynski (1961) introduced the valuable concept of the *point-water-level* to reduce head measurements in a multi-density field to the same density. Glover (1959) contributed to the influence of seepage across the interface, and Kohout (1960) from the United States as well as De Josselin de Jong (1960) from the Netherlands, and Bear (1961) from Israel, considered the multi-density interaction processes at the interface, including dispersion and diffusion. Numerous investigations followed in coastal areas on the behavior of the interface under the influence of external conditions (e.g., Verruijt, 1969). Research in this field in Europe was highly stimulated by the regularly organized Salt Water Intrusion Meetings (SWIM), which started in 1967 as a Dutch-German initiative.

1.3.3.3 Consolidation of Hydrogeology as a Scientific Discipline

Numerous solutions of the general flow equation for different schematizations of subsurface structures, liquid density differences, and the character of the boundary and initial conditions have been developed in the course of the half-century between the 1920s and 1970s. Many mathematical techniques were identified as being suitable to solve the Laplace equation, including graphical flow net analysis, separation of variables, integral transformations, conformal mapping, and numerical methods. Electric analogons were widely used to particularly study cause-and-effect relationships in complicated aquifer situations for steady- as well as transient flow (e.g., Karplus, 1958). These analogons were gradually replaced from the 1960s by the more flexible numerical solutions using computer simulation models (e.g., Fayers and Sheldon, 1962; Chun et al., 1964). Viscous parallel plate models were applied particularly for flow conditions with multi-density fluids. In the Netherlands, a unique, horizontal, parallel plate model with vertical storage was build by the National Bureau for Drinking Water Supply to simulate unsteady flow under water table conditions, and to assess the effect of groundwater extractions on the water table (Santing, 1957).

For many years, Morris Muskat's (1937) monumental book with theoretical solutions of flow problems for homogeneous and heterogeneous fluids (particularly meant for petroleum engineers) remained a reference in the English-speaking world; German-speaking countries mainly used a similar text of more limited scope by Dachler (1936). Geological aspects of groundwater behavior were emphasized in Cyrus F. Tolman's (1937) textbook. Theoreticians in search for new solutions often made use of Carslaw and Jaeger's standard text on analytical solutions of heat conduction, the first edition of which appeared in

1921. However, it was only in 1940 that Marion King Hubbert (1903–1989) discovered that the until then applied concept of the velocity potential (Kh) in the Laplace equation did not have general validity in the case of a 2D and 3D anisotropic flow field. He, therefore, proposed the generally applicable concept of the force potential (gh). He further examined Darcy's law in the light of the microscopic Navier–Stokes flow theory, and emphasized its character as a macroscopic law. Hubbert, who was familiar with earlier oil exploration research on the influence of fluid properties on percolation, (notably by R.D. Wyckhoff and others, including Muskat), further reviewed the concept of *intrinsic permeability* and the limits of the validity of Darcy's law. With a graphical solution for the pattern of flow lines between recharge areas and the associated drainage line, he explicitly illustrated the connection between water table elevation at a point and the value of the intersecting equipotential line. His well-known flow-net picture is, however, not essentially different from Pennink's case, as is depicted in Figure 1.5. It can be stated that Hubbert's rigorous treatise finally consolidated the scientific basis of groundwater hydrology within the general theory of hydrodynamics.

Information on the partly parallel developments in Eastern Europe and the former USSR, notably on the theory of groundwater flow, became available with the 1962 translation of Polubarinova-Kochina's (1952) well-known textbook *Theory of Groundwater Flow*, and subsequently through the text *Groundwater Hydraulics* by the Czechs Hálek and Sveč (1979). Until the early 1960s, the most complete volume on the theory and application of groundwater hydrology, including physical and chemical aspects, was probably Henri Schoeller's (1962) *Les Eaux Souterraines*. An accessible and popular English text on groundwater engineering, in the early 1960s, was David K. Todd's (1959) *Groundwater Hydrology*; this was followed by a more elaborate treatment by Milton E. Harr (1962). A large number of more advanced groundwater textbooks have subsequently been produced since the mid-1960s, and specialized international journals have appeared, including *Hydrologic Science Bulletin of the IASH* in 1956, *Journal of Hydrology* in 1963, *Ground Water* in 1963, and *Water Resources Research* in 1965. This reflects the increasing world wide interest in water resources management in this period, and the recognition of hydrology as a discipline in its own right.

These developments are also illustrated by the founding of the International Association of Hydrogeologists (IAH) in 1956, and the establishment in 1965 of UNESCO's International Hydrological Decade (IHD), forerunner of the present International Hydrological Programme (IHP). Another indication of emancipation of hydrology as a science was the 1971 decision by the International Association of Scientific Hydrology (IASH) to change its — somewhat defensive — name to International Association of Hydrological Sciences (IAHS). Well-known examples of broad and integrated texts on groundwater from the mid-1960s to the late 1970s include Davis and De Wiest (1966), Walton (1970), and notably Freeze and Cherry (1979). Bear et al. (1968), Verruijt (1970), and Bear (1972) produced an overview of the fundamentals of flow through porous media and theoretical approaches to the solution of groundwater flow problems. Remson et al. (1971) covered the application of numerical analysis, Kruseman and De Ridder (1970) developed a manual for analysis and interpretation of pumping test data, and Zötl (1974) wrote a comprehensive text on karst hydrology. Specialized volumes on chemical and contaminant hydrogeology, groundwater recharge, and isotope hydrogeology followed in the 1980s and 1990s.

1.3.4 Integrated Analysis of Regional Groundwater Systems

1.3.4.1 Basin-Scale Analysis and Evaluation

Regional reconnaissance studies in the early 20th century in the United States, included the delineation of recharge and discharge areas, flow pattern analysis, temperature measurements, monitoring of groundwater level fluctuations, and the determination of aquifer properties by steady-state pumping tests. Lysimeter experiments to assess elements of the water balance were already applied in the early 19th century, but their conditions were often very different from the undisturbed terrain and difficult to regionalize over larger areas. Veatch (1906) analyzed a large number of groundwater level fluctuations, particularly from Long Island, and reviewed the results of earlier lysimeter research. A first quantitative approach, based on water budget considerations, was presented by the chief of the Groundwater Branch

of the USGS, Mendenhall (1905), on an aquifer in South California. These studies were extended by Charles H. Lee (1914) into one of the first sound resource evaluations with regard to the *safe yield* concept.

The first extensive regional study in the Netherlands was carried out by the National Bureau for Drinking Water Supply in the coastal dune area. An analysis of the results from a large number of boreholes to depths of more than 100 m, with piezometers at different levels, produced a detailed 3D picture of the hydraulic head, corrected for water density, and the associated flow patterns (Van Oldenborgh, 1916). In addition, four large lysimeters (each with dimensions of 25×25 m and a depth of 2.5 m) were built in the 1930s in the dunes near Castricum to study the influence of vegetation on groundwater replenishment, in particular to assess the effect of afforestation of the dunes with pine trees. The results revealed a recharge reduction under pines of more than 50% relative to the annual 350 mm under natural dune shrub vegetation (cf. Stuyfzand, 1993).

Groundwater recharge and discharge components on a basin scale were studied during the first half of the 20th century almost exclusively by the water budget approach. Phreatic aquifer recharge was quantified by a combination of groundwater level rise and recession hydrographs, whereas river basins were analyzed using measurements of rainfall, stream runoff, and estimates of evapotranspiration. The study of base flow recession curves has been widely used to identify the hydrogeologic character of watersheds, and to separate fast (near-)surface runoff components of recent rainfall from storage prior to rainfall. This problem of hydrograph separation was extensively reviewed by Hoyt (1936), the results from an assessment of annual groundwater discharge from 22 US stream basins.

Numerous empirical studies have been carried out in the United States, as well as in Europe and the former USSR, to relate annual stream flow to rainfall. Most formulae were of the form $Q = aR - b$, where R is annual rainfall, and a and b are regional or watershed-specific parameters for average long-term conditions. In more advanced studies, the evaporation-related factors a and b are variables, expressed as functions of average temperature. Temperature-dependent annual evapotranspiration relations that have been widely used in watershed balance studies are, for example, the empirical potential evaporation formulae of Thornthwaite (1948) from the United States. In France, Turc (1954) proposed formulae for average discharge as a function of average temperature and rainfall based on a large number of watershed studies. These methods provide reasonable order of magnitude results, notably for the annual time scale. Thornthwaite also used his formulae for his well-known climate classification based on the monthly difference between rainfall and potential evaporation (Thornthwaite and Mather, 1955). More complex empirical rainfall-runoff relations, including watershed morphological characteristics, were developed for storm runoff prediction. R.E. Horton's papers on infiltration in the 1930s and 1940s, in particular, initiated numerous studies on the separation of surface runoff and groundwater discharge (cf. Hall, 1968). A well-known example of a semi-empirical runoff model with numerous parameters is the "Stanford Watershed Model," application of which only became feasible after the introduction of powerful digital computers in the late 1950s.

An analytical study of basin-scale dynamics of groundwater level fluctuations due to seasonal and inter-annual pulses of recharge was proposed by the Belgian Leon J. Tison (1951). (Tison was well-known for his leading position in the International Association of Scientific Hydrology.) Wright and Edmunds (1971) analyzed the depletion and residual flow in "fossil" groundwater basins in the Libyan Sahara on a time scale of 10,000 years (cited by Burdon, 1977, who similarly considered the whole Sahara and Arabian desert). More advanced time-series analysis of groundwater levels became available in the 1970s from application of the semi-empirical transfer function modeling procedures of Box and Jenkins (1970), originally developed to analyze economic trends. This method estimates groundwater level fluctuation as a response function of the external driving forces, on the basis of an empirical stochastic correlation and a deterministic auto-correlation (recession) function. It proved to be particularly applicable as a diagnostic tool to detect trend or to identify the influence of interfering processes. Evaluation of groundwater resources got a new perspective with the emergence in the 1960s of environmental isotopes and noble gas tracers as a tool to reconstruct paleohydrologic evolution, and to assess recharge and discharge components (cf. Section 1.3.5.2).

1.3.4.2 Topography Driven Groundwater Flow Systems

Until the 1960s emphasis was on the exploration and engineering requirements of groundwater hydrology, including the inventory of groundwater resources and groundwater balance studies. A more integrated regional approach to gain understanding of natural groundwater flow systems and associated hydrochemical process interaction with the subsurface, topography, vegetation, and climate, emerged in the 1960s from the work of Peter Meyboom and József Tóth, notably in the glacial hummocky terrain of the Canadian prairies. In these areas the water table follows the topography at a regional scale because the hydrogeology is dominated by clay-rich glacial till with low permeability. Most of the shallow water movement is therefore regulated by topography-controlled infiltration and evapotranspiration, which accordingly control the depth of the water table below ground surface.

Meyboom mapped the spatially distributed groundwater heads in the vicinity of glacial potholes, and recognized the seasonal dynamics of expanding and contracting seepage zones and shifts in groundwater flow patterns. He further identified discharge phenomena related to these “groundwater outcrops” such as springs, seeps, and salt tolerant vegetation in exfiltration zones with evaporation-induced increased salinity (Meyboom, 1966). Tóth (1962, 1963) derived an analytical solution (using a separation of variables technique) for the steady-state topography-driven flow systems with areas of recharge and discharge, the latter often characterized by seepage zones. It followed from the observations by Meyboom and the theoretical analysis of Tóth that, depending on the topography, nested flow systems of different order are generated which are characterized by shallow local systems superimposed on deep, regional systems.

Allen Freeze and Paul A. Witherspoon (1966, 1967) applied numerical analysis to examine and extend Tóth's flow systems by including subsurface heterogeneity. The concept of topography-driven and geologically transformed local and regional flow systems has subsequently been widely used in exploration, mapping, and analysis of groundwater basins at various scales. It should be noted, however, that the concept of nested flow systems was not new. For example, René d'Andrimont of Belgium, who was inspired by Pennink's work, recognized this regional flow pattern at the beginning of the 20th century and presented a picture of the phenomenon as a novelty at Brussels World Exhibition in 1910 (Figure 1.6; D'Andrimont, 1910).

There is a strong similarity between the Tóth–Meyboom approach and investigations in the Netherlands on groundwater drainage of shallow aquifers by stream systems of different order. Lodewijk F. Ernst of the Institute of Land and Water Management Research (ICW), in particular, analyzed the interaction of topography and the seasonal contraction and expansion of the drainage system caused by fluctuations in the groundwater table; he arrived at formulae, which were similar to those from Tóth (Ernst, 1962, p. 84). Both Meyboom and Tóth (who originated from the Netherlands and Hungary respectively) studied in the Netherlands, and were familiar with Ernst's work in general, but Tóth and Ernst derived their analytical formulas independently. Jacobus J. de Vries (1977) expanded Ernst's ideas to analyze the natural stream network under shallow aquifer conditions, and to explain them genetically as outcrops of groundwater flow systems of different order, driven by topography, climate, and seasonally fluctuating water tables.

1.3.4.3 Groundwater Discharge and Land Drainage

The Netherlands has more than 1000 years of tradition in land drainage, but the design of dewatering systems used to be based mainly on experience. This changed in the 1930s when a 250,000 ha land reclamation scheme in the former Zuiderzee (a North Sea embayment) was initiated, and a rational water management of this newly created land was proposed. Trail blazing theoretical analyses and experiments related to land drainage processes were carried out at laboratory, plot, and field scale by Symen B. Hooghoudt (1901–1953) at the Groningen Soil Science Institute and Experimental Station. These led to formulae for steady, horizontal flow through two-layer soils with radial converging flow components near the drainage channels (Hooghoudt, 1940). Hooghoudt simply transformed, by imaging, radial horizontal well flow theory into vertical radial flow, in order to evaluate the effective depth of drainage.

His work was further advanced theoretically by his students, the physicists Van Deemter (1950) and Ernst (1956), who included soil heterogeneity, drainage channel geometry, and nonlinear flow considerations.

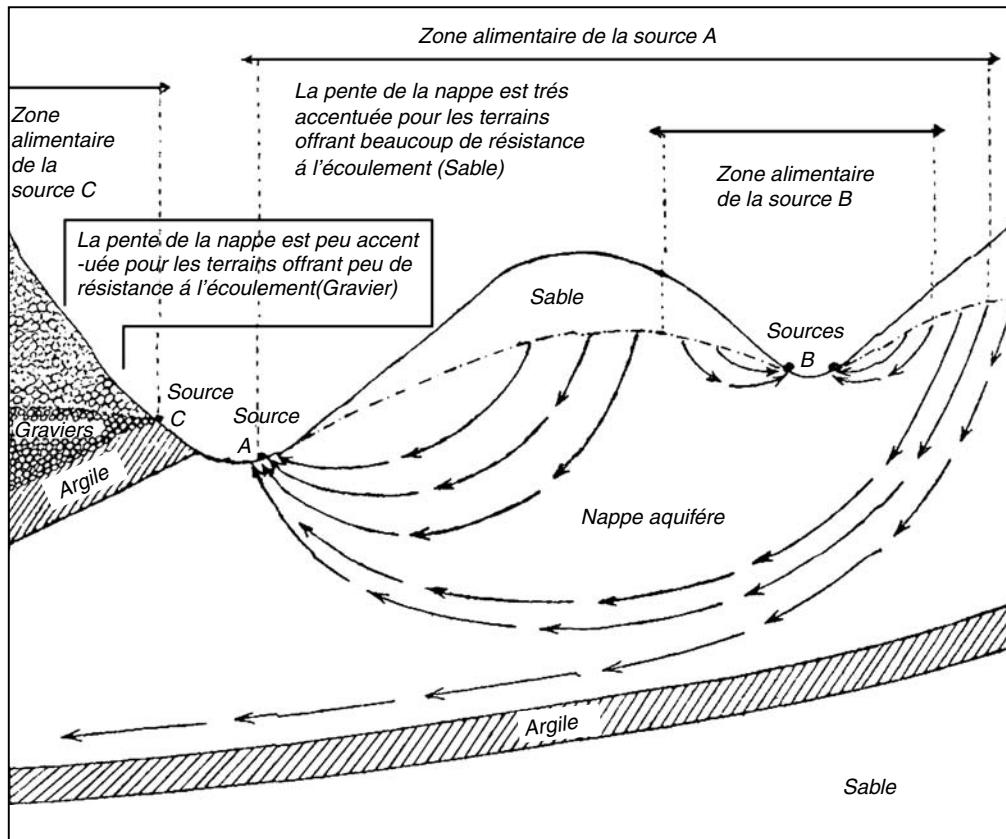


FIGURE 1.6 Groundwater flow systems; picture presented at the Brussels World Exhibition in 1910 by René d'Andrimont.

Their work was partly based on early analytical formulae by the Dutch professor of hydrodynamics, Burgers (1926) which they combined with the numerical relaxation method according to Southwell (1940). Similar situations were investigated in the United States by the soil physicist Don Kirkham and collaborators (e.g. Kirkham, 1954). Dutch-born Ph.D. students at Cornell and Iowa State Universities (notably Jan van Schilfgaarde) facilitated the export of Hooghoudt's ideas (which were published mainly in Dutch) to the United States (e.g., Van Schilfgaarde et al., 1956; cf. Raats and Van der Ploeg, 2005).

Previous to Hooghoudt's pioneering work, land drainage in Western Europe was based mainly on simple Dupuit assumptions of horizontal flow to ditches, and fed by uniform rain infiltration at the water table (cf. Section 1.3.1.1, Equation 1.4). This approach had already been applied in the second half of the 19th century by the Dane Ludvig Colding, who was in turn inspired by drainage experiments by the French engineer S.C. Delacroix in the 1850s (cf. Raats and Van der Ploeg, 2005).

Numerous studies have been carried out since the 1950s on time-dependent drainage processes from shallow flat-land aquifers with high drainage density, in connection with dewatering and groundwater table control. These processes are complex and strongly nonlinear because the shape of the water table, as well as the storage factor, are changing in time and space, and because of the influence of the drainage channel geometry on the flow resistance. An approximate, but reasonably adequate formula for the drainage process was derived by Dirk A. Kraijenhoff van de Lier (1958) in the Netherlands. He assumed a cascade of linear storage basins and combined this with the effect of the falling water table near the channel. This solution of the impulse-response process is related to the Glover-Dumm treatise of a falling groundwater table after an instantaneous delivered depth of irrigation water at the surface (Dumm, 1954),

and to Edelman's (1947) groundwater reaction following an instantaneous lowering of the water level in a drainage channel (cf. Section 1.3.3.2).

The physical processes of infiltration and the separation of shallow subsurface processes from percolation to the groundwater table are the domain of soil physics. Numerous studies by soil physicists, particularly in connection with agricultural water management and crop production, have been carried out on experimental plot scales. Many empirical formulae have been proposed for the relation between unsaturated hydraulic conductivity and soil moisture suction, based on textural properties of the soil. Well-known early formulae are those of Childs and Collis-George (1950) and Gardner (1958). Youngs' (1958) study of the vertical redistribution of moisture during and after the process of infiltration presented for the first time a coherent picture of the transient flow pattern in the unsaturated zone.

Paramount were the fundamental, physically based infiltration process studies by the Australian John R. Philip from the mid-1950s onwards (e.g., Philip, 1969). More relevant to regional hydrological processes were the numerical studies by Rubin and Steinhardt (1963) and Rubin et al. (1964) to predict infiltration from rainfall intensity and antecedent moisture conditions, as well as the integrated saturated–unsaturated modeling by Freeze (1969, 1971) to simulate the time-dependent response from steady ponding conditions resulting from constant rainfall.

Studies of capillary processes also led to a better understanding of evaporation from the soil. For example, Keen (1927) demonstrated that capillary uplift of moisture from the water table to the evaporating surface or root zone is strongly reduced if the water table is at a depth of about 35 cm for coarse sand and 85 cm for loam. More fundamental studies followed after the development of the physically-based Penman–Monteith evapotranspiration formulae (Penman, 1948; Monteith, 1965; Rijtema, 1965), which enabled to assess the actual evapotranspiration as a moisture–stress related function of the potential evapotranspiration.

Many experiments were carried out in the Netherlands from the 1950s through the 1970s to relate upward capillary flux and evapotranspiration to the depth of the groundwater table in shallow aquifers, in order to optimize groundwater depth with respect to crop yield (Wind, 1955). One of the applications was to determine crop yield reduction and damage to valuable wetlands caused by a lowering of the water table through groundwater extraction and land drainage. Similar studies have been carried out in arid areas in connection with irrigation requirements and the related hazards of salinization and clogging.

1.3.4.4 Runoff Generation on Sloping Areas

Land drainage studies are mainly connected with more or less artificially controlled conditions, often on low lying alluvial ground in more or less homogeneous situations. Under such conditions the drainage process can be characterized by lumped parameters, such as the *reservoir coefficient* of Kraijenhoff van de Leur (1958), that include the geometry of the drainage network as well as the subsurface permeability and storativity. Groundwater processes in natural watersheds and the generation of runoff, particularly on slopes and in heterogeneous areas, are much more complex. The separation of groundwater discharge from the total drainage basin runoff is a problem that is related to the distributed interaction of overland flow, saturated and unsaturated flow, and the process of infiltration.

Early investigations on infiltration of rainfall were carried out by Robert E. Horton (1933, 1945) as part of his pioneering studies on surface runoff, channel generation, and stream flow. He observed that infiltration capacity decreases rapidly over time in the course of a storm, and thus emphasized the dominance of surface processes. His work, however, referred mainly to semi arid conditions where overland flow is indeed common, especially because of the forming of infiltration-hampering salt crusts. At the same time, ideas about rapid runoff generation caused by interflow through piping, macro-pores, and perched saturated flow in stratified soils were, among others, advanced by Hursh (1936), Barnes (1939) and Hursh and Brater (1941), but their observations were probably neglected because of the authoritative concept of the so-called “Hortonian overland flow”.

It was only in the 1960s that it became evident that Horton's model did not apply to temperate areas with vegetation cover. Notably J.D. Hewlett of the US Forest Service, together with A.R Hibbert, developed

a new concept based on the observation that almost all rainwater normally infiltrates. This subsurface water subsequently runs off, partly through high permeability shallow horizons as flow in a perched saturated–unsaturated zone, and partly as deeper groundwater flow. Storm runoff is generated through converging discharge in the valley heads and other depressions with shallow groundwater tables, where the unsaturated zone becomes saturated and overland flow starts. The expanding *surface-saturated* zones develop notably along the upstream valley slopes and are responsible for the rapidly responding storm runoff (Hurst and Brater, 1941; Hewlett, 1961; Hewlett and Hibbert, 1963, 1967; Kirkby and Chorley, 1967). These so-called *partial-area* and *variable-source-area* concepts have been confirmed in later studies using tagged water and isotope tracer analysis (e.g., Fritz et al., 1976). There is a certain similarity with the previously mentioned contracting and expanding stream systems associated with shallow aquifers in level areas, where seasonal groundwater fluctuations determine the number and order of streams that participate in the groundwater drainage process (De Vries, 1977; Ernst, 1978).

Dunne and Black (1970a, 1970b) succeeded in simultaneously measuring the three components of discharge in an experimental watershed with an integrated set of instruments. The combined surface–subsurface, saturated–unsaturated flow approach has subsequently become the basis of runoff generation studies with reference to catchment morphologic conditions. Attempts in the 1960s to combine flow in saturated and unsaturated zones in numerical schemes paved the way for modeling of this saturated–unsaturated runoff system. One of the first integrated models was presented by Freeze (1972a, 1972b). Others applied environmental chemical and isotope tracers to separate the different systems, and to analyze their dynamics in field studies (e.g., Farvolden, 1963; Pinder and Jones, 1969).

These chemical tracer studies also confirmed the earlier assumptions that old water in storage is pushed out by newly recharged water. This combination of different approaches has eventually resulted in advanced physically based distributed modeling, in which the dynamic-contributed-area effect is combined with channel network topology; a well known model is TOPmodel (Bevin and Kirkby, 1979). Such catchment scale processes are evidently highly nonlinear in time and space, and subject to scaling problems.

These runoff generating processes are strongly related to slope erosion. Understanding of shallow subsurface flow and related surface runoff is essential for the understanding of the slope processes, and particularly of scarp retreat and the head wards extension of valley tributaries. Flow in a perched saturated zone can lead to seepage erosion and tunnelling by entrainment of grains in water emerging from a porous medium (Terzaghi, 1943; Buckham and Cockfield, 1950). The role of subsurface flow in stream channel initiation and development was proposed by, among others, Chorley and Morgan (1962), Dunne (1970), Jones (1971), and De Vries (1977).

These integrated analyses of runoff generation and slope erosion processes merged well with the systems approach in process-based landform evolution studies in physical geography, as proposed by Richard Chorley and Barbara Kennedy (1971). Systems analysis, and its dynamic feedback mechanisms, had evolved in the 1960s as a unifying conceptual framework in engineering as well as ecological studies to model cause–effect relationships. Patrick Domenico (1972) wrote a textbook on models in groundwater hydrology and groundwater management within the framework of this approach. It is clear that the analysis and modeling of integrated hydrological systems that comprise coupled nonlinear processes, only became feasible after the introduction of computer simulation models in the 1960s.

1.3.5 Evolution of Hydrochemistry and Isotope Hydrology

1.3.5.1 Hydrochemistry

The chemical processes in groundwater are much more complex than those studied by engineers in chemical reactors under controlled conditions. Groundwater on its way to the discharge area is normally interacting with aquifers, which are composed of different rock and soil minerals and which are subjected to the inflow of water, solutes, and gases from various origins, thus producing a new reactive mix. Equilibrium is therefore seldom reached. Until the mid-20th century, hydrochemistry was mainly restricted to groundwater inventories in order to determine its quality for use as drinking water and

for agricultural purpose. In hydrogeological exploration, chemical characteristics were attributed to petrographic rock type to enable determination of the interconnection between aquifers and origin of the waters. Classifications were developed for grouping water from a single aquifer, or from the same origin, and to identify dominant chemical processes and rock interaction, and to study chemical trends and evolution in connection with flow fields. The oldest classifications are based on Palmer's (1911) concepts of salinity and alkalinity, which are related to the contents of non-hydrolyzed salts and weak acids components, respectively.

Diagrams and graphs were developed to summarize large quantities of data, and to visually characterize water for interpretation and mapping. Well-known diagrams which are suitable to plot on maps are the 3-vector diagram of Rogers (1917), the bar diagrams according to Collins (1923), and the distinctive symbol diagrams from Stiff (1951). More genetically indicative are the triangular diagrams by Emmons and Harrington (1913), where similarity is characterized by the plot position of anion and cation ratios. Hill (1940) combined these triangular figures with a diamond-shaped diagram in which anion and cation ratios are combined into one plot; this procedure became well-known through the work of Piper (1944). Also genetically indicative are the logarithmic nomographs by Schoeller (1935), in which comparable compositions are indicated by parallel lines.

The dominant physico-chemical processes that could be identified in the early days were dissolution in connection with solubility of various minerals, sulphate reduction under anaerobic conditions, and cation exchange processes. Sulphate reduction and the role of bacteria were already known in the late 19th century, notably from groundwater in petroleum reservoirs. Exchange of sodium by calcium at the adsorption complex, as a product of freshening of originally saline aquifers, was proposed by Jan Versluys (1916) to explain sodium-bicarbonate water in the Dutch coastal area. This water type therefore became a useful tracer in reconstructing the hydrogeologic evolution in relation to coastal history. The same type of process had often been observed under the American Great Plains, and was characterized as "natural softening" because of the calcium depletion (Renick, 1924).

An important advancement for understanding the continuous dissolution of limestone along a flow path in the saturated zone was brought into perspective by discovery of the nonlinear relation between free CO₂ and the concentration of dissolved calcium. This explains the possibility of continued corrosion through inflow of new water with a different CO₂ content. If both existing and new water are in equilibrium with respect to CO₂ and calcite, the nonlinear relation causes their mixture to be undersaturated in this regard (Tillmans, 1932). Alfred Bögli (1964) applied this concept of mixing corrosion in his physically based karst-water studies. Numerous regional investigations have been carried out to explain the dynamics and development of karst systems, including the influence of geological structure on propagation of conduit systems and cave evolution.

A degree of zonation from the water table downward, and with distance from the intake area, could be recognized in the Great Plain aquifers in the United States. This zonation is dominated by a trend in which bicarbonate water is gradually converted into a mix of sulphate/chloride water and, eventually, into chloride dominated water, toward a sea water composition. This stratification was attributed to an increased ionic concentration induced by an increase in temperature and a prolonged circulation and retention, causing a relative abundance of the ions of the most soluble salts (Rogers, 1917).

Similar chemical evolution, caused by climatic zonation, was recognized by Schoeller (1941) from his studies of chemical evolution along flow trajectories in northern Africa. He observed an increase in solute concentration from the (semi-) humid to the dryer (semi-)arid regions. Schoeller explained this effect by postulating a net loss of water by evapotranspiration and lack of carbon-dioxide production in the latter areas, which had led to precipitation of calcium/magnesium carbonate and gypsum, and a concurrent enrichment of chlorides. In the USSR, with its clear climatic- and associated soil type zonation, comparable observations were made and explained in 1948 by Garmanov (Silin-Bekchurin, 1951; cf. Schoeller, 1962, Chapter VII) and Garmanov (1958). The Russian Chebotorev (1955) further included 10,000 analyses from Australia to analyze climatic zonation. The above cited Russian authors represent a large volume of literature from the USSR, including theoretical studies. Hydrochemistry in the USSR was largely based on the use of water composition in geochemical mineral exploration. This climatic zonation in chemical

character of natural waters was earlier recognized by Clark (1924) in his analysis and comparison of the average chemical composition of stream water from large river basins in various climatic belts.

Henri Schoeller of France was probably the first to carry out extensive, systematic calculations on chemical equilibrium and stability in the context of pH and redox potential (Eh) conditions for various rock types and environments (Schoeller, 1955, 1962). He applied the results in an extensive chemical facies classification, which was largely based on ion ratios, to identify chemical process and to explain observed groundwater compositions in terms of current conditions and paleo-hydrological evolution. The concept of chemical facies as a diagnostic tool to reflect the response of chemical processes within a hydrogeological framework, was also applied by Chebotarev. In addition mass-balance analysis of non-reactive solutes (notably chloride) and stable isotopes proved to be a powerful tool in groundwater budget studies, particularly to estimate groundwater recharge under arid conditions (e.g., Eriksson, 1976).

These applications of hydrochemistry in regional studies have been widely practised in the arid areas of North Africa and the Middle East in the 1950s and the 1960s, in projects, initiated by the UN-organizations, FAO and UNESCO.

In the United States, John D. Hem (1916–1994) compiled a wealth of data on the occurrence and properties of constituents from the huge number of reported water analysis, mainly from the United States (Hem, 1959, 1970). In addition, Hem and Cropper (1959) and Hem (1963) analyzed the stability of iron, manganese, and sulphur as functions of the environmental conditions characterized by the combined pH and Eh. The results were verified and confirmed from the field studies by Barnes and Back (1964) in the intensively investigated Floridian coastal aquifer. Studies on calcite solubility equilibria were carried out by William Back (1963). Around the same time, Garrels began his pioneering work on systematic application of well-known chemical principles in low-temperature aqueous solutions to study water–rock interaction. This led to the principles for a fundamental unifying understanding of hydrochemistry on the basis of chemical thermodynamics (Garrels, 1960). Back and Hanshaw (1965, 1970) applied the modern principles of equilibrium, thermodynamics, and interaction time in regional field studies. Georg Mattes (1973) wrote the first comprehensive textbook on hydrochemistry in western Europe; this German text was translated in English in 1982.

Modeling of combined transport and chemical reactions along a flow path is particularly complicated because it includes a distributed flow of water, matter, and energy, as well as physico-chemical and microbial processes in a heterogeneous environment. An integrated modeling approach only became feasible in the 1970s with the introduction of numerical simulations using powerful digital computers. Isotopic tracers proved to be invaluable tools to identify paleo-conditions and to calibrate the complex reactive-transport models. A well-known early hydrochemical transport and process model which accounts for mixing of waters is PHREEQE (Pakhurst et al., 1980), which marked the beginning of a new era in hydrochemistry. These developments were largely stimulated by pollution problems, and numerous studies on contaminant hydrochemistry have been published since the early 1980s.

1.3.5.2 Isotope Hydrology

Isotopes form a very special and powerful diagnostic tool, which has been used in groundwater studies since the 1950s. Initially, mainly short-lived radio-isotopes were added to surface water and groundwater as labels to trace pathways and to determine local parameters. However, because of the limited time and space scales of the experiments, as well as for safety reasons, this application was virtually abandoned in the 1960s and was replaced by the use of environmental isotopes. This opened the way for reconstruction of the “natural history” of water by determining its age and provenance, and to identify processes to which water had been subjected. The following outline is based mainly on reports by Letolle and Olive (2004) and Froehlich et al. (2004).

Important isotopes in hydrology, such as oxygen-18 (^{18}O) and deuterium (D), were already identified in the 1920s and 1930s, but the use of isotopes in hydrological research came into perspective only in the 1950s after the introduction of advanced mass-spectrometers. A catalogue of ^{18}O and D for waters of various origins was also established in the 1950s (Dansgaard, 1953; Epstein and Mayeda, 1953; Friedman, 1953). A further milestone was the discovery by Harmon Craig (1961) of the $^{18}\text{O}/\text{D}$ relation in precipitation, the

Global Meteorological Waterline, and his introduction of the use of *Standard Mean Ocean Water*. At the same time, analyses of thermal waters revealed their predominantly meteoric origin, and subsequent investigations showed that the juvenile component of waters associated with volcanic activity normally represents only a small fraction (Craig, 1963).

The application of radio-active tracers began in 1947 with the first radiocarbon dating by W.F. Libby (for which he was awarded the Nobel Prize in 1960); the first groundwater dating followed in 1957 by Karl-Otto Münnich and his team in Heidelberg. From 1952 to 1980 huge amounts of tritium were added to the atmosphere as a result of thermo-nuclear bomb tests. This made tritium the most important indicator to distinguish pre-1952 from post-1952 waters, with the various annual peaks in the groundwater used as markers for identifying its year of infiltration. The 1963-peak became the classical marker, when the tritium content of the atmosphere that year rose to about 3000 TU. Application of tritium, in combination with ^{18}O and D, has been particularly valuable to obtain insight to the complicated processes in the unsaturated zone. The decrease of the tritium concentration in the atmosphere after the 1980s, and the natural attenuation of the previous peaks, has reduced the value of this isotope.

Important for the stimulation of isotope hydrology was the establishment in 1958, in Vienna, of the International Atomic Energy Agency (IAEA) as an UN agency. One of its sub groups was the Isotope Hydrology Group (later changed to Section of Isotope Hydrology), initially headed by the Swede Eric Eriksson, with as mission to promote and support the development and practical application of nuclear and isotopic techniques in hydrology and related earth sciences. Of paramount importance was the establishment by IAEA of the Global Network of Isotopes in Precipitation, comprising of 150 selected WMO meteorological stations.

General application of isotopes in hydrology began in the early 1960s with, among others, Brown (1961) in Canada, Eriksson (1963) in Sweden, Münnich et al. (1967) in Germany, and Vogel (1967), and Mook (1968) in the Netherlands. Isotopes have been used worldwide from the 1970s in a variety of hydrological studies, including understanding the mechanisms of infiltration and evaporation, biodegradation, the provenance of water, the climatic conditions during infiltration, the separation of hydrograph components, salt water intrusion, hydro-geothermal processes, and dating groundwater in connection with paleohydrological evolution (notably in fossil aquifers and basins with low recharge). In addition, noble gas contents in groundwater have been used as recorders of water temperature during infiltration.

Application of environmental tracers, in combination with numerical transport modeling, particularly facilitated the direct determination of groundwater replenishment, including aspects like preferential flow paths and bypass flow. This applies specifically to (semi-)arid areas where the magnitude of recharge is far below the accuracy range of the evapotranspiration determination, so that establishing recharge as difference between rainfall and evaporation or by groundwater discharge analysis is normally not feasible (e.g., Zimmermann et al., 1967; Allison and Hughes, 1974; Dinçer et al., 1974). Several stimulating conferences on the use of isotopes in arid zone hydrology were organized by the IAEA. It was in response to this new perspective that UNESCO-IHP and the IAH launched in 1984 an international cooperation programme on groundwater recharge, targeting arid and semi arid areas. The main outcome of this initiative was a manual for understanding and estimate of natural recharge (Lerner et al., 1990).

Another very useful application of isotope tracers is in hydrochemical studies, notably for identifying physical processes and for calibrating inverse models. From the 1980s, isotope application was further boosted by the introduction of the accelerated mass spectrometer (AMS), which allows measurements of very low concentrations. This resulted in the use of, for example, ^{36}Cl for dating young as well as very old water, and a variety of “new” tracers to identify, for example, sources of contamination and biochemical degradation processes in landfill leachates.

1.4 Main Trends since the Early 1980s

The development of hydrogeological concepts and the mathematical and chemical tools to describe and analyze groundwater behavior, seem to have reached a mature stage by the early 1980s. The principal

remaining problem is the heterogeneity of the hydrogeological environment and the nonlinearity of the hydrogeological processes. Therefore, advances in the last quarter of the 20th century were dominated by the application of the numerical approach by digital computer simulation models. These made feasible the analysis of distributed transient flow problems in heterogeneous aquifers with complex boundary conditions, and to include such aspects as reactive transport, multi-phase flow, skeletal deformation, and microbial activity. This released groundwater studies from the necessity of drastic simplification and unrealistic assumptions in order to arrive at analytical solutions. Computer modeling, on the other hand, has often led to unrealistic simulation of reality by parameter generation through misleading optimization results.

The computer facilities and the shift of focus from quantity to quality and contamination studies has resulted in an explosion of transport and process models to explain present condition, to predict quality evolution, to design adequate monitoring schemes, and to develop remediation measures, including *in-situ* microbial techniques. Interest in aquifer behavior on a geologic time scale, including tectonic stress evolution and diagenetic processes on basin scale, has been stimulated by concern for the long-term behavior of radioactive waste repositories. Progress has been made in the 1980s in understanding the physical and geochemical behavior of an increasing number of environmental isotopes and other tracers, and in more refined detection techniques.

One of the main future hydrogeological challenges is to solve the problem of the geostatistical meaning of aquifer parameters and their scale dependency, including inverse modeling with its non-unique solutions (cf. Narasimhan, 1982, 1998). It is questionable however, to what detail we will be able to characterize the subsurface with its heterogeneous lithology and discontinue porous system. With respect to the interaction of groundwater and runoff generation is the challenge to develop distributed, physically based hydrological models at catchment scale. A principal problem here is the up scaling of hydrologic and geomorphologic parameters, and the synthesis of the various process components for nonlinear systems. Time-scale is another question; the need, for example, to isolate some types of waste for a very long period, requires prediction of aquifer system behavior on a time scale of hundreds of thousands of years.

Acknowledgments

The author expresses his gratitude to Ian Simmers, Peter Raats, and Garth van der Kamp for reading and amending the manuscript.

References

- Allison, G.B. and Hughes, M.W. 1974. Environmental tritium in the unsaturated zone: estimation of recharge to an unconfined aquifer. In *Proc. Sympos. Isotope Techn. in Groundwater Hydrology*, Vol. 1, pp. 57–72, IAEA, Vienna.
- Back, W. 1963. Preliminary result of a study of the calcium carbonate saturation of groundwater in central Florida. *Int. Assoc. Sci. Hydrol. Publ.* 8, 43–51.
- Back, W. 1981. Hydromythology and ethnohydrology in the New World. *Water Resour. Res.* 17, 257–287.
- Back, W. and Hanshaw, B.B. 1965. Chemical geohydrology. In *Advances in Hydroscience*, Vol. 2, pp. 49–109, Ven Te Chow, Ed. Academic Press, New York.
- Back, W. and Hanshaw, B.B. 1970. Comparison of chemical hydrogeology of the carbonate peninsulas of Florida and Yucatan. *J. Hydrology* 10, 330–368.
- Barenblatt, G.I., Zheltov, I.P., and Kochina, I.N. 1960. Basic concepts in the theory of seepage of homogeneous liquids in fissured rocks. *Sov. J. Appl. Math. Mech.* (Engl. transl.) 24, 852–864.
- Barnes, B.S. 1939. The structure of discharge-recession curves. *Trans. Am. Geophys. Union* 20, 721–725.
- Barnes, I. and Back, W. 1964. Geochemistry of iron-rich ground water of southern Maryland. *J. Geol.* 72, 435–447.
- Bear, J. 1961. On the tensor form of dispersion in porous media. *J. Geophys. Res.* 66, 1185–1197.
- Bear, J. 1972. *Dynamics of Fluids in Porous Media*. Elsevier, New York.

- Bear, J., Zaslavski, D. and Irmay, S. 1968. *Physical Principles of Water Percolation and Seepage*. UNESCO, Paris.
- Bevin, K.J. and Kirkby, M.J. 1979. A physical based variable contributing area model of basin hydrology. *Bull. Intern. Assoc. Hydrol. Sci.* 24, 43–69.
- Biot, M.A. 1941. General theory of three dimensional consolidation. *Appl. Phys.* 12, 155–164.
- Biswas, A.K. 1970. *History of Hydrology*. North-Holland, Amsterdam.
- Bögli, A. 1964. Mischungskorrosion; ein Beitrag zum Verkarstungsproblem. *Erdkunde* 18, 83–92.
- Bögli, A. 1980. *Karst Hydrology and Physical Speleology*. Springer, Berlin.
- Boulton, N.S. 1954. The drawdown of the water table under non-steady state conditions near a pumped well in an unconfined formation. *Proc. Inst. Civil Eng.* 3, 564–579.
- Boulton, N.S. and Streltsova, T.D. 1977. Unsteady flow to a pumped well in a fissured waterbearing formation. *J. Hydrology* 35, 257–270.
- Boussinesq, J. 1877. Essai sur la Théorie des Eaux Courantes. *Mém. présentés par divers savants à l'Acad. Sci.*, t. 23, 1–680.
- Boussinesq, J. 1904. Récherches théoriques sur l'écoulement des nappes d'eau infiltrées dans le sol. *J. Mathématiques Pures Appliquées* 10, 5–78; 363–394.
- Box, G.E.P. and Jenkins, G.M. 1970. *Time Series Analysis: Forecasting and Control*. Holden-Day, San Francisco.
- Bredehoeft, J.D. 1967. Response of well-aquifer systems to earth tides. *J. Geophys. Res.* 72, 3075–3087.
- Bredehoeft, J.D. and Papadopoulos, I.S. 1965. Rates of vertical groundwater movement estimated from earth's thermal profile. *Water Resour. Res.* 1, 325–328.
- Bredehoeft, J.D. and Hanshaw, B.B. 1968. On the maintenance of anomalous fluid pressure, thick sedimentary sequences. *Geol. Soc. Am. Bull.* 79, 1097–1106.
- Bredehoeft, J.D., Back, W., and Hanshaw, B.B. 1982. Regional ground-water flow concepts in the United States: historic perspective. In *Recent Trends in Hydrogeology*, Narasimha, T.N., Ed. Geol. Soc. Am. Special Paper 189, 297–316.
- Brown, R.M. 1961. Hydrology of tritium in the Ottawa valley. *Geochim. Cosmochim. Acta* 21, 199–216.
- Brown, G.O. 2002. Henry Darcy on the making of a law. *Water Resour. Res.* 38, 11.1–11.12.
- Brown, G.O., Garbrecht, J.D., and Hager, W.H. (Eds.). 2003. *Henry P.G. Darcy and Other Pioneers in Hydraulics*. Am. Soc. Civil Engineers, Reston, Virginia.
- Brutsaert, W. 2005. Afterword — A short historical sketch of theories about the water circulation on Earth. In *Hydrology — An Introduction*, Chap. 14, pp. 557–589. Cambridge University Press, Cambridge.
- Buckingham, E. 1907. Studies on the movement of soil moisture. *U.S. Dept. Agri. Bur. Soils, Bull.* 38.
- Buckham, A.F. and Cockfield, W.E. 1950. Gullies formed by sinking of the ground. *Am. J. Sci.* 248, 137–141.
- Burdon, D.J. 1977. Flow of fossil groundwater. *Q. J. Geol.* 10, 97–124.
- Burgers, J.M. 1926. Grondwaterstrooming in de omgeving van een net met kanalen. *De Ingenieur* 41, 657–665.
- Carlslaw, H.S. and Jaeger, J.C. 1921. *Introduction to the Mathematical Theory of the Conduction of Heat in Solids*. MacMillan, New York.
- Carlston, C.W. 1963. An early American statement of the Badon Ghyben–Herzberg principle of static fresh-water/salt-water balance. *Am. J. Sci.* 261, 88–91.
- Carman, P.C. 1937. Fluid flow through a granular bed. *Trans. Inst. Chem. Eng.* 15, 150–156.
- Carman, P.C. 1956. *Flow of Gases Through Porous Media*. Butterworth, London.
- Chamberlin, T.C. 1885. The requisite and qualifying conditions of artesian wells. *U.S. Geol. Survey 5th Ann. Rep.*, pp. 125–173.
- Chebotorev, I.I. 1955. Metamorphism of natural waters on the crust of weathering. *Geochim. Cosmochim. Acta* 8, 22–48; 137–170; 198–212.
- Childs, E.C. and Collis-George, N. 1950. The permeability of porous materials. *Proc. R. Soc.* 201, 392–405.
- CHO-TNO. 1980. Research on possible changes in the distribution of saline seepage in the Netherlands. *Comm. Hydrological Research-TNO, Proc. Inf.* 26, The Hague.

- Chorley, R.J. and Morgan, M.A. 1962. Comparison of morphometric features Unaka Mountains, Tennessee and North Carolina, and Dartmoor, England. *Geol. Soc. Am. Bull.* 73, 17–34.
- Chorley, R.J. and Kennedy, B.A. 1971. *Physical Geography, A System Approach*. Prentice-Hall, London.
- Chun, R.Y.D., Weber, E.M., and Mido, K. 1964. Computer tools for sound management of groundwater basins. *Int. Assoc. Sci. Hydrol. Publ.* 64, 424–437.
- Clark, F.W. 1924. The data of geochemistry, 5th ed. *U.S. Geol. Survey Bull.* 770.
- Collins, W.D. 1923. Graphic representation of analyses. *Ind. Eng. Chem.* 15, 394.
- Cooper, H.H. and Rorabaugh, M.I. 1963. Groundwater movement and bank storage due to flood stages in surface streams. *U.S. Geol. Survey Water-Supply Paper* 1536-J.
- Craig, H. 1961. Isotopic variations in meteoric waters. *Science* 133, 1702–1703.
- Craig, H. 1963. The isotope geochemistry of water and carbon in geothermal areas. In *Nuclear Geology on Geothermal Areas*, pp. 17–53, Tongiorgi, E., Ed. Pisa.
- Cvijić, J. 1893. Das Kartsphänomen. *Geogr. Abhandl. (Wien)* 5, 217–329.
- Cvijić, J. 1918. Hydrographie souterraine et évolution morphologique du karst. *Rec. Trav. Inst. Géogr. Alpine (Grenoble)* 6, 375–426.
- Dachler, R. 1936. *Grundwasserströmung*. Springer-Verlag, Wien.
- D'Andrimont, R. 1910. Circulation de l'eau dans le sol et le sous-sol. *Extrait du Bull. Soc. Belge de Géologie*, t. XXIV, exposé de l'Exposition Internationale de Bruxelles, 1910.
- Dansgaard, W. 1953. The abundance of O¹⁸ isotopes in atmospheric water and water vapour *Tellus* 5, 461–469.
- Darcy, H.P.G. 1856. *Les Fontaines Publiques de la Ville de Dijon*. V. Dalmont, Paris.
- Darton, N.H. 1909. Geology and underground waters of South Dakota. *U.S. Geol. Survey Water-Supply Paper* 227.
- Daubrée, A. 1887. *Les Eaux Souterraines, aux Époque Anciennes et à l'Époque Actuelle*. (3 volumes), Dunod, Paris.
- Davis, W.M. 1930. Origin of limestone caverns. *Geol. Soc. Am. Bull.* 41, 475–628.
- Davis, S.N. and De Wiest, R.J.M. 1966. *Hydrogeology*. Wiley, New York.
- De Glee, G.J. 1930. *Over Grondwaterstroomingen bij Wateronttrekking door middel van Putten*. Ph.D. dissertation, Delft University of Technology, Waltman, Delft.
- De Josselin de Jong, G. 1960. Singularity distributions for the analysis of multiple flow through porous media. *J. Geophys. Res.* 65, 3739–3758.
- Delhomme, J.P. 1979. Spatial variability and uncertainty in groundwater flow parameters: a geostatistic approach. *Water Resour. Res.* 15, 269–280.
- De Vos, H.C.P. 1929. Eenige beschouwingen omtrent de verweekingslijn in aarden dammen. *De Waterstaats-Ingenieur* 17, 335–354.
- De Vries, J.J. 1977. The stream network in the Netherlands as a groundwater discharge phenomenon. *Geologie & Mijnbouw* 56, 103–122.
- De Vries, J.J. 1981. Fresh and salt groundwater in the Dutch coastal area in relation to geomorphological evolution. *Geologie & Mijnbouw* 60, 363–368.
- De Vries, J.J. 2004. From speculation to science: the founding of groundwater hydrology in the Netherlands. In *Dutch Pioneers of the Earth Sciences*, pp. 139–164, Touret, J.L.R. and Visser, R.P.W., Eds. Royal Netherlands Academy of Arts and Sciences, Amsterdam.
- Dinçer, T., Al-Mugrini, A., and Zimmermann, U. 1974. Study of the infiltration and recharge through sand dunes in arid zones with special reference to the stable isotopes and thermonuclear tritium. *J. Hydrol.* 23, 79–109.
- Domenico, P.A. 1972. *Concepts and Models in Groundwater Hydrology*. McGraw-Hill, New York.
- Dooge, J.C.I. 2004. Background to modern hydrology. In *The Basis of Civilization — Water Science?* Rodda, J.C. and Ubertini, L., Eds. Intern. Assoc. Hydrol. Sci. Publ. 286, 3–12.
- Dos Santos Junior, A.G. and Youngs, E.G. 1969. A study of the specific yield in land-drainage situations. *J. Hydrol.* 8, 59–81.

- Downing, R.A., Eastwood, W., and Rushton, K.R. 2004. Norman Savage Boulton (1899–1984): civil engineer and groundwater hydrologist. In *200 Years of British Hydrogeology*, Mather, J.D., Ed. Geological Society Publishing House, Bath.
- Drabbe, J. and Badon Ghijben, W. 1889. Nota in verband met de voorgenomen putboring nabij Amsterdam. *Tijdschr. Kon. Inst. Ingenieurs, Verhandelingen* 1888/1889, 8–22.
- Du Commun, J. 1818. On the causes of fresh water springs, fountains, etc. *Am. J. Sci., 1st Ser.* 14, 174–176.
- Dumm, L.D. 1954. Drain spacing formula. *Agri. Eng.* 35, 726–730.
- Dunne, T. 1970. Runoff production in a humid area. *U.S. Dept. Agri. Rep. ARS* 41–160.
- Dunne, T. and Black, R.D. 1970a. An experimental investigation of runoff production in permeable soils. *Water Resour. Res.* 6, 478–490.
- Dunne, T. and Black, R.D. 1970b. Partial area contributions to storm runoff in a small New England watershed. *Water Resour. Res.* 6, 1296–1311.
- Dupuit, A.J.E.J. 1863. *Études Théoriques et Pratiques sur le Mouvement des Eaux dans les Canaux Découvertes et à Travers les Terraines Perméables*. 2^e éd., Dunod, Paris.
- Edelman, J.H. 1947. *Over de Berekening van Grondwaterstroomingen*. Ph.D. dissertation, Delft University of Technology.
- Edelman, J.H. 1972. *Groundwater Hydraulics of Extensive Aquifers*. Intern. Inst. Land Reclamation and Improvement (ILRI), Bull. 13, Wageningen.
- Emmons, W.H. and Harrington, G.L. 1913. A comparison of waters of mines and of hot springs. *Econ. Geol.* 8, 653–669.
- Epstein, S. and Mayeda, T. 1953. Variations of the $^{18}\text{O}/^{16}\text{O}$ ratio in natural waters. *Geochim. Cosmochim. Acta* 4, 213–224.
- Eriksson, E. 1963. The possible use of tritium for estimating groundwater storage. *Tellus* 10, 472–478.
- Eriksson, E. 1976. The distribution of salinity in ground waters in the Delhi region and recharge rates of ground water. In *Interpretation of Environmental Isotopes and Hydrochemical Data in Groundwater Hydrology*, pp. 171–178. Workshop IAEA, Vienna (1975).
- Ernst, L.F. 1956. Calculations of the steady flow of groundwater in vertical cross-sections. *Neth. J. Agri. Sci.* 4, 102–131.
- Ernst, L.F. 1962. *Grondwaterstromingen in de Verzadigde Zone en hun Berekeningen bij Aanwezigheid van Horizontale Evenwijdige Open Leidingen*. Ph.D. dissertation, Utrecht University, Pudoc, Wageningen.
- Ernst, L.F. 1978. Drainage of undulating sandy soils with high groundwater tables. *J. Hydrology* 39, 1–50.
- Farvolden, G.N. 1963. Geologic controls on groundwater storage and base flow. *J. Hydrology* 1, 219–249.
- Fayers, F.J. and Sheldon, J.W. 1962. The use of high-speed digital computers in the study of the hydrodynamics of geological basins. *J. Geophys. Res.* 67, 2421–2431.
- Ferris, J.G. 1951. Cyclic fluctuations of water levels as a basis for determining aquifer transmissibility. *Intern. Assoc. Hydrol. Sci. Publ.* 33, 148–155.
- Forchheimer, Ph. 1886. Über die Ergiebigkeit von Brunnen-Anlagen und Sickerschlitten. *Zeitschr. des Architekten- und Ingenieur-Vereines* 32, 539–564.
- Forchheimer, Ph. 1898. Grundwasserspiegel bei Brunnenanlagen. *Zeitschrift Öesterr. Ingenieur- und Architekten-Vereines*. 44, 629–635.
- Forchheimer, Ph. 1914. *Hydraulik*. Teubner, Leipzig.
- Freeze, R.A. 1969. The mechanism of natural groundwater recharge and discharge: 1. One-dimensional, vertical unsteady, unsaturated flow above a recharging or discharging groundwater flow system. *Water Resour. Res.* 5, 153–171.
- Freeze, R.A. 1971. Three-dimensional transient, saturated–unsaturated flow in a groundwater basin. *Water Resour. Res.* 7, 347–366.
- Freeze, R.A. 1972a. Role of subsurface flow in generating runoff: 1. Baseflow contribution to channel flow. *Water Resour. Res.* 8, 609–623.
- Freeze, R.A. 1972b. Role of subsurface flow in generating runoff: 2. Upstream source areas. *Water Resour. Res.* 8, 1272–1283.

- Freeze, R.A. 1975. A stochastic conceptual analysis of one-dimensional groundwater flow in non-uniform homogeneous media. *Water Resour. Res.* 11, 725–741.
- Freeze, R.A. and Witherspoon, P.A. 1966. Theoretical analysis of regional groundwater flow: 1. Analytical and numerical solutions to the mathematical model. *Water Resour. Res.* 2, 641–656.
- Freeze, R.A. and Witherspoon, P.A. 1967. Theoretical analysis of regional groundwater flow: 2. Effect of water-table configuration and subsurface permeability. *Water Resour. Res.* 3, 623–634.
- Freeze, R.A. and Cherry, J.A. 1979. *Groundwater*. Prentice-Hall, Englewood Cliffs, N.J.
- Friedman, I. 1953. Deuterium content of natural waters and other substances. *Geochim. Cosmochim. Acta* 4, 89–103.
- Fritz, P., Cherry, J.A., Weyer, K.U., and Sklash, M. 1976. Storm runoff analyses using environmental isotopes and major ions. In *Interpretation of Environmental Isotope and Hydrochemical Data in Groundwater Hydrology*, pp. 111–130. Workshop IAEA, Vienna (1975).
- Froehlich, K., Gonfianti, R., and Aggarwal, P. 2004. Isotope hydrology at IAEA: history and activities. In *The Basis of Civilization — Water Science?* Rodda, J.C. and Ubertini, L., Eds. Intern. Assoc. Hydrol. Sci. Publ. 286, 125–132.
- Gardner, W. 1920. A capillary transmission constant and methods of determining it experimentally. *Soil Sci.* 10, 103–126.
- Gardner, W.R. 1958. Some steady state solutions of the unsaturated moisture flow equation with application to evaporation from a water table. *Soil Sci.* 85, 228–232.
- Garmanov, I.V. 1958. Fundamental principles of hydrochemical zoning of underground waters in the European part of the Soviet Union. *Sympos. Ground Water, Calcutta (1955), Proc.* 4, 293–302.
- Garnier, F. 1822. *De l'Art du Fontenier Sondeur et des Puits Artésiens*. Huzard, Paris.
- Garrels, R.M. 1960. *Mineral Equilibria at Low Temperature and Pressure*. Harper and Bross, New York.
- Geertsema, J. 1957. The effect of fluid pressure decline on volumetric changes in porous rocks. *Trans. Am. Inst. Min. Metall. Petr. Eng.* 210, 331–340.
- Gelhar, L.W. 1977. Effects of hydraulic conductivity variations on groundwater flows. *Proc. 2nd Int. Symp. Stochastic Hydraulics, Lund (Sweden, 1976)*, pp. 409–428, Hjort, P., Jönsson, L., and Larsen, P., Eds. Intern. Assoc. Hydraulic Research, Water Res. Publ., Fort Collins, Colorado.
- Glover, R.E. 1959. The pattern of fresh-water flow in coastal aquifer. *J. Geophys. Res.* 64, 457–459.
- Graton, L.C. and Fraser, H.J. 1935. Systematic packing of spheres, with particular relation to porosity and permeability. *J. Geol.* 43, 785–909.
- Green, W.H. and Ampt, G.A. 1911. Studies on soil physics, part 1: the flow of air and water through soils. *J. Agric. Sci.* 4, 1–24.
- Gringarten, A.C. and Witherspoon, P.A. 1972. A method of analysing pump test data from fractured aquifers. *Proc. Sympos. Rock Mechanics*, Vol. 3-B, pp. 1–9, Int. Soc. Rock Mech. and Int. Assoc. Eng. Geol., Stuttgart.
- Grund, A. 1903. Die Karsthydrographie: Studien aus Westbosnien. *Geogr. Abhandl. (Wien)* 7, 103–200.
- Haines, W.B. 1930. Studies in the physical properties of soils V: the hysteresis effect in capillary properties and the modes of moisture distribution associated therewith. *J. Agric. Sci.* 20, 97–116.
- Hálek, V. and Švec, J. 1979. *Groundwater Hydraulics*. Academia, Prague.
- Hall, F.R. 1968. Base flow recession — a review. *Water Resour. Res.* 4, 973–983.
- Hantush, M.S. and Jacob, C.E. 1955. Non steady radial flow in an infinite leaky aquifer. *Trans. Am. Geophys. Union* 36, 95–100.
- Hantush, M.S. 1960. Modification of the theory of leaky aquifers. *J. Geophys. Res.* 65, 3713–3725.
- Hantush, M.S. 1964. Hydraulics of wells. In *Advances in Hydroscience*, Vol. 1, pp. 281–432, Ven Te Chow, Ed. Academic Press, New York.
- Hantush, M.S. 1967. Flow to wells in aquifers separated by a semi-pervious layer. *J. Geophys. Res.* 72, 1709–1720.
- Harr, M.E. 1962. *Groundwater and Seepage*. McGrawHill, New York.
- Hazen, A. 1893. Some physical properties of sands and gravels with special reference to their use in filtration. *Massachusetts State Board of Health, 24th Ann. Rep.*, pp. 539–556.

- Helm, D.C. 1982. Conceptual aspects of subsidence to fluid withdrawal. In *Recent Trends in Hydrogeology*, Narasimhan, T.N. Ed. Geol. Soc. Am. Special Paper 189, 103–139.
- Hem, J.D. 1959, 1970. Study and interpretation of the chemical characteristics of natural water. *U.S. Geol. Survey Water-Supply Paper* 1473.
- Hem, J.D. and Cropper, W.H. 1959. Survey of ferrous–ferric equilibria and redox potentials. *U.S. Geol. Survey Water-Supply Paper* 1459-A.
- Hem, J.D. 1963. Chemical equilibria and rates of manganese oxidation. *U.S. Geol. Survey Water-Supply Paper* 1667-A.
- Herzberg, A. 1901. Die Wasserversorgung einiger Nordseebäder. *Journal für Gas- beleuchtung und Wasserversorgung* 44, 815–819; 842–844.
- Hewlett, J.D. 1961. Soil moisture as a source of base flow from steep mountain watersheds. *Southeastern Forest Experimental Station, Paper* 132. U.S. Forest Service, Ashville, NC.
- Hewlett, J.D. and Hibbert, A.R. 1963. Moisture and energy conditions within a sloping soil mass during drainage. *J. Geophys. Res.* 68, 1081–1087.
- Hewlett, J.D. and Hibbert, A.R. 1967. Factors affecting the response of small watersheds to precipitation in humid areas. In *Forest Hydrology*, pp. 275–290, Sopper, W.E. and Lull, H.W., Eds. Pergamon, Oxford.
- Hill, R.A. 1940. Geochemical patterns in Coachella Valley, Calif. *Trans. Am. Geophys. Union* 21, 46–49.
- Holzmüller, G. 1882. *Einführung in die Theorie der Isogonalen Verwandtschaften und der Konformen Abbildungen*. Leibzig.
- Hooghoudt, S.B. 1940. Algemeene beschouwingen van het probleem van de detailontwatering en de infiltratie door middel van parallel loopende drains, greppels, slooten en kanalen. In *Bijdrage tot de Kennis van Eenige Natuurkundige Grootheden van den Grond*, 7, Versl. Landbouwk. Onderz. 46 B, 515–707, The Hague.
- Horton, R.E. 1933. The role of infiltration in the hydrologic cycle. *Trans. Am. Geophys. Union* 14, 446–460.
- Horton, R.E. 1945. Erosional development of streams and their drainage basins: hydrological approach to quantitative geomorphology. *Geol. Soc. Am. Bull.* 56, 275–370.
- Hoyt, W.G. 1936. Studies of relations between rainfall and run-off in the United States. *U.S. Geol. Survey Water-Supply Paper* 772.
- Hubbert, M.K. 1940. The theory of ground-water motion. *J. Geology* 48, 785–944.
- Hubbert, M.K. and Rubey, W.W. 1959. Role of fluid pressure in mechanics of overthrust faulting. *Bull. Geol. Soc. Am.* 70, 115–166.
- Huisman, L. and Kemperman, J. 1951. Bemaling van spanningsgrondwater. *De Ingenieur* 62 B13, 29–35.
- Hursh, C.R. 1936. Storm-water and absorption. Contribution to Report of Committee on absorption and transpiration 1935–1936. *Trans. Am. Geophys. Union* 17, 269–302.
- Hursh, C.R. and Brater, E.F. 1941. Separating storm hydrographs from small drainage areas into surface and subsurface flow. *Trans. Am. Geophys. Union* 22, 721–725.
- Hurst, W. 1934. Unsteady flow of fluids in oil reservoirs. *J. Appl. Phys.* 5, 20–30.
- Issar, A.S. 1990. *Water Shall Flow from the Rock — Hydrogeology and Climate in the Lands of the Bible*. Springer-Verlag, Berlin.
- Jacob, C.E. 1940. On the flow of water in an elastic artesian aquifer. *Trans. Am. Geophys. Union* 21, 574–586.
- Jacob, C.E. 1946. Radial flow in a leaky artesian aquifer. *Trans. Am. Geophys. Union* 27, 198–208.
- Jacob, C.E. 1950. Flow of groundwater. In *Engineering Hydraulics*, pp. 321–386, Rouse, H., Ed. Wiley, New York.
- Jones, A. 1971. Soil piping and stream initiation. *Water Resour. Res.* 7, 602–610.
- Karplus, W.J. 1958. *Analog Simulation*. McGraw-Hill, New York.
- Katzer, E. 1909. *Karst und Karsthydrographie*. Zur Kunde der Balkan Halbinsel 8, Serajevo-Wien.
- Keen, B.A. 1927. The limited role of capillarity supplying water to the plant roots. *Proc. 1st Int. Congr. on Soil Science*, pp. 504–511.
- Keilhack, K. 1912. *Grundwasser und Quellenkunde*. Bornträger, Berlin.

- King, F.H. 1899. Principles and conditions of the movements of groundwaters. In *19th Annual Report of the U.S. Geological Survey*, part 2, pp. 59–295.
- Kirkby, M.J. and Chorley, R.J. 1967. Throughflow, overland flow, and erosion. *Bull. Int. Assoc. Sci. Hydrol.* 12, 5–21.
- Kirkham, D. 1954. Seepage of artesian and surface water into drain tubes in stratified soil. *Trans. Am. Geophys. Union* 35, 775–790.
- Kirkham, D. 1967. Explanation of paradoxes in Dupuit–Forchheimer seepage theory. *Water Resour. Res.* 3, 609–622.
- Klute, A. 1952. A numerical method for solving the flow equation for water in unsaturated material. *Soil Sci.* 73, 105–116.
- Kohout, F.A. 1960. Cyclic flow of salt water in the Biscayne aquifer of Southeastern Florida. *J. Geophys. Res.* 65, 2133–2141.
- Kooper, J. 1914. Beweging van het water in de bodem bij onttrekking door bronnen. *De Ingenieur* 29, 697–706; 710–716.
- Kozeny, J. 1927. Über kapillare Leitung des Wassers im Boden. *Sitzungsber., Akad. Wiss. Wien*, 136, 271–306.
- Kraijenhoff van de Leur, D.A. 1958. A study of non-steady groundwater flow with special reference to a reservoir-coefficient. *De Ingenieur* 70, 87–94.
- Kruseman, G.P. and De Ridder, N.A. 1970. *Analysis and Evaluation of Pumping Test Data*. Intern. Inst. Land Reclamation and Improvement (ILRI), Bull. 11, Wageningen.
- Lee, C.H. 1914. Determination of the safe-yield of under-ground reservoirs of the closed basin type. *Proc. Am. Soc. Civil Eng.* 40, 817–887.
- Lehmann, O. 1932. *Die Hydrographie des Karstes*. Franz Deutike, Wien.
- Letolle, R. and Olive, Ph. 2004. A short history of isotopes in hydrology. In *The Basis of Civilization — Water Science?* Rodda, J.C. and Ubertini, L., Eds. Intern. Assoc. Hydrol. Sci. Publ. 286, 49–66.
- Lerner, D.N., Issar, A.S., and Simmers, I. 1990. *Groundwater Recharge — A Guide to Understanding and Estimating Natural Recharge*. Int. Assoc. Hydrogeologists, Vol. 8, Heise, Hannover.
- Lueger, O. 1883. *Theorie der Bewegung des Grundwassers in den Alluvionen der Flussgebiete*. Stuttgart.
- Luszynski, N.J. 1961. Head and flow of groundwater of variable density. *J. Geophys. Res.* 66, 4247–4256.
- Maasland, M. 1957. Soil anisotropy and land drainage. In *Drainage of Agricultural Lands*, Agr. Monographs 7, pp. 216–287, Luthin, J.L., Ed. Am. Soc. Agr., Madison, WI.
- Maillet, E. 1905. *Essais d'Hydraulique Souterraine et Fluviale*. Herman, Paris.
- Martel, E.A. 1894. *Les Abîmes, les Eaux Souterraines, les Cavernes, les Sources, la Spéléologie*. Delgrave, Paris.
- Martel, E.A. 1921. *Nouveau Traité des Eaux Souterraines*. Doin, Paris.
- Mattes, G. 1973. *Die Beschaffenheit des Grundwassers*. Bornträger, Berlin. (English translation: *The Properties of Groundwater*, Wiley, New York, 1982.)
- Mazure, J.P. 1943. Enkele vergelijkende berekeningen betreffende de gevolgen van boven- en diepwateronttrekking in het duingebied. *Water* 27, 117–124.
- Mead, D.W. 1919. *Hydrology*. McGraw-Hill, New York.
- Meinzer, O.E. 1923a. The occurrence of ground water in the United States. *U.S. Geol. Survey Water-Supply Paper* 489.
- Meinzer, O.E. 1923b. Outline of ground-water hydrology. *U.S. Geol. Survey Water-Supply Paper* 494.
- Meinzer, O.E. 1928. Compressibility and elasticity of artesian aquifers. *Econ. Geol.* 23, 263–291.
- Meinzer, O.E. and Hard, H.H. 1925. The artesian water-supply of the Dakota Sandstone in North Dakota. *U.S. Geol. Survey Water-Supply Paper* 520-E.
- Meinzer, O.E., Ed. 1942. *Hydrology*. Dover, New York.
- Mendenhall, W.C. 1905. The hydrology of the San Bernardo Valley, California. *U.S. Geol. Survey Water-Supply Paper* 142.
- Meyboom, P. 1966. Unsteady groundwater flow near a willow ring in hummocky moraine. *J. Hydrology* 4, 38–62.
- Meyer, A.F. 1917. *The Elements of Hydrology*. Wiley, New York.

- Monteith, J.L. 1965. Evaporation and environment. *Proc. Symp. Exp. Biol.* 19, 205–234.
- Mook, W.G. 1968. *Geochemistry of Stable Carbon and Oxygen of Natural Waters in the Netherlands*. Ph.D. dissertation, Groningen University, Netherlands.
- Münnich, K.O., Roether, W. and Thilo, L. 1967. Dating of groundwater with tritium and 14-C. In *Isotopes in Hydrology*, pp. 305–320. IAEA, Vienna.
- Muskat, M. 1934. The flow of compressible fluids through porous media and some problems of heat conduction. *J. Appl. Phys.* 5, 71–94.
- Muskat, M. 1937. *The Flow of Homogeneous Fluids Through Porous Media*. McGraw-Hill, New York.
- Narasimhan, T.N. (Ed.). 1982. Recent trends in hydrogeology. *Geol. Soc. Am. Special Paper* 189.
- Narasimhan, T.N. 1982. Preface to *Recent Trends in Hydrogeology*, Narasimhan, T.N., Ed. *Geol. Soc. Am. Special Paper* 189, v–viii.
- Narashiman, T.N. 1998. Hydraulic characterization of aquifers, reservoir rocks, and soils: a history of ideas. *Water Resour. Res.* 34, 33–46.
- Pakhurst, D.L., Thorstenson, D.C., and Plumber, L.N. 1980. PHREEQE — A computer program for geochemical calculations. *U.S. Geol. Survey Water Resources Inv.* 80–96.
- Palmer, C. 1911. The geochemical interpretation of water analyses. *U.S. Geol. Survey Bull.* 479.
- Pennink, J.M.K. 1904. De prise d'eau der Amsterdamsche duinwaterleiding. *De Ingenieur* 19, 213–223.
- Pennink, J.M.K. 1905. Over de beweging van grondwater. *De Ingenieur* 20, 482–492.
- Pinder, G.F. and Jones, J.F. 1969. Determination of groundwater component of peak discharge from chemistry of total runoff. *Water Resour. Res.* 5, 438–445.
- Penman, H.L. 1948. Natural evaporation from open water, bare soil and grass. *Proc. R. Soc. A* 193, 120–145.
- Philip, J.R. 1969. Theory of infiltration. In *Advances in Hydro Science*, Vol. 5, pp. 215–305, Ven Te Chow, Ed. Academic Press, New York.
- Piper, A.M. 1944. A graphic procedure in the geochemical interpretation of water analyses. *Trans. Am. Geophys. Union* 25, 914–923.
- Poland, J.F. 1961. The coefficient of storage in a region of major subsidence caused by compaction of an aquifer system. *U.S. Geol. Survey Profess. Paper* 424–B.
- Poland, J.F. and Davis, G.H. 1969. Land subsidence due to the withdrawal of fluids. *Rev. Eng. Geol.* 11, 187–269. *Geol. Soc. Am.*, Boulder, CO.
- Polubarnova-Kochina, P.Ya. 1952. *Theory of Groundwater Flow* (in Russian). Translated by Roger J.M. de Wiest, Princeton University Press, Princeton, NJ, 1962.
- Prinz, E. 1919. *Hydrologie*. Springer-Verlag, Berlin.
- Raats, P.A.C. 1969. The effect of finite response upon the propagation of oscillations of fluids in porous media. *Zeitschrift für angewandte Math. und Physik* 20, 936–946.
- Raats, P.A.C. and Van der Ploeg, R.R. 2005. Hooghoudt, Symen Barend. In *Encyclopedia of Soils in the Environment*, Vol. 2, pp. 188–194, Hillel, D., Ed. Elsevier, Oxford.
- Remson, I., Hornberger, G.M., and Molz, F.J. 1971. *Numerical Methods in Subsurface Hydrology*. Wiley, New York.
- Renick, B.C. 1924. Base exchange in groundwater as illustrated in Montana. *U.S. Geol. Survey Water-Supply Paper* 520–D, pp. 53–72.
- Richards, L.A. 1931. Capillary conduction of liquids through porous mediums. *Physics* 1, 318–333.
- Richardson, L.F. 1911. The approximate arithmetical solution by finite differences with the application stresses in masonry dams. *Phil. Trans. R. Soc. A* 210, 307–357.
- Rijtema, P.E. 1965. *An Analysis of Actual Evapotranspiration*. Ph.D. dissertation, Wageningen University, *Agric. Res. Rep.* 659, 1–107.
- Rogers, G.S. 1917. Chemical relations of the oil field waters in San Joaquin Valley, California. *U.S. Geol. Survey Bull.* 653.
- Rosenstein, J.S., Moore, J.E., Lohman, S.W., and Chase, E.B. 1976. *200 Years of Hydrogeology in the United States*. National Well Assoc., Dublin, OH.
- Rubin, J. and Steinhardt, R. 1963. Soil water relations during rain infiltration. I. Theory. *Soil Sci. Soc. Am. Proc.* 27, 246–251.

- Rubin, J., Steinhardt, R., and Reiniger, P. 1964. Soil water relations during rain infiltration. II. Moisture content profiles during rains of low intensities. *Soil Sci. Soc. Am. Proc.* 28, 1–5.
- Russel, W.L. 1928. The origin of artesian pressure. *Econ. Geol.* 23, 132–157.
- Santing, G. 1957. A horizontal scale model based on the viscous flow analogy of studying groundwater flow in an aquifer having storage. *Int. Assoc. Sci. Hydrol. Publ.* 44, 105–114.
- Schoeller, H. 1935. Utilité de la notion des échanges de bases pour comparaison des eaux souterraines. *Bull. Soc. Géol. Fr.* 5, t. 5, 651–657.
- Schoeller, H. 1941. L'influence du climat sur la composition chimique des eaux souterraines vadosees. *Bull. Soc. Géol. Fr.* 5, t. 11, 267–289.
- Schoeller, H. 1955. *Géochimie des Eaux Souterraines. Application aux Eaux de Gisements de Pétrole*. Technip., Vol. 1, 1956, Paris.
- Schoeller, H. 1962. *Les Eaux Souterraines*. Masson, Paris.
- Schofield, R.K. 1935. The pH of the water in soil. *3rd Int. Soil Sci. Congres, Oxford, Trans.* Vol. 2, pp. 37–48.
- Schultze, J. 1924. *Die Grundwasserabsenkung in Theorie und Praxis*. Springer, Berlin.
- Seelheim, F. 1880. Methoden zur Bestimmung der Durchlässigkeit des Bodens. *Zeitschr. Anal. Chem.* 19, 387–418.
- Silin-Bekchurin, A.I. 1951. *Hydrogéologie Spécial*. Moscow.
- Skempton, A.W. 1944. Notes on the compressibility of clays. *Q. J. Geol. Soc. Lond.* C, 119–135.
- Slichter, C.S. 1899. Theoretical investigations of the motion of groundwater. In *19th Annual Report of the U.S. Geological Survey*, part 2, pp. 295–384.
- Slichter, C.S. 1902. The motions of underground waters. *U.S. Geol. Survey Water-Supply Paper* 67, pp. 11–106.
- Slichter, C.S. 1905. Field measurements of the rate of the movement of underground water. *U.S. Geol. Survey Water-Supply Paper* 140.
- Smith, N. 1976. *Man and Water*. Davies, London.
- Southwell, R.V. 1940. *Relaxation Methods in Engineering Science*. Oxford University Press, London.
- Stallman, R.W. 1963. Computation of groundwater velocity from temperature data. In *Methods of Collecting and Interpreting Groundwater Data*, Bentall, R., Ed. U.S. Geol. Survey Water-Supply Paper 1544, pp. H36–H46.
- Steggewitz, J.H. 1933. *De Invloed van de Getijbewegingen van Zeeën en Getijrivieren op de Stijghoogte van het Grondwater*. Ph.D. dissertation, Delft University of Technology, Meinema, Delft.
- Stephens, D.B. and Ankeny, M.D. 2004. The missing link in the historical development of hydrogeology. *Ground Water* 42, 304–309.
- Stiff, H.A. 1951. The interpretation of chemical water analysis by means of patterns. *J. Petr. Techn.* 3, 15–17.
- Stuyfzand, P.J. 1993. *Hydrochemistry and Hydrology of the Coastal Dune Area of the Western Netherlands*. Ph.D. dissertation, Vrije Universiteit Amsterdam, KIWA Research, Nieuwegein, Netherlands.
- Terzaghi, K. 1925. *Erdbaumechanik auf Bodenphysikalischer Grundlagen*. Deuticke, Leipzig.
- Terzaghi, K. 1929. Origin of artesian pressure. *Econ. Geol.* 24, 96–97.
- Terzaghi, K. 1943. *Theoretical Soil Mechanics*. Wiley, New York.
- Terzaghi, K. and Peck, R.B. 1948. *Soil Mechanics in Engineering Practice*. Wiley, New York.
- Theis, C.V. 1935. The relation between the lowering of the piezometric surface and the rate and duration of discharge of a well using groundwater storage. *Trans. Am. Geophys. Union* 16, 519–524.
- Thiem, A. 1870. Die Ergiebigkeit artesischer Bohrlocher, Schachtbrunnen, und Filtergallerien. *J. Gasbeleuchtung Wasserversorgung* 14, 450–467.
- Thiem, G. 1906. *Hydrologische Methoden*. Ph.D. dissertation, Techn. Hochschule Stuttgart, Gebhardt, Leipzig.
- Thornthwaite, C.W. 1948. An approach towards rational classification of climate. *Geogr. Rev.* 38, 55–94.
- Thornthwaite, C.W. and Mather, J.R. 1955. The water balance. *Publ. Climatology* 8, 1–86.
- Tillmans, J. 1932. *Die Chemische Untersuchungen von Wasser und Abwasser*, 2nd ed. Knapp, Halle.

- Tison, L.J. 1951. Au sujet des fluctuations des nappes aquifères étudiées. *Intern. Assoc. Sci. Hydrol. Publ.* 33, 195–201.
- TNO-NITG. 2004. De waterput. *Informatie*, November 2004, pp. 27–30. Netherlands Institute of Applied Geosciences, Utrecht.
- Todd, D.K. 1959. *Groundwater Hydrology*. Wiley, New York.
- Todd, D.K. and Huisman, L. 1959. Groundwater flow in The Netherlands coastal dunes. *J. Hydraul. Div. Am. Soc. Civil Eng.* 85, Hy 7, 63–82.
- Tolman, C.F. 1937. *Groundwater*. McGraw-Hill, New York.
- Tóth, J. 1962. A theory of groundwater motion in small basins in Central Alberta, Canada. *J. Geophys. Res.* 67, 4375–4387.
- Tóth, J. 1963. A theoretical analysis of groundwater flow in small drainage basins. *J. Geophys. Res.* 68, 4795–4812.
- Turc, L. 1954. Calcul du bilan de l'eau. Évaluation en fonction des précipitations et des températures. *Intern. Assoc. Sci. Hydrol. Publ.* 38, 188–202.
- Van Deemter, J.J. 1950. Theoretische en numerieke behandeling van ontwaterings- en infiltratie-stromingsproblemen. *Versl. Landbouwk. Onderz.* 56(7), 's-Gravenhage.
- Van der Kamp, G.S.J.P. 1973. *Periodic Flow of Groundwater*. Ph.D. dissertation, Vrije Universiteit Amsterdam, Ropodi, Amsterdam.
- Van Oldenborgh, J. 1916. Mededeelingen omtrent de uitkomsten van door Rijksbureau voor Drinkwatervoorziening ingestelde geo-hydrologische onderzoeken in verschillende duingebieden. *De Ingenieur* 31, 458–467; 474–498.
- Van Schilfgaarde, J., Kirkham, D., and Frevert, R.K. 1956. Physical and mathematical theories of tile and ditch drainage and their usefulness in design. *Iowa Agric. Exp. Station, Res. Bull.* 436.
- Veatch, A.C. 1906. Fluctuations of the water level in wells, with special reference to Long Island, New York. *U.S. Geol. Survey Water-Supply Paper* 155.
- Veihmeyer, F.J. and Hendrickson, A.H. 1931. The moisture equivalent as a measure of the field capacity of soils. *Soil Sci.* 32, 181–193.
- Verruijt, A. 1969. An interface problem with a source and a sink in the heavy fluid. *J. Hydrology* 8, 197–206.
- Verruijt, A. 1970. *Theory of Groundwater Flow*. Macmillan, London.
- Versluys, J. 1916. Chemische werking in den ondergrond der duinen. *Versl. Meded. Kon. Ned. Akad. Wetensch., afd. Wis- en Natuurk.*, 24, 1671–1676, Amsterdam.
- Versluys, J. 1931. Over de oorsprong van het gebruik van eenige vaktermen der hydrologie. *De Ingenieur* 46, 27–31.
- Vogel, J.C. 1967. Investigation of groundwater with radiocarbon. In *Isotopes in Hydrology*, STP 152, pp. 355–369. IAEA, Vienna.
- Volger, O. 1877. Die wissenschaftliche Lösung der Wasser-, im besondere der Quellenfrage mit Rücksicht auf die Versorgung der Städte. *Zeitschr. Ver. Deutscher Ingenieure* 21, 482–502.
- Voss, C.I. (Ed.). 2005. The future of hydrogeology. *Hydrogeol. J.* 13, 1–349.
- Vreedenburgh, C.G.J. 1935. Over de stationaire waterbeweging door grond met homogeen-anisotropische doorlaatbaarheid. *De Ingenieur in Ned. Indië* 2, 140–143.
- Vreedenburgh, C.G.J. and Stevens, O. 1936. Electric investigation of underground water flow nets. *Proc. Int. Conf. Soil Mech. and Found. Eng., Cambridge (Mass.)*, 1936, 219–222.
- Walton, W.C. 1970. *Groundwater Resources Evaluation*. McGraw-Hill, New York.
- Weber, H. 1928. *Die Reichweite von Grundwasser-absenkungen mittels Rohrbrunnen*. Ph.D. dissertation, Springer, Berlin.
- Wenzel, L.K. 1932. Recent investigations of Thiém's method for determining permeability of water bearing materials. *Am. Geophys. Union, Trans. 13th Ann. Meeting*, pp. 313–317.
- Wenzel, L.K. 1936. The Thiém method for determining permeability of water bearing materials and its application to the determination of specific yield — results of investigations in the Platte River valley, Nebraska. *U.S. Geol. Survey Water-Supply Paper* 679A, pp. 1–57.

- Wind, G.P. 1955. A field experiment concerning capillary rise of moisture in a heavy clay soil. *Neth. J. Agric. Sci.* 3, 60–69.
- Wright, E.P. and Edmunds, W.M. 1971. Hydrogeological studies in Central Cyrenaica, Libya. *Sympos. Geology of Libya, University of Tripoli*, pp. 459–482.
- Youngs, E.G. 1958. Redistribution of moisture in porous material after infiltration. *Soil Sci.* 86, 117–125; 202–207.
- Zimmermann, U., Münnich, K.O., and Roether, W. 1967. Downward movement of soil moisture traced by means of hydrogen isotopes. In *Isotope Techniques in the Hydrologic Cycle*, pp. 28–36. *Geophys. Monograph Ser.* 11, Am. Geophys. Union.
- Zötl, J. 1974. *Karsthydrologie*. Springer, Wien.
- Zunker, F. 1923. Die Bestimmung der spezifischen Oberfläche des Bodens. *Landw. Jahrbuch* 58, 159–203.

Further Information

Biswas (1970) gives an extensive account on the history of hydrology from ancient times to the end of the 19th century. A historical sketch on the evolution of theories about the hydrological cycle is presented by Brutsaert (2005). Recent trends in historical perspective on several aspects of hydrogeology, notably from North America, are reviewed in Narasimhan (1982); Narasimhan (1998) gives an explanatory overview of the developments in determination and characterization of hydraulic subsurface parameters. Recent trends are also found in a Special of the *Hydrogeology Journal* on the future of hydrogeology (Voss, 2005). See *References* for complete citations.

2

Geological Occurrence of Groundwater

2.1	Introduction.....	2-2
2.2	Importance of Groundwater	2-3
	Groundwater and World's Available Fresh Water •	
	Groundwater Availability in North America • Brief History	
	of Groundwater Usage	
2.3	Earth Materials.....	2-7
	Minerals and Rocks • Unconsolidated Materials	
2.4	Important Hydraulic Properties of Earth Materials and Groundwater, with Examples	2-10
	Porosity • Moisture Content • Capillarity • Electrostatic	
	Forces of Attraction and States of Water in Pores •	
	Compressibility of Water • Compressibility of Solid Earth	
	Materials • Hydraulic Head • Storage • Intrinsic	
	Permeability • Viscosity • Hydraulic Gradient • Hydraulic	
	Conductivity of Saturated Media • Hydraulic Conductivity	
	of Unsaturated Media • Anisotropy and Heterogeneity •	
	Aquifers	
2.5	Water in the Unsaturated Zone	2-28
	Moisture Content vs. Depth • Recharge and Infiltration	
	Capacity • Hydraulic Conductivity and Specific Discharge	
	through the Unsaturated Zone • Residence Time •	
	Subsurface Stormflow or Interflow Zone • Atmospheric	
	Discharge and Seepage Face • Discharge to the Saturated	
	Zone	
2.6	Water in the Saturated Zone	2-33
	Residence Time • Surface Discharge • Gaining and Losing	
	Streams • Bank Storage	
2.7	Protection of Groundwater Supplies	2-42
	Water Laws • Natural Protective Barriers and Waste	
	Disposal • Artificial Protective Barriers • Well-Head	
	Protection Programs	

Darrell I. Leap
Purdue University

2.8	Groundwater and Modern Environmental Concerns	2-46
	Possible Effects of Climate Change on Groundwater Flow •	
	Groundwater and Wetlands • Groundwater and Land	
	Subsidence • Groundwater and Soil Liquefaction • Excess	
	Evaporation of Groundwater and Consequent Salt Buildup •	
	Artesian Pressure and Surface Effects • Groundwater	
	Overdraft and Its Consequences • Effects of Mining •	
	Effects of Artificial Drainage • Chemical Changes due to	
	Groundwater–Surface Water Interactions • Anthropogenic	
	Changes in Groundwater Recharge Areas • Groundwater	
	Influence on Construction • Effects of Irrigation	
	Glossary.....	2-53
	References	2-56
	Further Information	2-59

2.1 Introduction

Subsurface water is generally divided into two major types: phreatic water or soil moisture in the unsaturated zone, and groundwater in the saturated zone. This division is made mainly because of the differences in the physics of flow of water in the saturated vs. the unsaturated zone; these differences are discussed later in this chapter.

To understand the increasing importance of groundwater and the science explaining it in a modern, industrialized nation and economy, it is well to consider the experience of the United States of America. In the 18th century, when the United States was being formed into a new nation, little emphasis was given to groundwater because in the humid East, where most of the population was located, surface water met nearly all water needs. In the 19th century and into the first half of the 20th, groundwater was investigated as a water source, especially as the population moved westward and new farms, ranches, mines, and cities were created in the semiarid West. Lack of water caused intense competition among the various uses and users, which continues to this day at an even greater pace.

Anticipating a need for natural resource evaluation of all kinds, the U.S. Government in the waning years of the last century sponsored a number of exploratory and investigative studies of water resources, especially in the West. These studies continued until well into this century. Notable were the field efforts of Chamberlin (1885), Darton (1909), Lee (1915), Mendenhall (1905), and others.

In addition, laboratory studies and theoretical studies added to our understanding of the physics and chemistry of groundwater. Important breakthroughs in physical understanding were made by several investigators who are listed with representative significant references: Hazen (1911), King (1899), Theis (1935), Jacob (1940), Hantush (1956), Hubbert (1940), and many others.

The work of Meinzer (1923, 1939) of the U.S. Geological Survey is significant in that he was the prime motivator for many government-sponsored investigations.

The work of these scientists was built upon a solid foundation of theory developed in Europe during the 19th and early 20th centuries by such luminaries as Darcy (1856), Boussinesq (1904), DuPuit (1863), Forcheimer (1914), and Theim (1906).

In the latter half of this century, several refinements of physical groundwater concepts have been made, and there has been a considerable body of knowledge built up as the result of efforts to more safely dispose of hazardous and radioactive wastes. Numerous advances in remote sensing by satellites have helped in delineating water-bearing entities such as fracture zones, springs, and to a lesser extent, aquifers. Geophysical techniques and methods have improved significantly in their use for groundwater exploration tools. Pumping tests and other hydraulic tests have been developed that greatly improve our understanding of the hydraulics of aquifers and their potential for groundwater supplies.

Perhaps the most significant tool developed in this century for the study of groundwater has been the digital computer. As a consequence, numerical modeling of groundwater flow and contaminant transport has become a commonplace effort in nearly all groundwater studies.

Problems of quality and quantity of data, in quantitative assessment of groundwater properties are mainly related to the functions of the heterogeneity of most geologic formations and their hydraulic properties, and the difficulty and expense of drilling for good data. With the recent advent of stochastic methods for analyzing hydrogeologic data, there are now better ways of interpolating, extrapolating, and predicting trends in hydraulic parameters.

With all the advances, both theoretical and technical, there is still one overriding fact in hydrogeologic analysis and modeling that must not be forgotten. Equations of flow and transport are differential in nature and require boundary conditions for their complete solution. The goodness of the solution to a partial differential equation of groundwater flow or transport is directly dependent upon the accuracy of the boundary conditions used. Boundary conditions are geological in nature and are only as accurate as the knowledge of the geology of the area being investigated.

This may appear simplistic to some, but the facts stand, not only from theory, but from the experience of thousands of investigators working on this problem for the past several decades. A thorough understanding of groundwater hydrology, correct interpretations of the results of hydraulic tests and resulting model predictions can only be made after a thorough understanding of the geology.

Still, high accuracy is often difficult to obtain in the quest for hydraulic parameter values of aquifers. There are variations in geologic structure that are often too small to be determined, but do affect hydraulic conductivity on a significant scale. Even in fairly homogeneous systems, this problem still frustrates the analyst and the modeler, and much more research is needed. Recent work by Eggleston et al. (1996) gives an example of this problem in glacial outwash of Cape Cod.

Perhaps the most crucial area of groundwater research today is the search for new and improved methods to acquire, quantify, and properly utilize good subsurface hydrogeological data with minimal expenditures of time and money.

2.2 Importance of Groundwater

2.2.1 Groundwater and World's Available Fresh Water

The amount of water on, under, and above the earth's surface remains essentially constant. Although a minor amount of water vapor may escape into space, additional new water is constantly created as *juvenile* water by chemical reactions during volcanic emanations. Table 2.1 shows estimates of the volumes of various kinds of water on, above, and under the surface of the earth as reported by Maidment (1993). Examination of the figures reveals that fresh water of all kinds comprises only 35,029,210 km³ in contrast to the 1,338,000,000 km³ sea water. Thus, fresh water makes up only 2.5% of all the water on earth, but not all of this water is available for human use. The water in polar ice caps, other forms of ice and snow, soil moisture, marshes, biological systems and the atmosphere are not readily available. As a result, only the 10,530,000 km³ of groundwater, 91,000 km³ of freshwater in lakes, and the 2,120 km³ of water in rivers are considered attainable for use and comprise a total of 10,623,120 km³. Consequently, groundwater comprises 99% of the earth's available fresh water.

2.2.2 Groundwater Availability in North America

Heath (1984) describes the groundwater regions of the United States. Heath (1988) has made a thorough compilation of groundwater areas in all of North America, including the United States, Canada, and Mexico. Figure 2.1 shows the major ground water regions of the conterminous United States.

The United States, because of its size and varied and complex geology, contains groundwater in many different kinds of rocks and unconsolidated materials. Table 2.2 summarizes the hydraulic characteristics of each of the regions in Figure 2.1. A major aquifer system in the United States is worth a brief discussion because of its importance in supplying water for major agricultural uses and because of problems associated with heavy pumping from it.

TABLE 2.1 Estimates of Relative Volumes of Water of Various Kinds on Earth

Item	Area 10^6 km^2	Volume km^3	Percent of total water	Percent of fresh water
Oceans	361.3	1,338,000,000	96.50	
Groundwater:				
Fresh	1,34.8	10,530,000	0.76	30.10
Saline	134.8	12,870,000	0.93	
Soil moisture	82.0	16,500	0.0012	0.05
Polar ice	16.0	24,023,500	1.7	68.6
Other ice and snow	0.3	340,600	0.025	1.0
Lakes:				
Fresh	1.2	91,000	0.007	0.26
Saline	0.8	85,400	0.006	
Marshes	2.7	11,470	0.0008	0.03
Rivers	148.8	2,120	0.0002	0.006
Biological water	510.0	1,120	0.0001	0.003
Atmospheric water	510.0	12,900	0.001	0.04
Total water	510.0	1,385,984,610	100	
Fresh water	148.8	35,029,210	2.5	100

Source: Adapted from Maidment, D.R. 1993. Hydrology, in Maidment, D.R., ed. *Handbook of Hydrology*, 1.1-1.15. McGraw-Hill, Inc., New York, NY. Data from UNESCO. 1992. *International Glossary of Hydrology*. 2nd Ed. WMO Rept. 385, World Meteorological Association. Geneva, Switzerland.

The High Plains Aquifer occupies the area marked as "High Plains" in Figure 2.1. It is composed of several hundred feet of Tertiary sediments that were laid down by streams flowing eastward from the present site of the Rocky Mountains as they were being uplifted. Sandy layers in the present aquifer provide irrigation water for several states in the High Plains.

Although there are a number of separate geological formations present, the Ogallala formation, consisting of alluvium, is the principal water-yielding unit and thus, is the one most often tapped by wells. Gutentag et al. (1984) reported that in 1978 there were 170,000 wells extracting water from this aquifer, and the pumping rate was from 2 to 100 times the recharge rate.

As a result, the water table has dropped 100 ft or more causing an increase in pumping expense. Careful management practices are now being put into place to better manage the resource. This is a significant example of ground-water mining and shows the effects of overuse of water resources.

Detailed descriptions of the hydrogeology of each region and in all of North America can be found in Heath (1988). In addition, the U.S. Geological Survey has performed detailed studies on several major aquifer systems in the United States through a major investigative effort known as the Regional Aquifer Systems Analysis (RASA). Specific reports of these studies are available from the U.S. Geological Survey National Headquarters in Reston, Virginia. Summaries of results of the RASA studies can be found in Bennet (1979), Sun (1986), and Chapters 25, 26, and 27 of Back et al. (1988).

2.2.3 Brief History of Groundwater Usage

The following paragraphs give only a brief synopsis of the history of groundwater usage, but a more complete history can be found in Chapter 1 of, "*History of Scientific Developments in Groundwater Hydrology*" by Jacobus J. de Vries.

It is uncertain when mankind first started extracting groundwater by artificial means such as wells or infiltration galleries. Early humans most likely drank from surface streams. They may also have discovered groundwater through the discharge of natural springs and in some parts of the world, used this source in addition to surface streams. As streams dried up in hot weather, people learned to dig into the alluvium to find water below the surface.

There was much speculation about the origin of water emanating from springs, and early theories proposed a hollow earth filled with water. The ancient Greeks were the first people to record their theories,

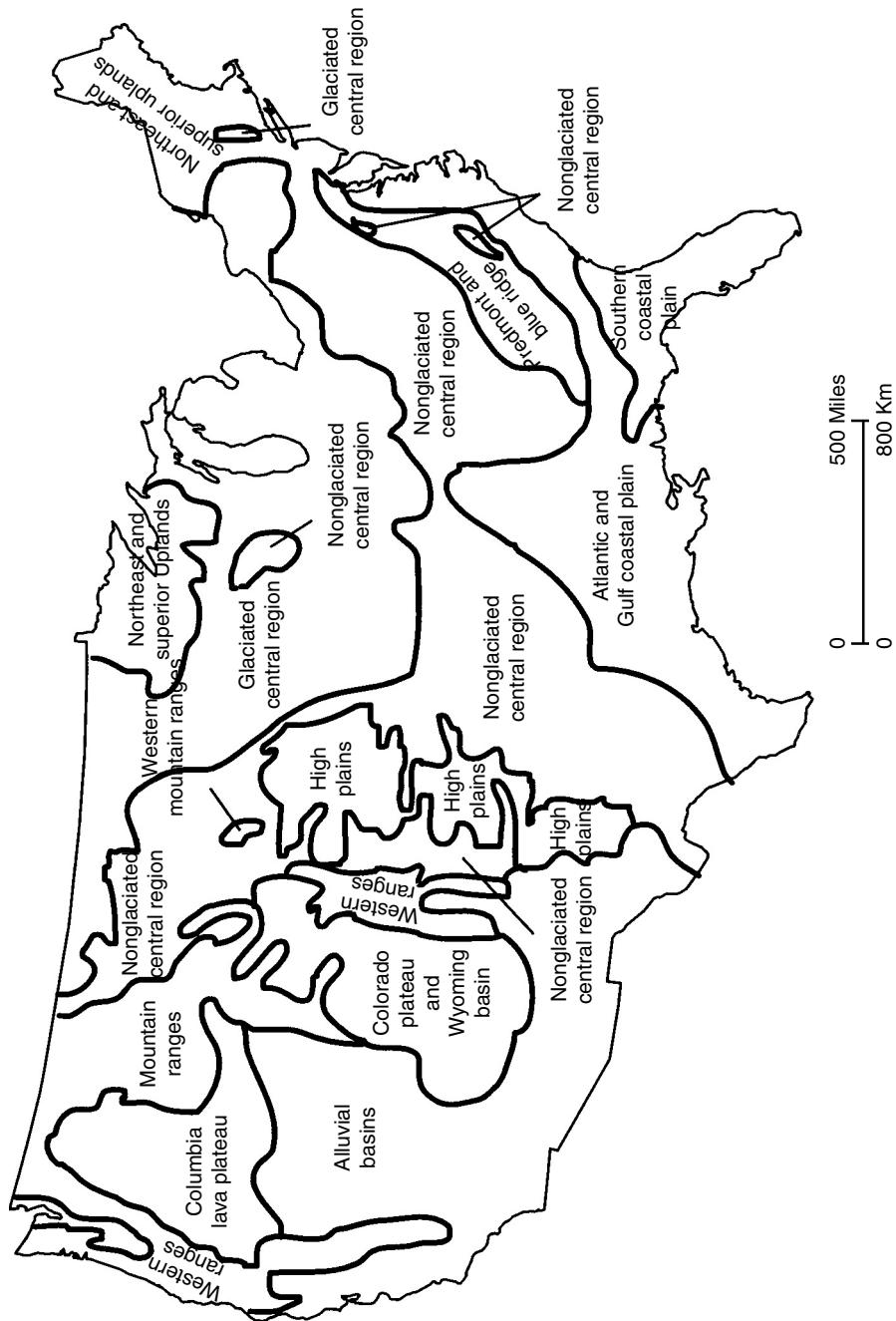


FIGURE 2.1 Ground Water regions in North America. (Adapted from Heath, R.C. 1984. Ground-Water Regions of the United States. U.S. Geological Survey Water-Supply Paper 2242, 1-78. U.S. Government printing Office, Washington, D.C.)

TABLE 2.2 Hydraulic Characteristics of Groundwater Regions in the United States

Region No.	Name	Components of the system				Characteristics of the dominant aquifers							
		Unconfined aquifer	Confining beds	Confined aquifers	Presence and arrangement	Water-bearing openings		Composition		Storage and transmission properties		Recharge and discharge conditions	
						Primary	Secondary	Degree of solubility	Porosity	Transmissivity	Recharge	Discharge	
1	Western Mountain Ranges	×	×	Hydrologically insignificant	Dominant aquifer	Minor aquifer or not very productive	Hydrologically insignificant	Not highly productive	Multiple productive aquifers	The dominant productive aquifer	Single unconfined aquifer	Two interconnected aquifers	Complex interbedded sequence bed, confined aquifer, containing bed, cooling cracks in lava
2	Alluvial Basins	×	×	Hydrologically insignificant	Hydrologically insignificant	Hydrologically insignificant	Hydrologically insignificant	Highly productive	Multiple productive aquifers	Not highly productive	Single unconfined aquifer	Two interconnected aquifers	Mixed soluble and insoluble
3	Columbia Lava Plateau	×	×	Hydrologically insignificant	Hydrologically insignificant	Hydrologically insignificant	Hydrologically insignificant	Highly productive	Multiple productive aquifers	Not highly productive	Two interconnected aquifers	Two interconnected aquifers	Large (>0.2)
4	Colorado Plateau and Wyoming Basin	×	×	Hydrologically insignificant	Hydrologically insignificant	Hydrologically insignificant	Hydrologically insignificant	Highly productive	Multiple productive aquifers	Not highly productive	Two interconnected aquifers	Two interconnected aquifers	Moderate (0.01-0.2)
5	High Plains	×	×	Hydrologically insignificant	Hydrologically insignificant	Hydrologically insignificant	Hydrologically insignificant	Highly productive	Multiple productive aquifers	Highly productive	Single unconfined aquifer	Single unconfined aquifer	Large (<0.01)
6	Nonglaciated Central Region	×	×	Hydrologically insignificant	Hydrologically insignificant	Hydrologically insignificant	Hydrologically insignificant	Highly productive	Multiple productive aquifers	Highly productive	Two interconnected aquifers	Two interconnected aquifers	Small (<0.01)
7	Glaciated Central Region	×	×	Hydrologically insignificant	Hydrologically insignificant	Hydrologically insignificant	Hydrologically insignificant	Highly productive	Multiple productive aquifers	Highly productive	Two interconnected aquifers	Two interconnected aquifers	Large (<0.01)
8	Piedmont and Blue Ridge	×	×	Hydrologically insignificant	Hydrologically insignificant	Hydrologically insignificant	Hydrologically insignificant	Highly productive	Multiple productive aquifers	Highly productive	Two interconnected aquifers	Two interconnected aquifers	Large (<0.01)
9	Northeast and Superior Uplands	×	×	Hydrologically insignificant	Hydrologically insignificant	Hydrologically insignificant	Hydrologically insignificant	Highly productive	Multiple productive aquifers	Highly productive	Two interconnected aquifers	Two interconnected aquifers	Large (<0.01)
10	Atlantic and Gulf Coastal Plain	×	×	Hydrologically insignificant	Hydrologically insignificant	Hydrologically insignificant	Hydrologically insignificant	Highly productive	Multiple productive aquifers	Highly productive	Two interconnected aquifers	Two interconnected aquifers	Large (<0.01)
11	Southeast Coastal Plain	×	×	Hydrologically insignificant	Hydrologically insignificant	Hydrologically insignificant	Hydrologically insignificant	Highly productive	Multiple productive aquifers	Highly productive	Two interconnected aquifers	Two interconnected aquifers	Large (<0.01)
12	Alluvial Valleys	×	×	Hydrologically insignificant	Hydrologically insignificant	Hydrologically insignificant	Hydrologically insignificant	Highly productive	Multiple productive aquifers	Highly productive	Two interconnected aquifers	Two interconnected aquifers	Large (<0.01)
13	Hawaii	×	×	Hydrologically insignificant	Hydrologically insignificant	Hydrologically insignificant	Hydrologically insignificant	Highly productive	Multiple productive aquifers	Highly productive	Two interconnected aquifers	Two interconnected aquifers	Large (<0.01)
14	Alaska	×	×	Hydrologically insignificant	Hydrologically insignificant	Hydrologically insignificant	Hydrologically insignificant	Highly productive	Multiple productive aquifers	Highly productive	Two interconnected aquifers	Two interconnected aquifers	Large (<0.01)
15	Puerto Rico and Virgin Islands	×	×	Hydrologically insignificant	Hydrologically insignificant	Hydrologically insignificant	Hydrologically insignificant	Highly productive	Multiple productive aquifers	Highly productive	Two interconnected aquifers	Two interconnected aquifers	Large (<0.01)

Source: Adapted from Heath, R.C. 1984. *Ground Water Regions of the United States*, U.S. Geological Survey Water-Supply Paper 2242, 1-78, U.S. Government Printing Office, Washington, DC.

germinated no doubt from their observation of disappearing rivers and emerging springs in the limestone *karst* areas of the Greece and the rest of the Balkan Peninsula. Davis and DeWiest (1966) summarize the history of groundwater use and the development of the body of knowledge now surrounding it.

The earliest recorded use of groundwater is found in the 16th chapter of the Book of Genesis where it tells of a spring in the desert. In Chapter 26, mention is made of the digging of wells.

Egyptians are reported to have perfected core drilling by 3,000 b.c. Ancient Chinese invented the churn drill (similar to the modern cable-tool drill) and drilled to a depth of 1500 m (Tolman, 1937).

Western Europe did not engage in drilling until the 12th century when churn drilling was developed there. Much of the impetus for development of drilling methods was generated in the area around Artois, France, where artesian water was in demand.

In the latter part of the 19th century, the rotary drilling technique was developed, and by 1890 the *hydraulic rotary method* had been perfected. This was probably the single most significant advance in drilling technology (Davis and DeWiest, 1966).

As the water-well and oil-well drilling businesses grew in various parts of the world, rotary drilling methods were continuously improved and introduced worldwide. Today, it is the most widely used method and has been the primary reason that groundwater is now available to even primitive tribes. Present groundwater availability to people on every continent has been a prime reason for increasing the health and standard of living for untold millions.

Groundwater is now recognized as a national treasure and a most important natural resource. For example, approximately half the water used in the United States of America is groundwater. In some areas, groundwater is the sole water supply for all purposes.

2.3 Earth Materials

2.3.1 Minerals and Rocks

The inorganic part of the earth is made up of rocks and their weathering products. A rock, by definition, contains one or more *minerals*, and a mineral is defined by Hurlburt (1970) as follows:

"In addition to being natural and inorganic, a mineral must meet another requirement: it must be a chemical element or compound. It cannot be a random mixture of elements; the atoms that make it up must have definite ratios to each other, so that its composition can be expressed by a chemical formula. Not only are the proportions of the various atoms of a given mineral fixed, but so are their relative positions. These attributes give to each mineral a set of properties that characterize it so uniquely that one can distinguish it from all other minerals."

A *rock* can be composed of one mineral or a mixture of several. For example, sandstone may contain grains of the mineral, quartz (silica or SiO_2), and a cement between the quartz grains composed of the mineral, calcite (CaCO_3). Granite is normally composed of crystals of the minerals, feldspar, quartz, mica, and others. By definition then, frozen water (ice) is both a mineral and a mono-mineralic rock.

The most fundamental of the three classes of rocks are *igneous rocks* that are formed as cooling products from the molten state. Igneous rocks are primordial in many parts of the world and are some of the world's oldest, exceeding three billion years in age. One way they can be formed is when molten rock is intruded into other rock formations and then cooled to the solid state. If the cooling is slow enough, various minerals will crystallize out into an interlocking solid mass that is characteristic of particular rocks such as granite.

Extrusive igneous rocks like dense basalt are forced from fissures in the earth's crust and harden into vast sheets of solid material, usually containing very small crystals (due to the rapid cooling), or perhaps no crystals at all (obsidian glass). Other extrusives like lighter lava or pumice are ejected during volcanic eruptions and are highly charged with gases to form very porous and even frothy glasses resembling a

sponge. Some of the ejecta may fall from the air to settle as a sediment. This particular kind of deposit is known as a *pyroclastic* rock, that is, part igneous and part sedimentary.

A *sedimentary rock*, the second class, is deposited from either air or water as grains of rocks and minerals. These sediments in turn, may have been derived from the weathering of igneous, metamorphic, or other sedimentary rocks.

If any kind of rock, igneous, metamorphic, or sedimentary, is subjected to intense heat and pressure, such as exists at great depths in the earth's crust, at the edge of tectonic plates, or in rising mountain ranges, the parent rock will be transformed into the third class of rock — a *metamorphic rock*. A metamorphic rock may contain the same chemical composition as the parent rock, but the mineral composition and structure may be changed drastically from the parent. For example, limestone containing amorphous or cryptocrystalline calcite (CaCO_3) is often metamorphosed into marble that has a definite crystal structure and is much harder than the original limestone. Granite may be metamorphosed into gneiss with the same overall chemical composition as the original granite, but with new minerals and mineral structures.

Groundwater can be found in all three classes of rocks, but in general, the sedimentary rocks contain by far the greatest amounts of water due to their greater *porosity*.

2.3.2 Unconsolidated Materials

Unconsolidated materials are those earth materials that have not been indurated. That is, the grains have not been fused together by heat and pressure, as in the cases of igneous granite or metamorphic gneiss; or by cement, as in the case of sedimentary rocks. Unconsolidated materials can be the non-indurated products of weathering of all three classes of rocks, or sediments laid down by running water, ponded water, the sea, or ejecta from volcanoes.

Most unconsolidated materials are young, geologically speaking, and are at or near the earth's surface. Thus, they have not been exposed to pressure, heat, and migrating cementing fluids long enough to become consolidated or hard. Hence, they generally contain high porosity and are the sources for much groundwater.

A common unconsolidated deposit in glaciated areas of the world is *glacial drift*. It is any kind of earth material that was deposited directly by glaciers or by meltwater from glaciers. As such, it can range in size from the finest silt or clay to the largest boulders, and can be mixtures of all sizes. The name "drift" was given to this material when it was believed that it was depositional material, or "drift," from the great Noachian Flood, described in the Book of Genesis. Drift can generally be subdivided into the more specific lithologies such as *till* and *outwash*.

Glacial till is a generally a heterogeneous mixture of many different lithologies and particle sizes. Typically, in the midwestern United States, till contains a preponderance of clay and silt with additional amounts of ground-up rocks and boulders that may vary in size from small pebbles to erratics the size of a house or larger. On rare occasions, geologists find tills composed of one lithology indicating local sources of material. Glacial till is not generally utilized as a source of groundwater because of its low permeability.

Glacial outwash is material deposited from high-energy streams of water that originated from melting glaciers. This process can be seen today at the toe of any mountain glacier on different continents.

Glacial ice, more often than not, contains entrained rock and soil material which it has eroded from the surrounding valley sides or ground beneath it. This material varies widely in size from clay to large boulders. As the ice melts and leaves the toe of the glacier, normally in great flow rates and very turbulent, this material is moved with the flowing water.

As the stream loses energy, materials settle out with larger sizes coming out first, followed by gradually lessening sizes. Therefore, along a stretch of an outwash stream one may find coarse gravel and boulders settling out first, followed by finer gravel, then sand and gravel mixtures, then sand, silt, and finally clay (in still water).

Many present-day stream courses were glacial spillways for outwash water and sediment during the Pleistocene Epoch (Ice Age). For example, the Wabash River Valley in Indiana contains outwash sand and gravel deposits in excess of 300 ft in thickness, which were deposited by the melting of two and possibly

three different ice sheets. The Big Sioux River valley in South Dakota and Iowa is another example. A very extensive deposit of outwash sand is the Cape Cod peninsula, which was deposited in an interlobate outwash between two lobes of ice – one to the east and one to the west of the site.

Outwash sand and gravel deposits are frequently exploited for groundwater because of their high porosity and permeability. Well yields in many of these deposits often exceed 5500 m³/d.

Other water-lain deposits not directly deposited from glacial meltwater may be comprised of reworked glacial detritus (glacially-transported material), or they may be found in areas where glaciers never occurred. The most common example of such deposits is *alluvium*. Flood plains along large streams are created of this material as the streams flood over their banks and deposit the material. Stream beds also contain alluvium.

In fast-flowing streams with high gradients, as in mountainous areas, alluvium may be absent because the stream is eroding rather than depositing material. If alluvium is found in and along such streams, it is generally very coarse-grained gravel with large boulders. On the other hand, mature streams such as the Ohio, Missouri, and Mississippi Rivers deposit their loads of fine silt and clay over broad flood plains.

Alluvium deposits may serve as important groundwater sources, but in large river valleys, the yield of such deposits may be low to moderate, depending on the grain size and the resulting permeability and porosity.

Lacustrine materials are silts and clays that are deposited from relatively still bodies of water such as lakes and lagoons. These materials, being so fine-grained, are not utilized for groundwater supplies because of their low permeability.

Peat is the remains of mostly water plants that die and accumulate in ponded water and marshes over long periods of time. The top part of a peat deposit is very porous and permeable, but it becomes more compact with depth. The lower layers of peat are often sticky masses of black organic material with little resemblance to the original plant material.

Peat is not generally utilized as a groundwater source, but it can serve, under the right conditions, as a natural cleansing agent to remove organics and heavy metals from water that passes through it — the large and complex organic molecules in the peat attract such contaminants.

Chemical precipitates of most importance include limestone and marl that precipitate directly from sea water or even from fresh water bodies. Major deposits of limestone were deposited in many parts of the world during the Cretaceous Period (the “Chalk Period”) of the Mesozoic Era. Examples include the Chalk Beds of Dover, England; the limestones of the Balkan Peninsula; and limestones of the High Plains in the United States. Other vast limestone deposits were formed in earlier times and are found across most of the Midwestern United States and the Appalachians.

Limestone is composed primarily of calcite. Entrained silt and clay and other materials may also be present. After deposition, a process known as *diagenesis* often takes place in which the rock incorporates magnesium to become dolostone, or CaMg(CO₃)₂; this is also the formula for the mineral, dolomite. Pure MgCO₃ is the mineral, magnesite.

Collectively, limestone and dolostone are called “carbonate rocks”, or “carbonates.” Carbonates, especially limestone, often undergo solution along bedding planes and fractures to form caves and sinkholes, which in an interconnected system is known as *karst* terrane, after the Karst region in Yugoslavia. Networks of such caves and tunnels may exceed hundreds of miles in length and may contain large streams that emanate from springs in the rock.

Examples of such systems are found in the Balkan Peninsula, the Mediterranean area, France, Kentucky (Mammoth Caves), New Mexico (Carlsbad Caverns), to name just a few. Karst systems are often sources of very large quantities of groundwater, and due to the very high permeability, can be productive aquifers. A drawback, though, is the ease with which water in karst systems can be contaminated by surface sources. Thus, care must be taken to protect such sources.

Aeolian deposits are fine-grained materials, such as silt and sand, which may have been deposited originally from water, but which have been reworked and redeposited by wind. Examples of such active deposits today can be found in sand-dune areas of the Sahara, the Middle East, New Mexico, Nevada, and many other places.

Ancient dunes from the geological past are often found as sandstone bodies and may have some potential for groundwater extraction if coarse enough to allow sufficient porosity and permeability. Generally, wind-blown deposits are fine-grained, and when cemented with precipitates from circulating groundwater, may possess low porosity and permeability. The finest grained aeolian deposits are composed of silt or "rock flour" known as *loess*. The silt source is usually a wide river bed with braided channels where large dry areas of fine-grained materials are exposed to the wind. In most cases, the silt was deposited in such rivers as the end product in the long chain of deposition of glacial outwash. Prevailing winds then pick up the silt and transport it downwind where it is deposited on the lee sides of river valleys.

Significant loess deposits are found on the east side of the Missouri River, which acted as a glacial spillway during the Pleistocene. North of Sioux City, Iowa, this material forms bluffs that are ten meters high. Other noteworthy deposits are found along the Mississippi River (another spillway) and in the Gobi Desert of China.

Loess, after deposition, will often be reworked by frost action to form columnar structures with vertical fractures. The grains of silt are then oriented with their long axes vertically to form such features. With this alteration, it will allow fast vertical movement and drainage of water, and is fairly solid material to build upon, but its permeability is too low to utilize it for groundwater supply.

2.4 Important Hydraulic Properties of Earth Materials and Groundwater, with Examples

2.4.1 Porosity

Porosity (n) is defined as the volume of the pores of a rock or soil sample (V_p) divided by the total volume (V_t) of both pores and solid material, that is,

$$n = V_p/V_t \quad (2.1)$$

When a rock is first formed by precipitation, cooling from an igneous melt induration from loose sediments; or when a soil is first formed by weathering of rock materials and possibly subsequent biological action, the new entity will contain a certain inherent porosity known as *primary porosity*. This porosity may later be reduced by cementation from precipitates from circulating groundwater, or from compaction accompanying burial by later sediments. However, fractures or solution cavities formed in the rock, or root tubes or animal burrows in soils may later form and are known as *secondary porosity*. Thus, the total porosity of a sample will be the sum of the primary and secondary porosities.

Porosity of a consolidated sample can be determined quite simply by first cutting the sample to a known dimension such as a cylinder or cube and measuring the total volume. Next, the sample is submerged in a known volume of water and allowed to saturate. After saturation, the volume of water displaced will be the volume of solids in the sample. The volume of voids is simply the difference between the total volume and the volume of solids, and porosity can be calculated by the above formula. More accurate and more sophisticated methods for testing earth materials for various properties are given in various publications of the American Society for Testing and Materials (ASTM).

If all the pores in a rock are not connected, only a certain fraction of the pores will allow the passage of water, and this fraction is known as the *effective porosity*. An example is pumice, a glassy volcanic ejecta (a solidified froth) which may float on water because the total porosity is so high and it contains so much entrained gas that its bulk density is less than water. It may take some days to sink because the effective porosity is so low that water cannot easily pass through it.

Coarse gravel may contain a porosity of only 25%, but that porosity will practically be effective. Thus, gravel is an excellent conductor of water.

Porosity of a rock or soil is determined largely by the packing arrangement of its grains and the uniformity of the grain-size distribution. The greatest ideal porosity that could be attained in a material

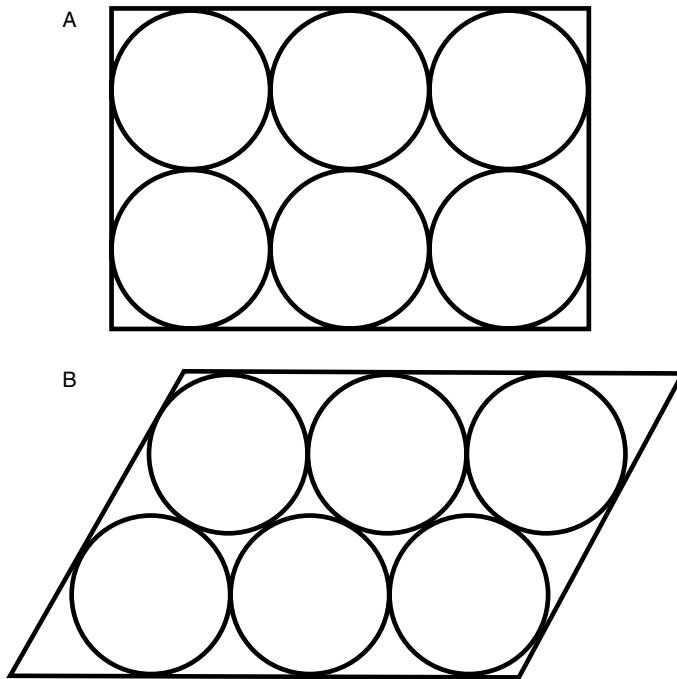


FIGURE 2.2 (A) Cubic packing; (B) Rhombohedral packing.

with uniform spherical grains is 47.65%; this is known as *cubic packing* (Figure 2.2A) because the centers of eight such grains form the vertices of a cube. The least porosity that can be attained with the same grains is found in *rhombohedral packing* with a value of 25.95% (Figure 2.2B). The centers of the eight adjacent spheres form the vertices of a rhombus.

These are ideal extremes because no natural sediment or rock contains spherical grains nor are the grains of uniform size. If small grains are situated in the spaces between large grains, the porosity will be reduced below that for the large grains only. This is demonstrated by the *uniformity coefficient* (C_u) which is a measure of how well or how poorly the sample grains are sorted:

$$C_u = D_{60}/D_{10} \quad (2.2)$$

where D_{60} is the grain diameter below which 60% of the grains are finer, and D_{10} (the effective grain size) is the diameter below which 10% of the grains are finer. If the C_u is less than four, the sample is well sorted; if it is greater than six, it is poorly sorted and will have a lower porosity than the first sample.

The *void ratio*, e , used in soil mechanics, is defined as

$$e = V_p/V_s \quad (2.3)$$

where V_s is the volume of solids, so that

$$n = 1/(1 + 1/e) \quad (2.4)$$

Typical values of porosity are given in Table 2.3.

TABLE 2.3 Important Physical Properties of Soil and Rock

Lithology	Porosity (percent)	Hydraulic conductivity (cm/sec)	Compressibility, $\alpha(m^2/N \text{ or } Pa^{-1})$
Unconsolidated			
Gravel	25–40	10^{-2} – 10^2	10^{-8} – 10^{-10}
Sand	25–50	10^{-4} –1	10^{-7} – 10^{-9}
Silt	35–50	10^{-7} – 10^{-3}	no data
Clay	40–70	10^{-10} – 10^{-7}	10^{-6} – 10^{-8}
Glacial Till	10–20	10^{-10} – 10^{-4}	10^{-6} – 10^{-8}
Indurated			
Fractured Basalt	5–50	10^{-5} –1	10^{-8} – 10^{-9}
Karst Limestone	5–50	10^{-4} –10	not applicable
Sandstone	5–30	10^{-8} – 10^{-4}	10^{-11} – 10^{-10}
Limestone, Dolomite	0–20	10^{-7} – 10^{-4}	$< 10^{-10}$
Shale	0–10	10^{-11} – 10^{-7}	10^{-7} – 10^{-8}
Fractured Crystalline Rock	0–10	10^{-7} – 10^{-2}	$—10^{-10}—$
Dense Crystalline Rock	0–5	10^{-12} – 10^{-8}	10^{-9} – 10^{-11}

Source: Adapted from Domenico, P.A. and Schwartz, F.W. 1990. *Physical and Chemical Hydrogeology*, John Wiley and Sons, Inc., New York; Freeze, R.A. and Cherry, J.A., 1979. *Groundwater*, Prentice-Hall, Inc., Englewood Cliffs, NJ; Fetter, C.W., 1994. *Applied Hydrogeology*, 3rd ed. Macmillan College Publishing Co. Inc., New York; Narashimhan, T.N., and Goyal, K.P., 1984. Subsidence due to geothermal fluid withdrawal, in Man-Induced Land Subsidence, *Reviews in Engineering Geology*, v. VI, 35–66, Geological Society of America, Boulder, Co.

2.4.2 Moisture Content

Moisture content (ϕ) can be measured and described either *gravimetrically* or *volumetrically*. The gravimetric equation is

$$\phi = W_w/W_t \quad (2.5)$$

where W_w is the weight of the water contained in a sample and W_t is the total weight of the solids and water in the sample. This definition, although useful for some purposes, does not indicate the degree of saturation of the rock or soil. Thus, the volumetric definition is more widely used:

$$\phi = V_w/V_t, \quad (2.6)$$

where V_w = volume of water in the sample, and V_t = total volume of water and solids. Closely related terms are the *saturation ratio* which is the volume of contained water divided by the volume of voids; and the *degree of saturation*, which is the saturation ratio times 100 (V_w/V_v) \times 100. If the saturation ratio is less than unity, or the degree of saturation is less than 100%, the sample is unsaturated and the pores are then partially filled with air.

Gravimetric moisture content can be determined by weighing a sample to obtain the total weight and then drying it in an oven to drive out the moisture. The dry sample is weighed to obtain the weight of the solids. The difference is the weight of the entrained water.

2.4.3 Capillarity

Capillary forces play a major role in the movement of water through unsaturated materials. Water is attracted to solid grains by *adhesion*. The familiar example of water rising in a soda straw is also a good example for capillary rise, h_c , in a small tube of radius, r , (Figure 2.3) and the rise is calculated as

$$h_c = 2\sigma \cos \alpha / \gamma_w r \quad (2.7)$$

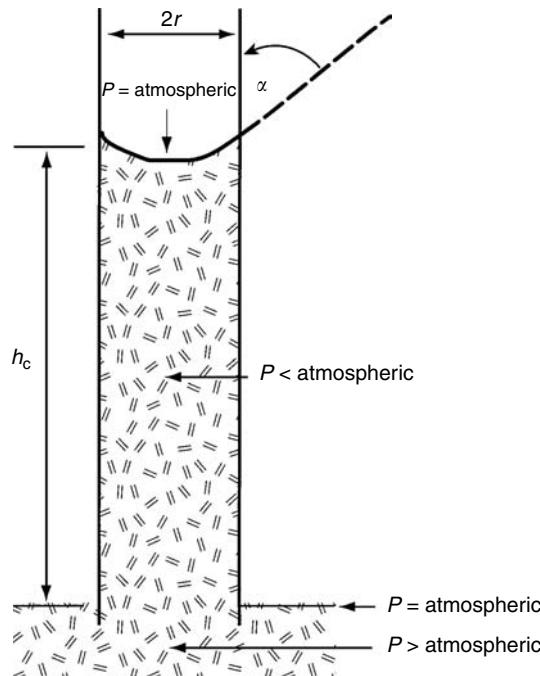


FIGURE 2.3 Capillary rise in a tube. (Adapted from Lohman, S.W. 1972. *Ground-Water Hydraulics*. U.S. Geological Survey Prof. Paper 708, U.S. Government Printing Office, Washington, DC.)

where σ is the surface tension of water (0.0756 N/m at 0°C), γ_w is the specific weight of water (9.805 kN/m³ at 0°C), $\alpha = 0^\circ$ for water. Thus, for water at 0°C in a clean tube, Domenico and Schwartz (1990) give the formula for capillary rise in cm as

$$h_c = 0.153/r. \quad (2.8)$$

Mavis and Tsui (1939) developed the following equation to estimate the height of capillary rise in millimeters in soils,

$$h_c = (2.2/d_H)[(1 - n)/n]^{2/3} \quad (2.9)$$

where d_H is the harmonic mean grain diameter in millimeters and n is the porosity. Capillary rise in coarse gravels may be only a few millimeters, but in clay it may be as much as three or four meters (Davis and DeWiest, 1966; Lohman, 1972).

In the root zone, normally in the upper part of an unsaturated soil, there is a natural competition for water between capillary forces in the soil and osmotic suction in the plant roots. The moisture content at which the capillary forces become greater than osmotic forces is known as the *wilting point*, that is, the moisture content at which the plants will cease to take in water and will start to wilt.

Due to the fact that capillary forces in coarse-grained materials are not as great as those in fine-grained materials, the wilting point in sand is only about 5% whereas in clay it is around 25 to 30% (USDA, 1955).

The wilting point will be reached more rapidly during hot, dry weather, especially if rainfall is insufficient, because capillarity causes water in the soil to rise where evapotranspiration from the surface will accelerate the loss of water from the soil.

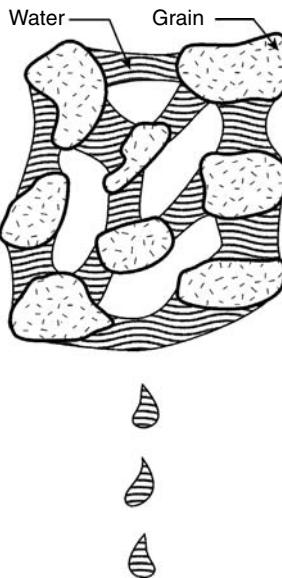


FIGURE 2.4 Porous medium under drainage.

2.4.4 Electrostatic Forces of Attraction and States of Water in Pores

Capillarity is caused by a combination of two forces: (1) *molecular attraction*, which is responsible for water adhering to soil or rock particle surfaces, and (2) *surface tension*, which is due to cohesion of water molecules toward each other when water is exposed to air. In saturated systems, the first force is balanced in all directions and thus, canceled out, and surface tension is also balanced.

An unsaturated rock or soil, if allowed to drain under the force of gravity, will not lose all its water by this means (Figure 2.4). Instead, some water will be held inside the pores by these forces. The drained water is known as *gravitational water* and the retained water as *capillary water* (Todd, 1980). Earlier literature may refer to the latter as *pellicular water*. The capillary or pellicular water will stay in the pores unless it is subjected to additional stresses such as centrifugation in the laboratory, or excessive heat as during a hot dry period.

If moisture reduction continues, the minimal water that will be held will exist as thin films around grains and will be held there by adhesive forces of molecular attraction; this water is known as *hygroscopic water* and is unavailable to plants. The wilting point is always a higher moisture content than the hygroscopic water content.

Often, initially dry soils near the surface will first attract and hold hygroscopic water if the air humidity increases significantly. It is only after the hygroscopic moisture content exceeds the *hygroscopic coefficient* (the maximum water content that can be held hygroscopically) that any additional water from precipitation or irrigation travel through the soil. Adhesion and cohesion are strongly affected by the mineralogy of soil or rock and chemical content of soil and water.

2.4.5 Compressibility of Water

Water is only very slightly compressible. At conditions of constant temperature and mass, Domenico and Schwartz (1990) define the *isothermal compressibility* of water as

$$\beta_w = 1/K_w = -(1/V_w)(\partial V_w / \partial P) \quad (2.10)$$

where β_w is the fluid compressibility in units of reciprocal of pressure, K_w is the bulk modulus of compression for water, V_w is the bulk volume of water, and P is the pressure. At 25°C, groundwater possesses a β_w of $4.8 \times 10^{-10} \text{ m}^2/\text{N}$ or $2.3 \times 10^{-8} \text{ ft}^2/\text{lb}$.

2.4.6 Compressibility of Solid Earth Materials

The compressibility of water-bearing rock and soil at some internal point is affected by both external and internal stresses and pressure of entrained water within the pores. The stress-balance equation is given as

$$\sigma_t = \sigma_e + P_p \quad (2.11)$$

where σ_t is the total vertical stress acting downward on the point of interest and includes the pressure of overlying soil or rock and its contained water as well as that from buildings, trees, and the like on the surface. The effective stress or resisting stress from the skeleton of the solid grains, that is, the matrix, is σ_e , and P_p is the pore pressure of the water in the pores. Figure 2.5 illustrates this relationship.

Any increase in total vertical stress must be balanced by the same increase on the right side of the equation; this occurs when depth increases because total vertical stress naturally increases with depth. If a well is pumped, the extraction of water will suddenly decrease the pore pressure and the porosity with an attendant increase in the effective stress exerted by the matrix. As a result, the matrix will compress. If water is injected into the well, the reverse occurs.

Such expansion and compression of the matrix, at constant temperature and assuming incompressible grains, can be quantified as shown by Domenico and Schwartz (1990) as

$$\beta_b = 1/K_b \quad (2.12)$$

$$= -(1/V_b)(\partial V_b / \partial \sigma_t) \quad (2.13)$$

$$= -1/V_b(\partial V_p / \partial \sigma_t) \quad (2.14)$$

$$= \beta_p = 1/H_p \quad (2.15)$$

where β_b is the bulk matrix compressibility in units of the reciprocal of pressure, K_b is the rock bulk modulus of compression, β_p is the vertical compressibility, H_p is a modulus of vertical compression related only to pores, V_b is the bulk volume, and V_p is the pore volume. The negative sign in the equation refers to the fact that volume decreases with increasing pressure. Equation 2.9–2.12 show that if the grains are incompressible then β_b is equivalent to β_p . Table 2.3 lists the compressibility of several lithologies.

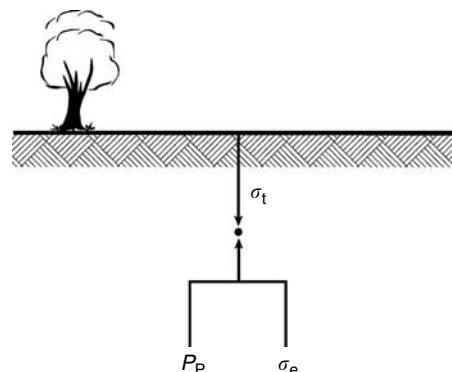


FIGURE 2.5 Stress-balance diagram.

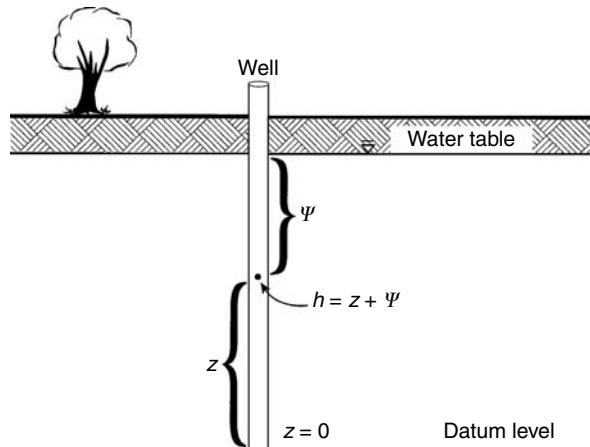


FIGURE 2.6 Illustration of hydraulic head.

2.4.7 Hydraulic Head

Hydraulic head at a point in a groundwater system (Figure 2.6) is expressed as

$$h = Z + P/\rho g \quad (2.16)$$

where Z is the elevation head or the distance of the reference point above a datum plane (normally mean sea level), P is fluid pressure at the point exerted by the column of water above the point, and ρg is the specific weight of water, γ , or more simply stated

$$h = Z + \Psi \quad (2.17)$$

where Ψ is also called the pressure head. The relationship has the dimension of length.

Development of this relationship can be found in Chapter 3 by J.W. Delleur and in Hubbert (1940).

Thus, hydraulic head has the dimension of length that makes it convenient for calculations based on the elevation of water above sea level, the generally accepted zero datum. Fluid potential and hydraulic head are equivalent and both are used, except that the head is the most popular entity for groundwater studies. It can be shown, often to the surprise of the unsuspecting, that water can flow from a region of lower pressure to a region of higher pressure if the total head at the starting point is greater than at the ending point.

In the field of petroleum and natural gas engineering, pressure is generally used in place of the head because pressures at great depths are normally so great that elevation heads are often insignificant.

2.4.8 Storage

A simple mass balance equation for groundwater flow through a unit volume of porous medium is given as

$$Q_{\text{out}} \Delta t = Q_{\text{in}} \Delta t \pm \Delta S \Delta t \quad (2.18)$$

where Q is the total flow rate in volume per unit time, and ΔS is the volume per unit of time going into or coming out of storage in time, t .

In a saturated porous medium that is confined between two transmissive layers of rock or clay, water will be stored in the pores of the medium by a combination of two phenomena: these are water compression and aquifer expansion. As water is forced into the system at a rate greater than that it is being extracted, the

water will compress and the matrix will expand to accommodate the excess. In a unit volume of saturated porous matrix, the volume of water that will be taken into storage under a unit increase in head, or the volume that will be released under a unit decrease in head is called *specific storage*, and is shown as

$$S_s = \rho g(\alpha + n\beta) \quad (2.19)$$

where α is aquifer compressibility, ρ is fluid density, g is gravitational acceleration, n is porosity, and β is water compressibility. This unit has the dimension of $1/L$ and is quite small, usually 0.0001 or less.

The storage coefficient of an aquifer, or simply, the storativity, S , is given as

$$S = S_s b \quad (2.20)$$

where b is the saturated thickness of the aquifer. Storativity is defined as the volume of water per unit aquifer surface area taken into or released from storage per unit increase or decrease in head respectively (Figure 2.7). It is a dimensionless quantity. In confined aquifers the value of storativity ranges from 0.005 to 0.00005.

In unconfined porous media, that is, where there is no overlying confining cover, storage of water in its upper part is defined as specific yield, S_y . This is the ratio of the volume of water that drains from a saturated porous matrix under the influence of gravity to the total volume of the matrix, per unit drop in the water table.

Specific yield is normally much greater than specific storage, as water released from elastic storage leaves the pores still saturated. Specific yield is often in the range of 0.2 to 0.3, or three to four orders of magnitude greater than elastic storage.

Gravity drainage will proceed until the forces of surface tension and molecular attraction to the matrix grains become equal to the force of gravity. The ratio of the volume of water retained in the pores to the total matrix volume is known as *specific retention*, or S_r . In the upper parts of an unconfined porous

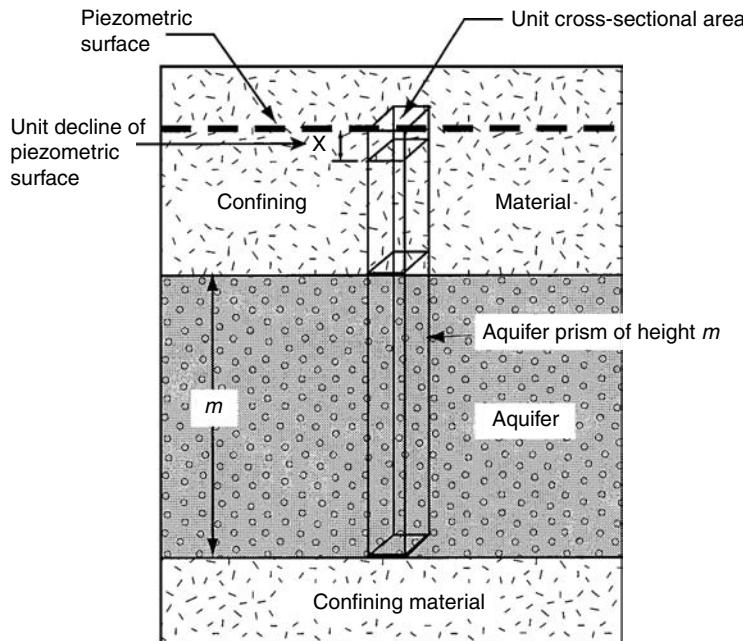


FIGURE 2.7 Illustration of specific storage and storativity. (Adapted from Ferris, J.G., Knowles, D.B., Browne, R.H., and Stallman, R.W., 1962. *Theory of Aquifer Test*. U.S. Geological Survey Water-Supply Paper 1536E.)

medium, where elastic storage is not significant, the sum of specific yield and specific retention equals porosity,

$$S_y + S_r = n \quad (2.21)$$

When an unconfined porous medium is very thick, the lower parts of the medium may also contain water under elastic storage, owing to the increase of pressure and consequent water compressibility and matrix expansion with increasing depth. In this case the total storativity of the medium is expressed as

$$S = S_y + bS_s \quad (2.22)$$

where b is the saturated medium thickness.

Field capacity is used to describe essentially the same phenomena as specific retention, but it is normally used in agricultural soil-moisture studies. It is a function not only of specific retention, but also of the evaporation depth and the unsaturated permeability of the soil (discussed below). A good discussion of all the above relationships can be found in de Marsily (1986).

2.4.9 Intrinsic Permeability

There are several basic properties of both porous media and fluid that will determine the ease with which the medium will transmit a fluid. The most fundamental of all these properties is known as *intrinsic permeability*, k . It is simply a function of the average pore size of the medium and is related to this property as follows,

$$k = Cd^2 \quad (2.23)$$

where d is the average pore diameter, and C is an empirical constant which depends upon packing, sorting, and other factors.

Intrinsic permeability is strictly a function of the medium and has nothing to do with the temperature, pressure, or fluid properties of a particular fluid passing through the medium. It is commonly measured in terms of the *darcy*, after Henri Darcy who developed the relationship known today as Darcy's Law (discussed later). The millidarcy is the commonly used unit, but many authors also use the units of cm^2 .

Although independent of the fluid, the darcy is defined with a fluid standard in order to quantify it as the area through which a fluid with a dynamic viscosity of one centipoise will flow at a rate equivalent to one cubic centimeter per second per square centimeter under a pressure gradient of one atmosphere per centimeter; or

$$1 \text{ darcy} = [(1 \text{ centipoise} \times 1 \text{ cm}^3/\text{sec})/1 \text{ cm}^2]/(1 \text{ atmosphere}/1 \text{ cm}) \quad (2.24)$$

One centipoise is $0.01 \text{ dyne}\cdot\text{sec}/\text{cm}^2$, and one atmosphere is $1.0132 \times 10^6 \text{ dynes}/\text{cm}^2$. If these values are substituted in to the above equation, the result is that

$$1 \text{ darcy} = 9.87 \times 10^{-9} \text{ cm}^2 \quad (2.25)$$

and the same value will apply to the flow of Newtonian liquids or gases of any kind through the medium.

2.4.10 Viscosity

The resistance of a flowing fluid to shear is known as *dynamic viscosity*. Figure 2.8 shows the conceptual and mathematical relationships in the development of the term. Note that a liquid fills the space between two plates; the bottom plate is stationary and the top plate moves unidirectionally at a velocity, U . The

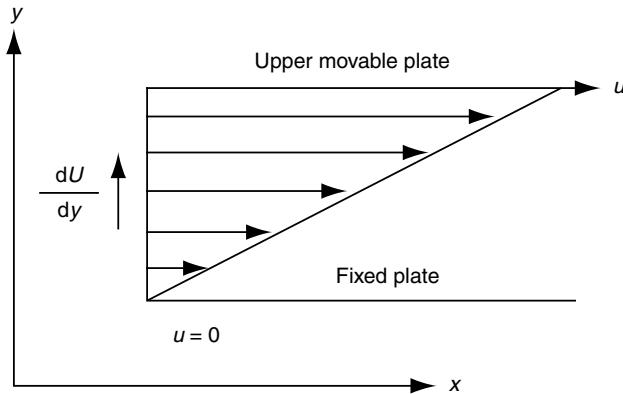


FIGURE 2.8 Illustration of dynamic viscosity as a resistance of a fluid to shear.

liquid is undergoing laminar flow and thus there is a linear change of velocity upward from zero on the bottom plate to U on the top plate. The rate of change of velocity with vertical direction is then dU/dy which is also called the vertical rate of strain. This is caused by the shearing stress, τ , which is

$$\tau = F/A \quad (2.26)$$

where F is the force that pulls the upper plate along at the velocity, U , and A is the surface area of the upper plate.

From these observations it is apparent that

$$\tau \propto dU/dy \quad (2.27)$$

and we assume that there is a constant, μ , that will convert this proportionality into an equation, namely

$$\tau = \mu(dU/dy) \quad (2.28)$$

and

$$\mu = \tau/(dU/dy) \quad (2.29)$$

where the constant μ , is known as the dynamic viscosity. It is specific to the fluid and to temperature, generally decreasing with increasing temperature. The dynamic viscosity has the units of $[FT/L^2]$.

Dynamic viscosity is defined in terms of the *poise* (0.1 Nsec/m^2), after the French mathematician and fluid dynamicist, Poiseuille. The centipoise (0.001 Nsec/m^2) is the most commonly used unit. At 15°C , water possesses a dynamic viscosity of 0.011404 poise, or $1.1404 \times 10^{-3} \text{ Nsec/m}^2$, or approximately one centipoise. In contrast, the dynamic viscosity of the earth's mantle is estimated at 10^{23} centipoise.

Dynamic viscosity is not to be confused with *kinematic viscosity*, η , which is dynamic viscosity divided by fluid density,

$$\eta = \mu/\rho \quad (2.30)$$

The kinematic viscosity has the units of L^2/T . The stoke, named after G.G. Stoke, the British physicist, has the units of cm^2/sec , but the centistoke ($0.01 \text{ cm}^2/\text{sec}$) is a more convenient unit. Water at 15°C has a kinematic viscosity of $1.139 \times 10^{-6} \text{ m}^2/\text{sec}$ or 1.139 centistokes.

2.4.11 Hydraulic Gradient

The hydraulic gradient is simply the loss of head per unit length of flow along a stream line, and is given as

$$I = dh/dl \quad (2.31)$$

a dimensionless number. Development of this relationship can be found in Chapter 3 by J.W. Delleur of this publication and in Hubbert (1940).

2.4.12 Hydraulic Conductivity of Saturated Media

Although the intrinsic permeability describes the water-transmitting property of a porous medium, it does not completely describe the relative ease with which a particular liquid will flow through the medium. The complete description is given by the *hydraulic conductivity*, K , which combines both medium and fluid properties,

$$K = k\rho g/\mu. \quad (2.32)$$

This parameter has the dimension of velocity, generally cm/sec or feet per day, and is a second-order tensor quantity. Its physical meaning is stated as, "The volume of liquid flowing perpendicular to a unit area of porous medium per unit time under the influence of a hydraulic gradient of unity." Earlier literature described this phenomenon as *field coefficient of permeability* with units of gallons per day per square foot. This name and definition are now only rarely used. Table 2.3 lists the hydraulic conductivities of several lithologies.

In the mid-19th century, a French engineer, Henri Darcy, was experimenting with sand filters for a water supply system for the city of Dijon, France. Through a series of experiments with sand-filled tubes (Figure 2.9), he was able to determine the factors that controlled the flow rate of water through the sand (Darcy, 1856).

He discovered one of the most important physical relationships in the science of porous-media hydrodynamics which became known as *Darcy's Law*,

$$q = -K(dh/dL) \quad (2.33)$$

where q is the *specific discharge* or volumetric flow rate per unit area of porous medium perpendicular to the direction of flow, and dh/dL is the hydraulic gradient along the flow path, L . By convention, the negative sign implies that flow is along the direction of decreasing gradient.

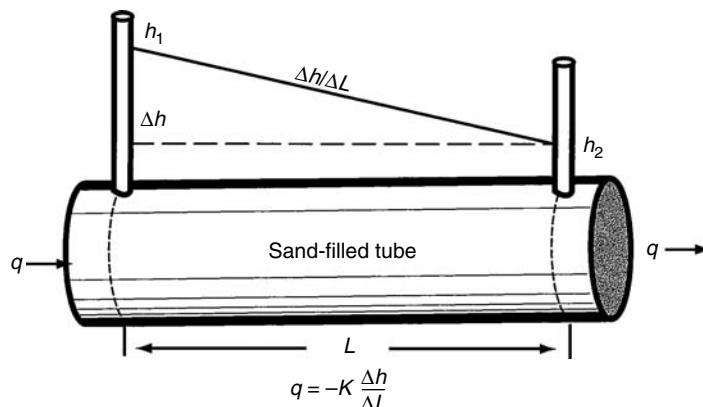


FIGURE 2.9 Figure illustrating tube experiments of Henri Darcy.

It should be noted that q has the dimension of velocity and is called *Darcian velocity*. This term or its usage is not the same as the true *average linear velocity* or *average pore velocity* or *seepage velocity* of flow through a porous medium; the latter property is given as

$$v = q/n_e \quad (2.34)$$

where n_e is the effective porosity. This difference should always be kept in mind when estimating the true velocity of groundwater flow and solute transport.

Although it is common to employ hydraulic conductivity in a general sense in studying or describing the hydraulic properties of a porous medium, it is more advantageous to use the term, *transmissivity* (defined below) to describe the ease with which water moves through a large porous medium body such as a horizontal or layered aquifer. Transmissivity, T (sometimes called transmissibility) is simply the product of hydraulic conductivity and saturated thickness of the aquifer,

$$T = Kb \quad (2.35)$$

and has the dimensions of L^2/T . This dimensional characteristic derives from the definition of transmissivity is “the volume of water per unit time passing through a unit width area of aquifer perpendicular to flow integrated over the thickness of the aquifer,” or $[L^3/(TL^2)]L$ (Figure 2.10).

Transmissivity is usually reported in units of square feet per day or square meters per day. The total rate of flow (Q) through any area (A) of the aquifer perpendicular to the flow direction under the gradient (I)

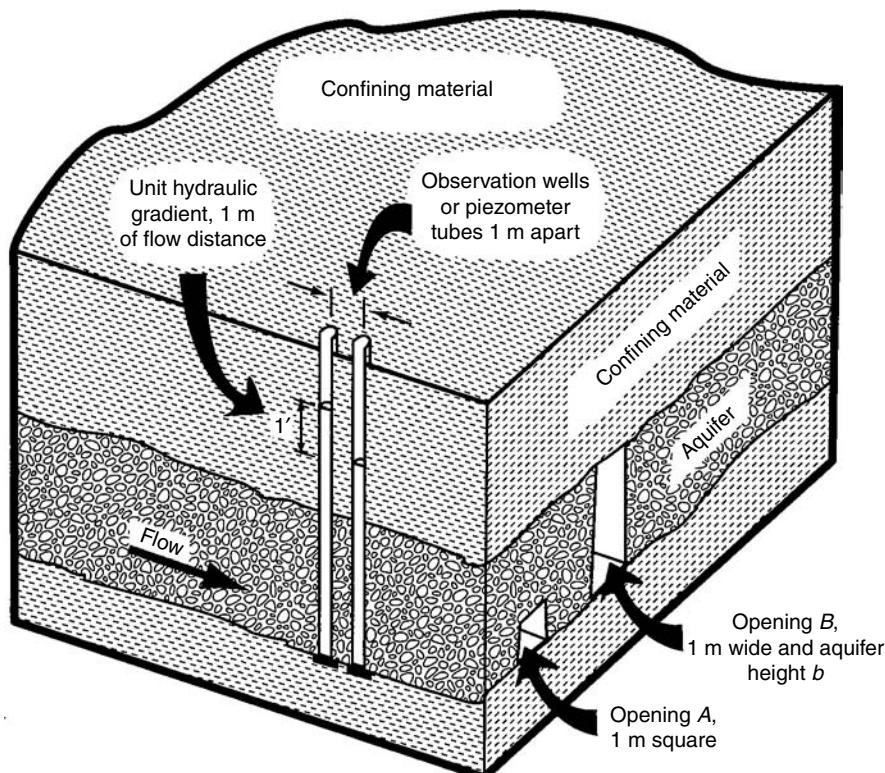


FIGURE 2.10 Illustration of the concept of transmissivity. (Adapted from Ferris, J.G., Knowles, D.B., Browne, R.H., and Stallman, R.W., 1962. *Theory of Aquifer Tests*. U.S. Geological Survey. Water-Supply Paper 1536E.)

is then given as

$$Q = TIA \quad (2.36)$$

2.4.13 Hydraulic Conductivity of Unsaturated Media

Water in unsaturated media is subject to the influences of not only the hydraulic gradient, but also of molecular attraction and surface tension, as described previously. Therefore, the unsaturated hydraulic conductivity is a function of the pressure head (Ψ), which, in unsaturated media will always be negative. Consequently, we write unsaturated hydraulic conductivity as $K(\Psi)$. This is discussed later in the section on the unsaturated zone.

2.4.14 Anisotropy and Heterogeneity

The discussion to this point has assumed that the porous media is *isotropic*, that is, the permeability and hence, the hydraulic conductivity and transmissivity, are equal in all directions at any point in the porous medium. If the parameters differ in value directionally at a point, the medium is then said to be *anisotropic*.

These directional properties are described three-dimensionally in cartesian tensor notation using nine general terms, a_{ij} , which can represent permeability, hydraulic conductivity or transmissivity, but transmissivity is a two-dimensional term that is applied only in a horizontal sense.

$$A_{ij} = \begin{matrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33}, \end{matrix} \quad (2.37)$$

If the principal directions of the tensor coincide with the major cartesian axes, the off-diagonal terms will cancel out and the only terms of interest will be a_{11} , a_{22} , and a_{33} . Specific discharge, velocity, and hydraulic conductivity are vectors. Thus, for example, in a three-dimensional porous medium with anisotropic hydraulic conductivity, specific discharge will be given as

$$\mathbf{q} = q_1 \mathbf{i} + q_2 \mathbf{j} + q_3 \mathbf{k}, \quad (2.38)$$

where \mathbf{i} , \mathbf{j} , and \mathbf{k} are unit vectors in the x , y , and z directions respectively, and q_1 , q_2 , and q_3 are specific discharge components in the x , y , and z directions respectively.

Expansion of this equation using, K_{xx} , K_{yy} , and K_{zz} to represent the main diagonal terms of the hydraulic conductivity tensor, the above equation can be expanded to the following form,

$$\mathbf{q} = K_{xx}(\partial \mathbf{h}/\partial x)\mathbf{i} + K_{yy}(\partial \mathbf{h}/\partial y)\mathbf{j} + K_{zz}(\partial \mathbf{h}/\partial z)\mathbf{k} \quad (2.39)$$

If the principal directions are along the cartesian directions, it is also common to express anisotropy graphically by use of the equation and figure for an ellipsoid in three dimensions, and an ellipse in two dimensions. Using a two-dimensional example, the semi-axes of an ellipse represent the square root of the property along those axes, and the length in any given direction, r , is the square root of the property in that direction (Figure 2.11).

The equation for an ellipsoid is

$$(x^2/a^2) + (y^2/b^2) + (z^2/c^2) = 1 \quad (2.40)$$

where a , b , and c are the semi-axes. If the medium is isotropic, the ellipsoid degenerates into a sphere for which the equation is

$$x^2 + y^2 + z^2 = r^2 \quad (2.41)$$

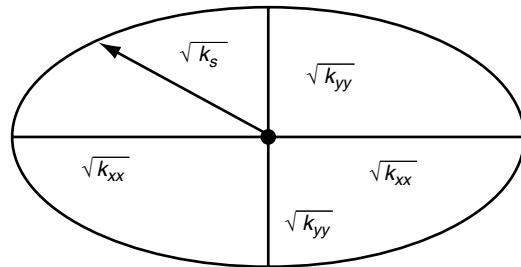


FIGURE 2.11 Elliptical representation of permeability and hydraulic conductivity.

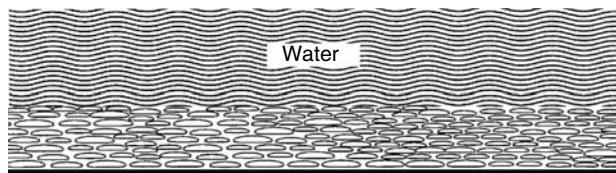


FIGURE 2.12 Natural hydraulic anisotropy in water-lain sediments.

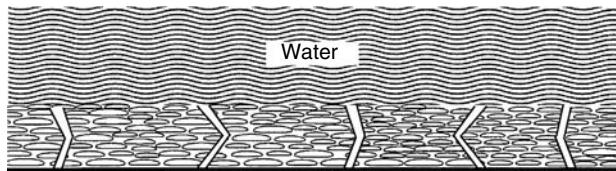


FIGURE 2.13 Anisotropy due to generation of secondary porosity by fracturing.

Isotropy or anisotropy of hydraulic conductivity and transmissivity is controlled specifically by the isotropy or anisotropy of the intrinsic permeability. The latter parameter, being a function strictly of the porous medium, is usually determined by the medium structure, which in turn, is the result of the medium's geological origin and subsequent alterations.

For example, sedimentary deposits almost always possess a lower permeability in the vertical direction than in the horizontal. This is due to the fact that as sediments are laid down from water, they assume a more stable position if possible where the longer axes of the grains and pebbles are oriented horizontally (Figure 2.12).

Subsequent fracturing by tectonic forces in a given direction will add a secondary permeability to the medium that is oriented by the forces and may result in greater vertical permeability than horizontal (Figure 2.13).

If the condition of directional equality of properties is the same from point to point anywhere in the medium, the medium is termed *homogeneous*. If the condition of either isotropy or anisotropy varies from point to point, the medium is then said to be *heterogeneous*. Figure 2.14 demonstrates these four possible descriptions of a medium – (1) homogeneous and isotropic, (2) homogeneous and anisotropic, (3) heterogeneous and isotropic, and (4) heterogeneous and anisotropic.

The average hydraulic conductivity (K_p) perpendicular to a layered sequence of m beds (Figure 2.15a), each of which is either isotropic or anisotropic, can be determined using a weighted harmonic average as

$$K_p = d \left/ \sum_1^m d_m / K_{pm} \right. \quad (2.42)$$

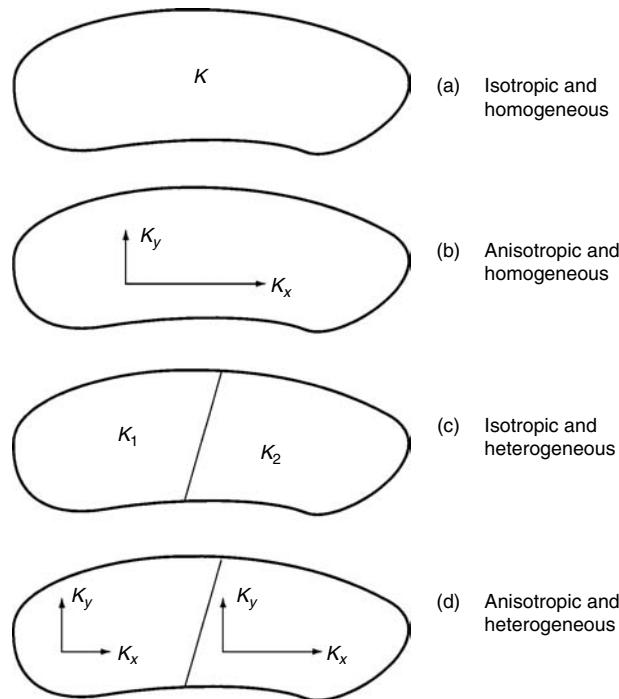


FIGURE 2.14 Summary of possible combinations of isotropy, anisotropy, homogeneity, and heterogeneity.

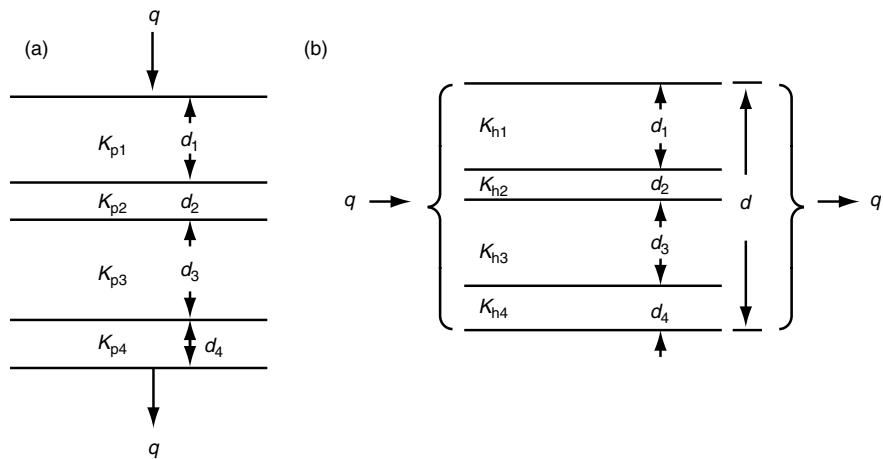


FIGURE 2.15 Hydraulic conductivity of layered systems. (a) flow perpendicular to layers; (b) flow parallel to layers.

where d is the total thickness, d_m is the thickness of each layer, and K_{pm} is the perpendicular hydraulic conductivity of each bed.

The average horizontal hydraulic conductivity parallel to the beds is given as a weighted linear average (Figure 2.15b)

$$K_h = \sum_1^m K_{hm} d_m / d \quad (2.43)$$

where K_{hm} is the horizontal hydraulic conductivity of each bed.

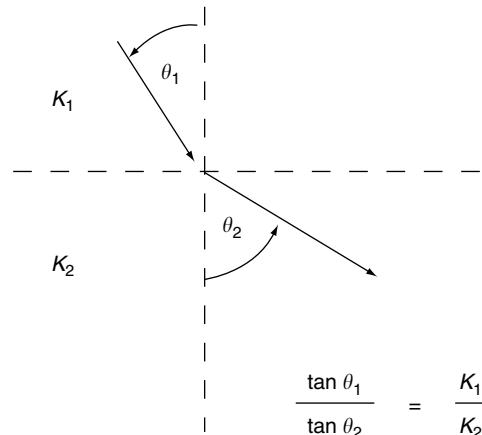


FIGURE 2.16 The Law of Tangents.

If two isotropic beds of porous media of differing hydraulic conductivities are separated by a plane surface, then a particle of water traveling through one bed will refract upon entering the other (Figure 2.16) according to the *Law of Tangents*,

$$\tan \theta_1 / \tan \theta_2 = K_1 / K_2 \quad (2.44)$$

This phenomenon occurs because of the laws of conservation of energy and mass, that is, one side of a particle will move faster or slower than the other at the interface, depending upon whether or not a more permeable or less permeable medium respectively is encountered. This phenomenon is discussed in Bear (1972).

2.4.15 Aquifers

An *aquifer* is defined by Davis and DeWiest (1966) as "... natural zone (geological formation) below the surface that yields water in sufficiently large amounts to be important economically." This definition is very relative and subjective, for a thin bed of sandstone may economically yield water to a well at a rate of $5.5 \text{ m}^3/\text{d}$ for a home but would not be sufficient to supply an irrigation well that required $2700 \text{ m}^3/\text{d}$. Yet, it could be called a aquifer by strict definition.

The most productive aquifers are generally deposits of glacial outwash, karstic carbonates, permeable sandstones, and highly-fractured rocks of all kinds.

An aquifer that is sandwiched in between two impermeable layers or formations that are impermeable is called a *confined aquifer* if it is totally saturated from top to bottom (Figure 2.17a).

If the recharge area for the aquifer is located at a higher elevation than the top of the aquifer, and a well is drilled into the aquifer, the water level will rise above the top as shown. Such an aquifer is known as an *artesian* aquifer; it is named after Artois, France where such wells are common. It should be noted, however, that the well does not have to be flowing to be termed "artesian," although that is the popular conception. A flowing well is known as a "flowing artesian well."

The water level above the top is known as the *piezometric surface* (pressure surface) which is the locus of the piezometric head, and it is not to be confused with the water table discussed below. The piezometric surface occurs above the ground surface because the higher elevation of the recharge area causes the pressure head to rise to such an elevation. The water within the aquifer will be partly under elastic storage. Pumping a well or allowing it to flow will release the water from storage.

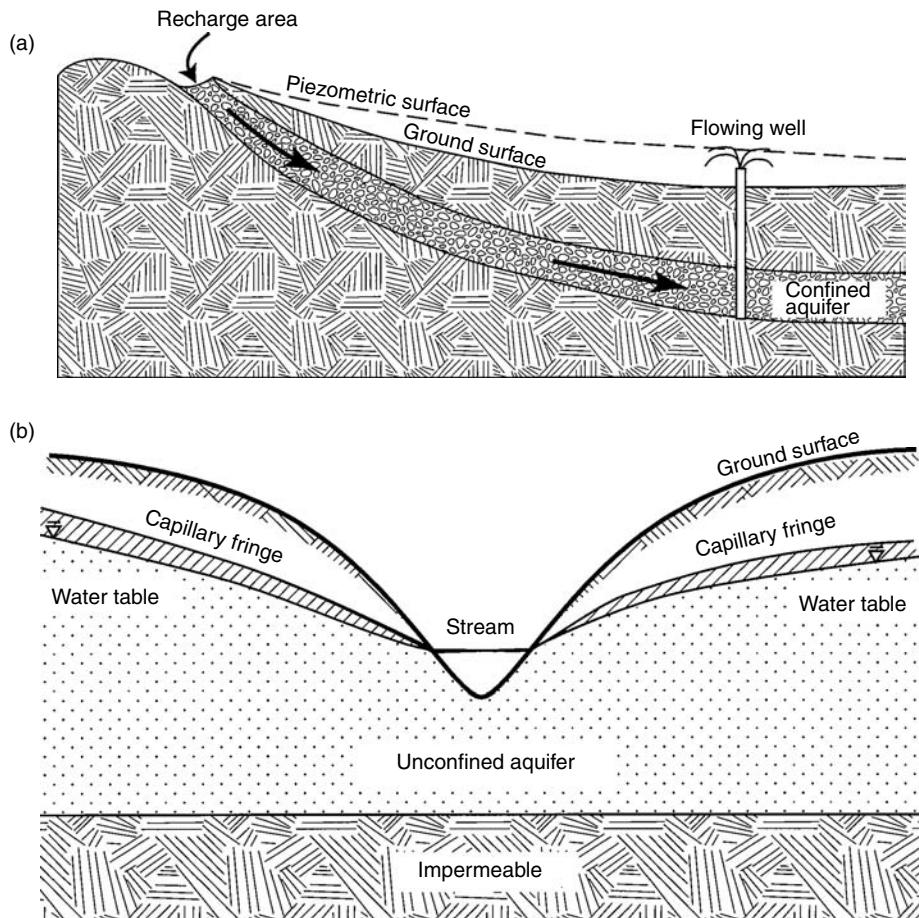


FIGURE 2.17 Kinds of aquifers. (a) Confined aquifer; (b) unconfined aquifer.

Artesian or confined aquifers are common in glaciated regions of the world where a body of outwash sand and gravel may have been covered by clay-rich till or lacustrine sediments from a subsequent glaciation. They may also occur in layered bedrock such as the famous Dakota aquifer of Cretaceous age which rises in the west to lap onto the crystalline rocks of the Black Hills in South Dakota and Wyoming where it is recharged by rain and snowmelt, and extends for a few hundred miles east to Minnesota. Along the way, it is sandwiched between impermeable or slightly permeable shales which maintain its confined and artesian condition. Wells drilled into this aquifer, even hundreds of miles from its recharge area, often flow from the pressure within the aquifer.

An *unconfined aquifer* possesses no overlying confining layer, but may sit upon an impermeable or slightly permeable bed. Therefore, the top of the unsaturated zone of an unconfined aquifer is most often the ground surface, and the top of the saturated zone is usually under negative pressure or tension. This latter property gives rise to the definition of the *water table* which is simply the surface where the relative pressure is zero, that is, the absolute pressure is atmospheric (Figure 2.17b). Immediately above the water table, the medium is still saturated but the water is held by capillary forces, thus creating a negative pressure head, or tension. This tension can exist even though the pores may be saturated between the water table and the top of the capillary fringe. Below the water table, the water pressure increases with depth.

Figure 2.18 shows the cross section of a perched aquifer. These are pockets of sand or gravel or other material that sit on top of impermeable materials such as clay. During periods of high recharge rates, these

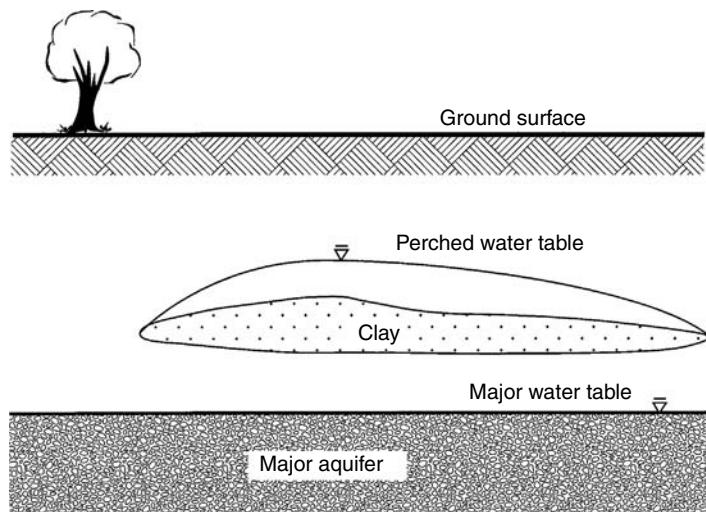


FIGURE 2.18 A perched aquifer.

aquifers may become saturated and actually contain enough water for usage on a temporary basis. Often though, after the onset of dry weather, and after pumping, the perched aquifers will become dry.

Many people make the mistake of finishing a well in a perched aquifer, thinking they are saving money by not having to drill deeper for water, but end up running out of water when the perched aquifer is depleted. Perched aquifers are common in glaciated regions because a till sheet may contain numerous sand and gravel deposits at various depths that are actually perched aquifers.

A map of the water table in a region can be constructed from water levels measured in wells in the region. In most cases the water table will be a subdued replica of the surface topography. Although the term, “piezometric surface,” is generally reserved for confined aquifers with a high pressure head, both water table and piezometric surface can be grouped into the term, *potentiometric surface*.

There is yet another kind of unit between the permeable and the impermeable, known as an *aquitard*, which is semi-permeable. A totally impermeable bed is termed an *aquiclude*. An aquifer sandwiched in between an aquitard and an aquiclude, or between two aquitards, is called a *semi-confined* or *leaky* aquifer. It usually possesses properties common to a confined aquifer, such as a high piezometric surface, but some water will flow into or out of the aquifer through the aquitard(s) (Figure 2.19).

In reality, few rocks or formations are entirely impermeable and the above definitions may be relative to the time period and intensity of pumping stresses imposed on the aquifer. For example, a low-permeability stratum overlying an aquifer may indeed serve as a confining layer under normal and natural pressures. However, when the added stress of heavy pumping is added to the aquifer, the change in head within the aquifer may be great enough to cause some water to move through the overlying bed and the system will then be semi-confined.

Pumping a well in an confined aquifer will release water from elastic storage determined by the elastic specific storage, as discussed above. Pumping from an unconfined aquifer releases water from gravity drainage of the actual media itself, and this is controlled by the specific yield. Recalling that specific yield may be as much as three or four orders of magnitude greater than the elastic storage coefficient implies that much more water can be obtained from an unconfined aquifer per unit of energy expended than from a confined aquifer, providing that the permeability and thickness of the two different aquifers is the same.

Well yields of a semi-confined aquifer with the same properties as the confined and unconfined aquifers will generally yield more water per unit of energy expended than the confined aquifer, but less than the unconfined, because the semi-confined will yield water from elastic storage plus some from leakage through the semi-confining bed(s), but neither sources are as great as the yield from an unconfined aquifer in most cases.

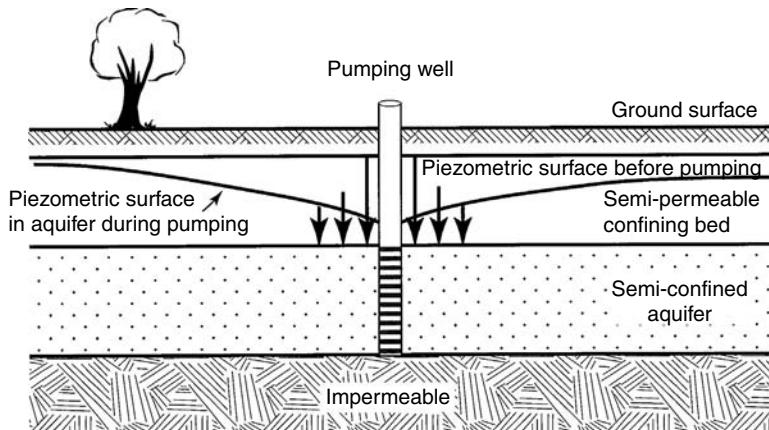


FIGURE 2.19 A semiconfined or leaky aquifer. Vertical arrows indicate relative rates of leakage into aquifer as a function of the distance from the well.

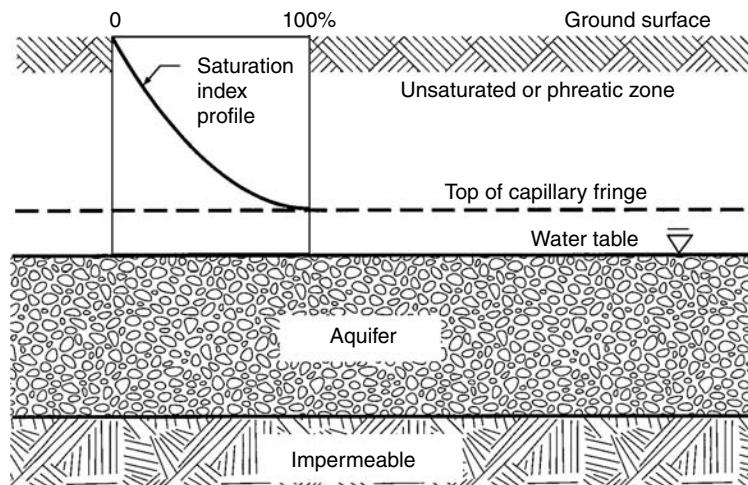


FIGURE 2.20 Saturation index profile vs. depth in an unsaturated soil.

Although more water can be obtained from unconfined aquifers than the other types, they are more vulnerable to contamination from surface sources than the others because of the lack of a protective confining bed above them. It is for this reason that special care must be taken to insure the removal and/or remediation of contamination sources above an unconfined aquifer in use or being considered for use as a water supply. This is discussed further under the subject, *Well-Head Protection Programs* below.

2.5 Water in the Unsaturated Zone

2.5.1 Moisture Content vs. Depth

Figure 2.20 shows the general profile of the saturation index from the surface to the water table in an ideal homogeneous and isotropic aquifer. For purposes of discussion, this figure assumes that there have been no recent rains or snowmelt events and that at the surface the moisture content is zero.

From the ground surface to the top of the capillary fringe, the saturation ratio increases from zero to unity and will remain so to the bottom of the aquifer. Below the capillary fringe is the water table at

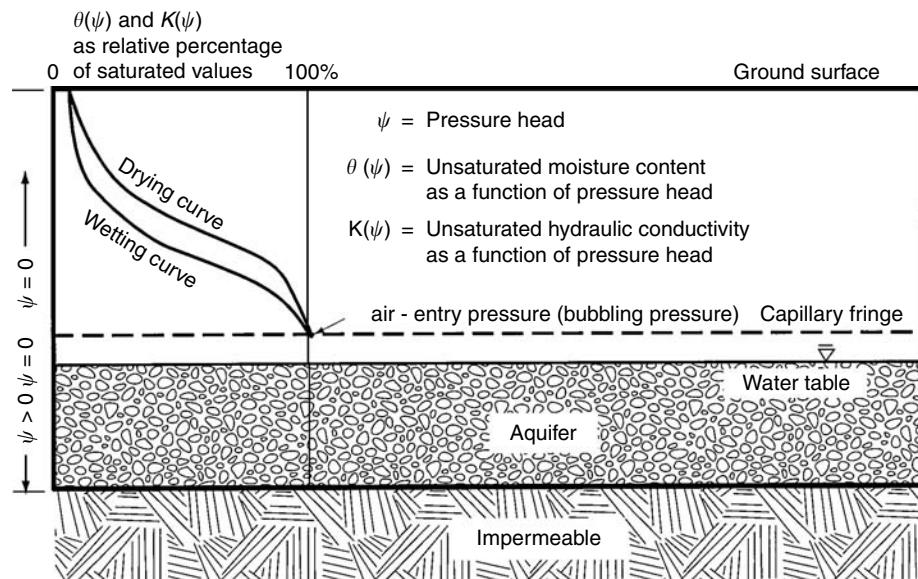


FIGURE 2.21 Soil moisture characteristic curves.

atmospheric pressure which marks the surface delineating the tension zone above from the pressure zone below. The relationships are all the result of a dynamic equilibrium of several forces and conditions — capillary attraction, porosity, initial moisture recharge at the surface, gravitational acceleration, vertical permeability, and evapotranspiration from the surface.

The functional relationship of the moisture content and hydraulic conductivity of the unsaturated profile can be demonstrated by the use of *characteristic curves* as shown in Figure 2.21.

At very low (negative) values of the pressure head, ψ , both the moisture content and hydraulic conductivity are at minimal values for the system. With increasing values of ψ , they increase to become constant at the top of the capillary fringe where the saturation ration is unity, indicating full saturation. Here also, the hydraulic conductivity is no longer a function of the pressure or moisture content, but attains a maximum value at which it remains, providing no changes in the porous media occur. The pressure head at this point is called the *bubbling pressure* or the *air-entry pressure*; at a lower pressure, air will enter the system causing loss of saturation. As the pressure increases with increasing depth in the system, it will become atmospheric, or zero relative to positive and negative extremes of the pressure range, marking the water table. With still increasing pressure, both moisture content and hydraulic conductivity remain constant.

The lag between the wetting and drying curve shapes is called *hysteresis*. This is due to the fact that initial wetting of the porous medium is enhanced by strong capillary and adhesive forces, but these same forces tend to hold water in the unsaturated pores and it requires more suction or greater negative pressures to drain the pores, as illustrated by the offset of the drying curve from the wetting curve.

It should be noted that characteristic curves of fine-grained materials such as silt and clay will be much steeper in shape than those of coarse sand and gravel. This is due to the fact that in the coarse-grained material with relatively large pore diameters, the effects of capillarity and adhesion are not as great as in the fine-grained material. It is for this reason that the capillary fringe in clay is much thicker (one or two meters) than in coarse materials (a few centimeters at most).

In reality, soil moisture is not distributed evenly in the unsaturated zone, but varies both vertically and horizontally due to variations in soil types, crop types, infiltration rates, etc. Recently, new work has been done in expressing the variability of soil moisture in unsaturated chalk with electrical tomography (Andrews et al., 1995). This method, although in the experimental stage, does allow visual imaging of soil moisture contents based on subsurface electrical resistivity of various zones.

2.5.2 Recharge and Infiltration Capacity

Recharge is the actual entering of water into an aquifer, whereas *infiltration* is movement of water from the surface into the ground. Usually, the recharge source is surface water from precipitation and to a lesser degree, irrigation or artificially constructed recharge pits or losing streams (discussed below). The actual recharge rate is controlled by several factors: (1) the amount and rate of precipitation not lost to surface runoff and evapotranspiration, (2) the initial soil moisture content or saturation ratio of the soil, (3) the elevation of the recharge surface relative to the discharge area, (4) the horizontal hydraulic conductivity of the aquifer being recharged and its hydraulic gradient which determine the rate at which recharged water will be carried out of the recharge area, (5) the vertical hydraulic conductivity of the soil being recharged, and (6) the presence of man-made alterations to the subsurface such as drainage tiles that carry water away to run off in surface streams.

People are often surprised to learn how little precipitation is actually recharged in even moist climatic zones. For example, the state of Indiana receives approximately 91.4 cm of precipitation water per year. Yet, a statewide average of only about 10% (9.14 cm) is actually recharged into the groundwater system, the rest being lost to evapotranspiration, surface runoff, and subsurface drainage tiles. Recharge is not evenly distributed with perhaps eight to ten times the amount going into more permeable sands and gravels than into till.

Figure 2.22 shows a hypothetical cross section of a recharge/discharge system with recharge taking place in the higher elevation and discharge in the lower. The stream lines are perpendicular to the equipotential lines. Discharge areas are often seen on the ground as wet spots with seeps or springs and often with lush vegetation.

Recharge rates are not easy to determine due to the variability of soils and the necessity of extensive field instrumentation. One very effective method has been the actual tracking of infiltrating water by analyzing the concentrations of atmospheric tritium at depth in the soil. This procedure was very successful in Indiana (Daniels et al., 1991), and in Minnesota (Kanivetsky and Rumynin, 1993). Tritium in the atmosphere, due to atmospheric nuclear weapons tests in the 1950s and 1960s, is incorporated into precipitation which is eventually recharged into the ground. The concentration of tritium in a vertical drill core can be used to estimate the rate of recharge.

In addition to elevation and soil type and its variability, recharge is controlled by the *infiltration capacity* of the soil. This phenomenon was studied by Horton (1933) and is the maximum rate at which a soil will permit the entry of water. It is generally a time-dependent parameter because it depends upon the rate of change of the saturation ratio and hence, the moisture content.

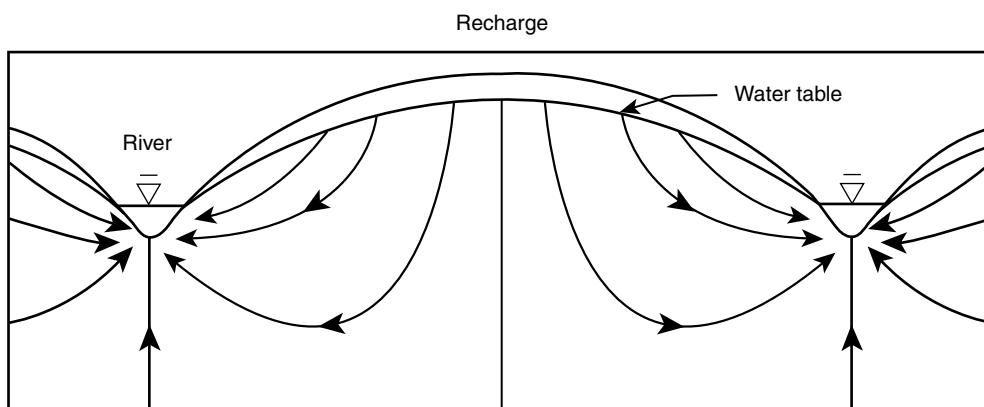


FIGURE 2.22 Hypothetical recharge/discharge system. (Adapted from Lohman, S.W. 1972. *Ground-Water Hydraulics*. U.S. Geological Survey Prof. Paper 708, U.S. Government Printing Office, Washington, DC.)

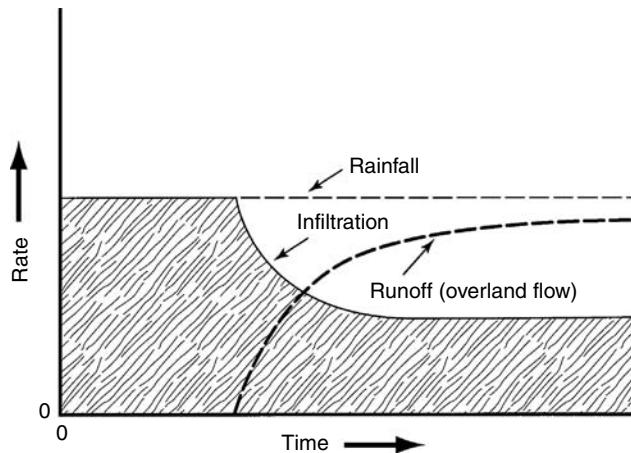


FIGURE 2.23 Illustration of infiltration capacity.

This relationship is illustrated with Figure 2.23. Here, an initially dry soil is rained upon at a constant rate. Initial infiltration into the soil is rapid as the empty pores begin filling. As infiltration continues, the pores gradually lose capacity for additional water at the initial rate. As a result, the infiltration capacity decreases to an eventual steady-state rate where the infiltrating water enters as the same rate that it is transported downward by porous media flow.

Once the infiltration rate falls below the precipitation rate (or snowmelt rate), the excess water will become runoff known as *overland flow*. Generally, this will happen after all surface storage requirements are met, for example, leaf storage in trees, crops, and grasses; and depression storage in hollows in the ground. From these considerations, it is obvious that the most effective rain for watering plants and recharging groundwater supplies is the slow steady kind, and not the intense, high-rate storm, most of which runs off and does not infiltrate.

2.5.3 Hydraulic Conductivity and Specific Discharge through the Unsaturated Zone

When the infiltration capacity is reached, the unsaturated hydraulic conductivity is then at its maximum. When the soil is initially dry, water will not flow any significant distance until the water content in the soil is sufficient and the pressure head becomes less negative, as demonstrated with the characteristic curves above. The specific discharge is then given as

$$q = -K(\psi) \quad \partial h / \partial l, \quad (2.45)$$

or

$$q = -K(\theta) \quad \partial h / \partial l \quad (2.46)$$

where θ is the water content of the soil. Three methods for computing unsaturated hydraulic conductivity as a function of water content are given below and are summarized in Rawls et al. (1993). The equations below are dimensionless making them applicable to any consistent set of units.

Brooks and Corey (1964) give the formula for hydraulic conductivity as

$$K(\theta)/K_s = [(\theta - \theta_r)/(n - \theta_r)]^m \quad (2.47)$$

where K_s is saturated hydraulic conductivity, θ_r is residual water content, n is porosity, and

$$m = 3 + 2/\lambda \quad (2.48)$$

where λ is the pore-size index (Brooks and Corey, 1964).

Campbell (1974) presents the following formula:

$$K(\theta)/K_s = (\theta/n)^m \quad (2.49)$$

where

$$m = 3 + 2/\lambda \quad (2.50)$$

Van Genuchten (1980) presents the formula:

$$K(\theta)/K_s = [(\theta - \theta_r)/(n - \theta_r)]^{1/2} \{1 - [((\theta - \theta_r)/(n - \theta_r))^{1/m}]^m\}^2 \quad (2.51)$$

where

$$m = \lambda/(\lambda + 1) \quad (2.52)$$

2.5.4 Residence Time

Water exists in different forms and in different places and for widely varying times in the hydrologic cycle. Freeze and Cherry (1979) state that soil moisture may stay in place for two weeks to one year. This rate is an average and depends on a wide variety of conditions such as permeability, infiltration rate, plant use, hydraulic gradient, and anthropogenic withdrawal rates.

2.5.5 Subsurface Stormflow or Interflow Zone

In many regions of the world, especially in tilled crop land and forest litter on sloping ground, infiltration into the upper one or two meters of soil will be quite rapid, but continued infiltration into the substratum will be much lower. The upper part often possesses much greater permeability than the lower part due to tilling, root activity, and movement of soil by worms, insects, and small mammals.

If the hydraulic conductivity of the lower part is less than the upper, then subsurface runoff will occur in the upper part, or the *subsurface stormflow or interflow zone*, as it is called. This zone may carry water all the way to a stream or it may discharge it along the way at a lower elevation. Figure 2.24 illustrates the interflow zone.

2.5.6 Atmospheric Discharge and Seepage Face

In a discharge area, at the juncture point where groundwater enters a stream or lake, the water table will coincide with the elevation of the surface water. Beyond this point, the hydraulic gradient is zero and groundwater has completed its journey from recharge to discharge.

Also at this point, there will generally be a zone of wet soil extending above the juncture point (Figure 2.25). This zone, called the *seepage face*, is wetted by capillary attraction of the water above the water table and it is essentially the surface exposure of the capillary fringe. A small fraction of the discharge from the groundwater system will escape into the atmosphere at this point. The height of the seepage face will depend upon the grain size of the soil – greater in fine-grained and lesser in coarse-grained sediments.

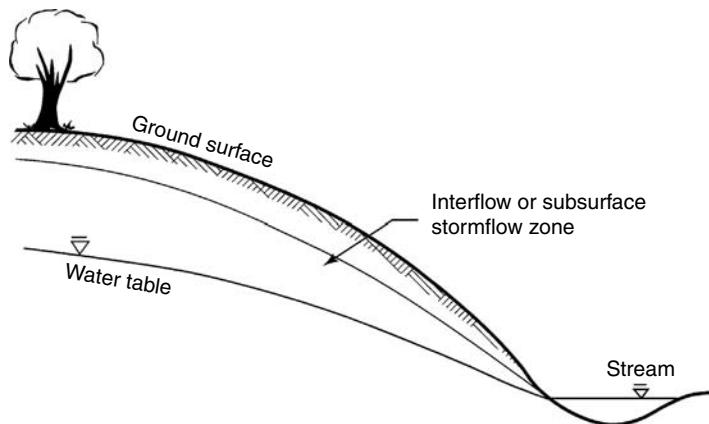


FIGURE 2.24 The interflow or subsurface stormflow zone.

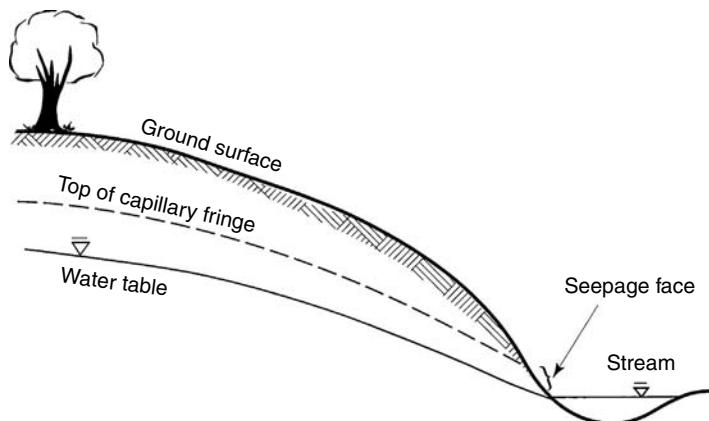


FIGURE 2.25 The seepage face.

2.5.7 Discharge to the Saturated Zone

Water in the unsaturated zone will percolate downward from the surface and gradually increase in content to a maximum, where the saturation ratio is unity at the top of the capillary fringe. At this point, it will flow downward to the water table under saturated flow conditions. The rate of flow to the saturated zone will depend to a large degree on the vertical permeability and the rate at which the groundwater in the saturated zone is being carried away by lateral Groundwater flow.

2.6 Water in the Saturated Zone

In saturated pores, the forces of adhesion are equal in all directions, so there is no directional attraction. The only dynamic or motive forces in action are gravity and the force represented by the gradient of the fluid potential. Resistive forces due to viscosity (resistance to shear) work in opposition to the dynamic forces.

Darcy's Law is assumed to be in effect in porous media flow and the flow is assumed to be laminar, that is, the *Reynold's number*

$$R_e = \rho v d / \mu, \quad (2.53)$$

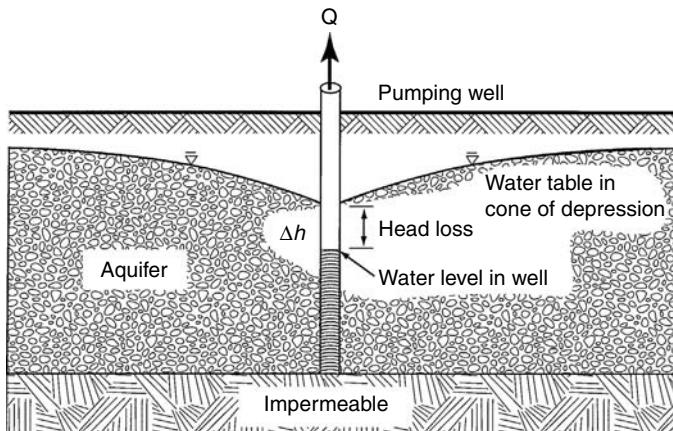


FIGURE 2.26 Flow to pumping well with head loss.

has a value from 1 to 10, where ρ is water density, v is average pore velocity, d is average pore diameter, and μ is dynamic viscosity of water at a given temperature.

The average pore diameter is probably the best estimate of pore properties that can be attained on a large scale. However, the pore size in a medium can range over even a few orders of magnitude in very heterogeneous media with high uniformity coefficients (greater than 4). Therefore, it is unlikely that laminar flow occurs in all pores at the same time. Large pores (including fractures and karst conduits) may allow passage of water at velocities high enough to be non-linear laminar or transitional ($4 \leq R_e \leq 100$), and in some cases, turbulent ($R_e > 100$). In small pores, the frictional resistance may be too great to allow flow faster than laminar. In spite of these uncertainties, hydrogeologists generally accept the assumption of laminar flow throughout a granular porous media with $R_e < 100$ for the sake of simplifying the mathematics of flow.

Groundwater may flow for some distance over a wide area in the laminar state, but as it approaches a discharge point (e.g., a spring or well) much narrower than its upgradient flow field, the streamlines will crowd together and the flow velocity will increase to maintain the same volumetric discharge rate. Near or at the point of discharge the velocity will often be great enough to be actually turbulent.

Under such conditions, the specific discharge will be expressed as

$$q = -K(\partial h/\partial l)^m, \quad (2.54)$$

where $m > 1$.

Figure 2.26 illustrates a situation where flow can change from laminar to turbulent as it approaches a pumping well. Assuming that the aquifer is homogeneous and isotropic and that the flow is steady-state, the flow lines converge as they approach the well.

At some distance from the well the equation of steady-state flow is given as

$$q = -K(dh/dr) \quad (2.55)$$

where r is the radial distance from the well center. At or near the well, the equation becomes

$$q = -K(dh/dr)^m \quad (2.56)$$

where $m > 1$ and generally closer to 2.

Steady-state laminar flow into the well will require a gradient such that the water level in the well casing is the same elevation as that immediately outside the well casing in the aquifer. Turbulent flow into the

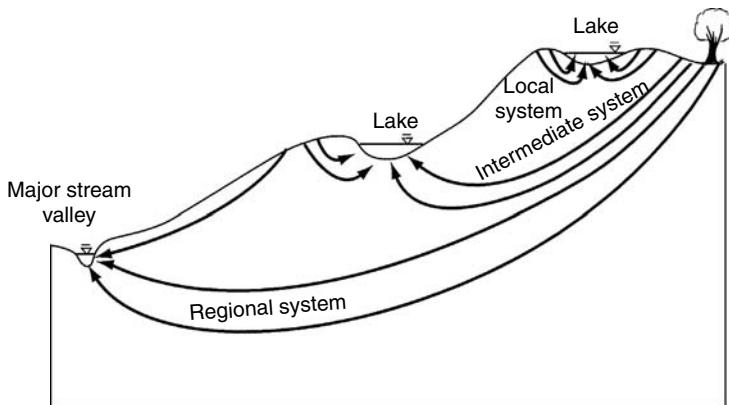


FIGURE 2.27 Local, intermediate, and regional flow systems.

well requires more energy and thus a steeper gradient than laminar flow. To accomplish this, the water level in the well will have to drop below the level required for laminar flow as demonstrated in Figure 2.26.

This extra head drop is known as head loss or *well loss*, and is undesirable. Well loss means extra energy is needed which reduces the efficiency of the pumping system and drives up the cost of pumping water. In addition, because a greater pressure drop occurs with head loss, gasses will be more likely to come out of solution and cause precipitation of lime (CaCO_3) and iron and manganese oxides (Fe_2O_3 and MnO_2 respectively) onto the well screen which causes clogging of the screen and further reduces the efficiency of the well. For further discussion, see chapter on *Well Hydraulics*.

2.6.1 Residence Time

Groundwater occurs over a wide range of depths, from near the surface to even thousands of meters below. It has been estimated that the residence time of groundwater can vary from as little as two weeks to more than 10,000 years (Freeze and Cherry, 1979). Some water presently being extracted from aquifers in the north central United States is believed to have been recharged during the Pleistocene Epoch, more than 10,000 years ago.

Tóth (1962) introduced the concept of flow system lengths and applied it to the prairie pothole regions of the Canadian high plains. The same concept (the "Prairie Profile") applies in many other regions and to many other Groundwater regimes. Tóth's three systems are the (1) local systems, (2) intermediate systems, and (3) regional systems, as demonstrated by Figure 2.27.

The local flow system consists of relatively short cells with discharge a few tens to a few hundreds of meters from the recharge areas. Such systems generally contain water of the best quality, that is, the lowest concentration of total dissolved solids, because of its short residence time and the short flow paths. Residence times may be only a few months or a few years. Such systems are commonly recharged on the interfluvial high ground between glacially-created ponds, and they discharge to the ponds a few tens or hundreds of meters away.

Intermediate systems cover a larger area of perhaps a few hundred to a few thousand meters and residence time may be tens of years. Dissolved solids concentrations are greater than in the local systems.

Regional systems may cover many kilometers and be as large as a state. A typical example is the Dakota Aquifer which is recharged in the Black Hills of South Dakota and yet extends for a few hundred miles eastward.

Most regional systems discharge into major rivers or large lakes, such as the Mississippi or the Great Lakes, respectively. In regional systems, the total dissolved solids can be quite high. For example, it is not uncommon to find as much as 2,500 to 3,000 mg/l of TDS in the Dakota Aquifer. Residence times can exceed several thousands of years.

Fossil water is that water trapped deep in the earth and incorporated with original sediments. Generally, it is of geologic time in age and contains such high concentrations of dissolved minerals that it is often considered a brine.

Connate water is groundwater which has circulated so deeply in the flow system that it too is excessively old, approaching geologic time in many cases. It also is briny in most cases.

Both kinds of water may contain large concentrations of magnesium, calcium, sodium, potassium, chloride, sulfate, carbonate, and bicarbonate. Often these brines contain dissolved matter in concentrations exceeding 100,000 mg/l, and as such are known as *bittern brines* (Hem, 1985).

2.6.2 Surface Discharge

When streams are incised into aquifers, or even into saturated soil or rocks which cannot be classified as aquifers, the streams normally serve as discharge lines or sinks for groundwater as it leaves the groundwater system. Groundwater is seldom static and streams carry water back to lakes and oceans where it is naturally recycled by evaporation.

Between storms and runoff events, stream flow is maintained by groundwater discharge known as *baseflow*, as long as the water table remains above the stream bottom. After a storm or snowmelt event, the stream will flow at its highest rate and will continue to flow at an ever decreasing rate until the next precipitation event. During dry weather, the water table may even drop below the stream bottom which will cause stream flow to cease. The history of these events and relationships is shown in a *hydrograph* in Figure 2.28.

The falling limb of the hydrograph is a record of both overland flow and baseflow. Over a period of many years, such falling-limb records of a particular stretch of stream can be analyzed by regression analysis to derive a number unique to that particular stretch and its upstream contributing basin. This number, a , the *base flow recession constant*, is used in the following equation to compute flow rates,

$$Q = Q_0 e^{-at}, \quad (2.57)$$

where Q is the flow rate at any time, t , Q_0 is the initial flow rate, and t is elapsed time since time of the initial flow rate. The reader may note that this equation, a decay-curve equation, is of the same form as the classical decay curve equation which is also applicable to radioactive decay, using different units.

The baseflow recession constant of a stream or a certain stretch of a stream may change only slightly over a long period of time if neither the environment nor the climate does not change appreciably. However,

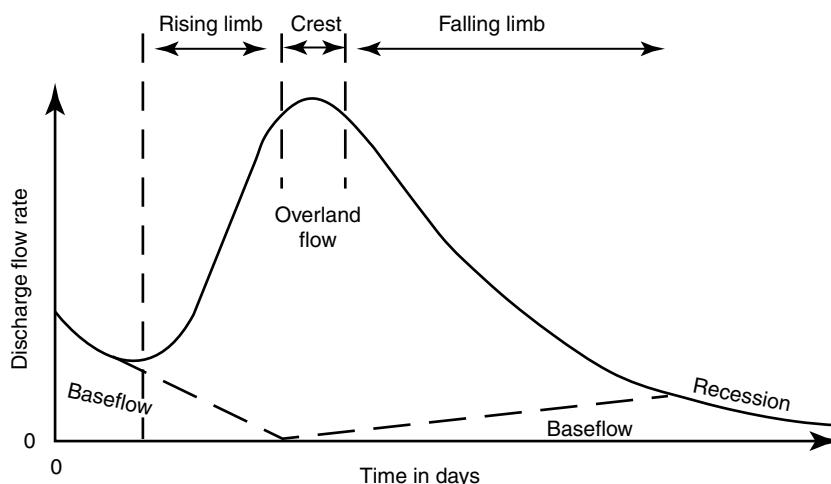


FIGURE 2.28 Hydrograph of stream flow.

anthropogenic alterations such as timbering, farming, tiling of fields, construction of roofs and pavements, and damming of the stream or diversion of its water can drastically change the recession constant in a fairly short time.

In water-management and water-budget studies, it is often necessary to separate out the various components of the hydrograph to determine the contribution of groundwater, surface water, etc. This is not an easy task, but four commonly-employed methods are described in the following paragraphs:

- (1) *Direct-measurement method:* Accurate knowledge of the groundwater hydraulic gradient on both sides of the stream and an accurate estimate of the hydraulic conductivity must be known beforehand. The first can be gotten by installation of piezometers, in which the water levels are regularly measured. Hydraulic conductivity can be determined with pumping tests or from laboratory determinations from grain-size distributions. Direct measurement of stream flow is taken carefully, and evapotranspiration is estimated by proper formulas. Direct precipitation is also recorded.

Using Darcy's law, the specific discharge from groundwater can be calculated to yield an estimate of groundwater contribution to total stream flow. The groundwater discharge, precipitation, and evapotranspiration are then known, they can be subtracted from total stream flow at various times in the time period of the hydrograph to determine the overland flow component during these times.

- (2) *The curve tangent method:* Pilgrim and Cordery (1993) describe a graphical estimation method in which tangents are drawn to recession curves at point where overland flow starts and is assumed to end (breaks in the hydrograph slopes). Under conditions where maximum baseflow discharge is below 10% of the maximum discharge, a straight line can be drawn between the two tangent points which will be an acceptable estimate of the baseflow discharge component.
- (3) *The basin-area method:* Linsley et al. (1975) gives the formula for calculating D , the number of days between the storm peak of a hydrograph and the end of overland flow, for a basin of area, A (in square kilometers), as,

$$D = 0.827A^{(0.2)} \quad (2.58)$$

The hydrograph recession limb that existed before the storm is extended until it is under the hydrograph peak. A straight line drawn from this point to the point on the hydrograph corresponding to D is then considered to be the graph of the baseflow component.

- (4) *The chemical and isotope method:* Stream water and groundwater supplying the stream's base flow can be analyzed for a variety of chemicals and isotopes during periods of normal flow to obtain normal background levels. Typical analytes include ^2H , ^3H , ^{18}O , Na^+ , Ca^{2+} , Mg^{2+} , Cl^- , SO_4^{2-} , and HCO_3^- . Electrical conductance has also been used. During periods of high flow, the stream water is sampled again and analyzed for the same isotopes (Freeze and Cherry, 1979).

If the chemicals or isotopes that were high in concentration in the groundwater are also high in the stream water, then it can be assumed that a significant percentage of the surface flow is actually being contributed by groundwater. On the other hand, if the groundwater isotopes are low in the stream water, then it can be assumed that most flow originates from surface runoff.

Highly permeable soils will contribute significantly to surface flow during precipitation. Low-permeability soils will contribute less, and the isotope analyses will show these differences.

A *spring* is a natural discharge point for groundwater to the surface and the atmosphere. Several geological factors can cause springs of various types to occur, as listed below:

- (1) *Geological contact springs:* When a saturated permeable formation sits on a stratum of low permeability, it will often result in a spring at the contact. This is a common occurrence in glaciated regions where saturated sand and gravel sits on till. If the water has not been contaminated by surface sources, the water from such springs is often of good quality because it has not traveled far before discharge and therefore, little mineral matter has been dissolved.

- (2) *Karst springs:* In limestone and dolomite formations, springs will often discharge from the points on a valley side where caves have been intersected by incising stream valleys. These are common in areas of the world where the valleys are deep enough and precipitation and recharge are great enough to allow a water table to exist well above the streams.

The Fontaine de Vaucluse in southern France discharges between 8 and 150 m³/sec and is one of the largest springs in the world. It is supplied by precipitation on the fissured carbonates of the Vaucluse Plateau. The extensive network of karst conduits ends at the point of discharge. Another kind of karst spring is the sinkhole spring. This feature is simply a spring where a sinkhole intersects karst conduits. Springs of this kind are very common in the limestone of the state of Florida, with names like Silver Springs, Tarpon Springs and others. Many large lakes in central Florida are actually large sinkhole springs.

- (3) *Structural springs:* When there is significant movement along faults such that permeable and water-bearing formations are juxtaposed against impermeable rocks; or fracturing opens pathways into aquifers, water can emanate from these features to discharge at the surface. Such springs are not uncommon in the Great Basin of the United States of America where tension in the earth's crust has produced a large area of block faulting Nevada, Utah, and parts of Arizona and Idaho. One such example is the Ash Meadows Springs in the Amargosa Desert of Nevada, just east of Death Valley, where discharge along a fault spring in carbonates is approximately 2.1×10^7 cubic meters per year.
- (4) *Depression springs:* If the land surface is lowered below the water table, a depression spring will form and will actually create a small local flow system with the spring being the discharge point. Such depressions can be created by wind in areas of fine sand, or by land slides, or tectonic activity, or even the collapse of land into underlying sinkholes or mine tunnels.

Many springs are known for their size and have been utilized for water supplies, spas, and tourist attractions. Examples are Steamboat Springs, Colorado; Hot Springs, Arkansas; Big Springs, Texas; White Sulfur Springs, West Virginia; and many others.

Ponds and lakes can form in many ways. Commonly in glaciated regions like the Midwestern United States, Canada, and northern Europe, present-day lakes and ponds are the results of ice blocks that were covered with drift and later melted to form permanent depressions in the new landscape. Glaciers in mountain valleys may leave large debris dams at their termini which cause ponding of water. Sinkhole lakes are common in Florida, Kentucky, Indiana, and other karst areas. Landslides can dam up rivers. Tectonic forces can cause certain areas, as in the Basin and Range areas of the United States, to subside relative to the surrounding areas, for example, Death Valley which contained a large lake during the Pleistocene Epoch. Volcanic craters may contain ponded water like Crater Lake, Oregon. Oxbow lakes can form when large rivers meander and eventually, a meander loop is cut off from the main stream.

Surface bodies of water are nearly always groundwater discharge areas. Very large lakes, for example the Great Lakes in North America, serve as discharge areas for regional, intermediate and local systems upgradient from them. The smallest ponds of few acres in area discharge water from local systems.

The dynamic role of lakes and ponds in the movement of groundwater is most significant for two reasons. In the first place, the generally low topographic positions of these bodies in relation to the upgradient flow system feeding them cause water to flow toward them under the force of gravity, and the size of the lake and its elevation relative to other bodies of water in the area can determine the relative size of the groundwater system feeding it; that is, very large and low-elevation lake surfaces can be sinks for regional systems, whereas small ponds at relatively high elevations will generally serve only local systems.

Second, lakes, especially those with large surface areas, can allow significant evaporation to take discharge water out of the system and put it back into the atmosphere. This process is especially important in hot summer months when actual evapotranspiration can equal potential evapotranspiration. For example, Leap (1988) discusses an area on the Coteau Des Prairies in northeastern South Dakota where 6% of the surface area is covered with glacially-derived ponds and lakes, ranging from a few hundred

meters to a few kilometers in diameter, that drive local and intermediate systems respectively. The potential evapotranspiration rate from May through September, when transpiration is highest, was calculated to be 0.61 m.

Assuming, for example, that in one square kilometer of the area, 6% of the area, or $60,000 \text{ m}^2$ is covered with water, then the total amount lost from the lakes is approximated as $60,000 \text{ m}^2 \times 0.61 \text{ m}$ or 36,600 cubic meters of water during the period from the first of May through the last of September. This is a conservative estimate because minimal evapotranspiration occurs before and after this time before the winter freeze sets in.

For illustrative purposes, assume that the one square kilometer area is underlain by an aquifer five meters thick with a porosity of 25% (characteristic of this area). If fully saturated, this aquifer would then contain $1,250,000 \text{ m}^3$ of water. Thus, the $36,600 \text{ m}^3$ of water lost by evapotranspiration would equal 3% of the total water contained in the aquifer.

This amount does not seem like much over the entire one-square-kilometer area, but the lowering of the water surface of many ponds by even a small amount by evapotranspiration sets up gradients sufficient to drive many local flow systems. The same amount of loss in one large lake of the same area can cause gradients sufficient to drive even intermediate systems.

Meyboom (1967) discusses the role of evapotranspiration in prairie pothole lakes in western Canada. His study showed that the role of vegetation, especially water-using phreatophytes like willow, around the edge of a lake can be quite significant. In many areas, the water table will actually be depressed beneath these trees.

The depression of the water table below the root zone can be studied quantitatively with a method by White (1932). Using his equation, one can estimate the amount of water removed from the groundwater system per day by evapotranspiration. An example is given in Todd (1959).

Thus, it is the interaction of the dynamics of gravity pulling rain and snow to the surface and then moving groundwater to the lakes (and streams), and evapotranspiration removing it from the earth's surface that maintains the circulation of groundwater. If this were not so the residence time in the ground would be so long that groundwater would dissolve so much mineral matter that it would become unpotable. Through this interactive process, water on earth is constantly cleansed before being re-introduced into the system.

Definitive modeling studies of the three-dimensional physical dynamics of groundwater flow to and from lakes were made by Winter (1976). Two contrasting examples of several modeled scenarios are shown in Figure 2.29a and b in order to illustrate the importance of permeability contrasts in the subsurface below lakes.

In Figure 2.29a, the subsurface is anisotropic, but homogeneous. Figure 2.29b shows the case where the subsurface is anisotropic and heterogeneous. Given that there is an ample supply of recharge, a high-permeability layer at depth can be great enough to divert much of the flow toward the lower layer. Thus, high-permeability layers at depth are significant in causing water from the surface to recharge the subsurface. Knowledge of the positions and hydraulic characteristics of such layers are critical to estimating recharge rates for water-management purposes, and for estimating the potential for contaminants from the surface to reach potable groundwater supplies.

Ideally, there will be a point beneath a lake undergoing both recharge and discharge, to and from the subsurface respectively, where the hydraulic gradients will be in opposition to the extent that flow does not take place. This is called the *stagnation point* and it is shown in Figure 2.29a. The stagnation point is in dynamic equilibrium with the head and gradients that exist at any one time, but will change position with changes in recharge rates from the surface, or discharge from the subsurface. Therefore, any attempts to dispose of waste at a stagnation point, as has been suggested in the past, would not be wise due to the fact that the location will change under most circumstances and soluble waste will then migrate.

An additional factor that can greatly affect the inflow/outflow characteristics of lakes is the presence low-permeability sediments on the bottom. This is often a problem in artificial recharge pits where such sediments can lower the infiltration rate of water from the pits to the subsurface. The influence of the

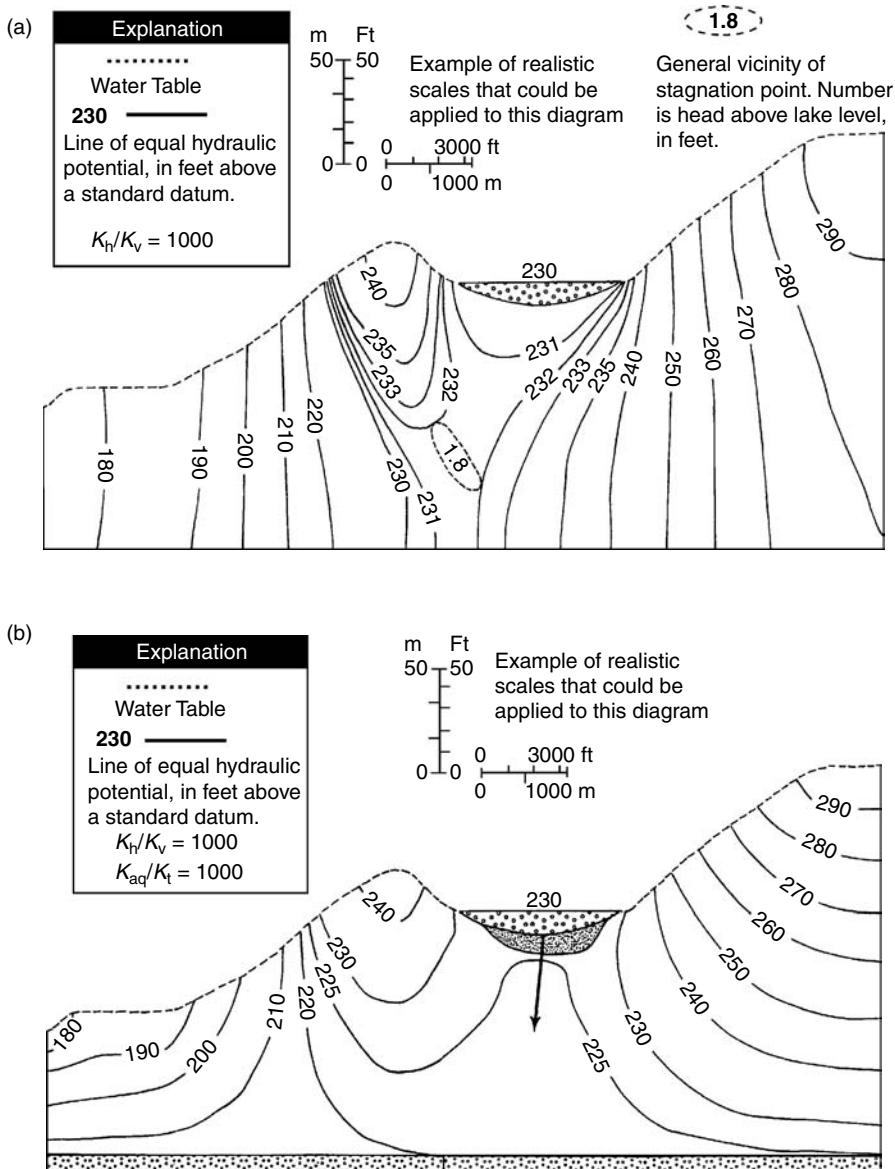


FIGURE 2.29 Cross section showing lake/aquifer interactions. (a) Anisotropic and homogeneous aquifer; (b) Anisotropic and heterogeneous aquifer. (Adapted from Winter, T.C. 1976. *Numerical Simulation Analysis of the Interaction of Lakes and Ground Water*. U.S. Geological Survey Prof. Paper 1001. U.S. Government Printing Office, Washington, DC.)

sediments can be quantified by Equation 2.42, from which it can be shown that even a thin layer of low-permeability sediments can make a very large difference in the infiltration rate.

2.6.3 Gaining and Losing Streams

Streams also gain and lose water in the same manner as lakes and ponds (Figure 2.30a and b). *Perennial streams*, that is, streams that flow year-round, are usually *gaining streams*. They gain water from base flow through the sides of the streams as it flows through from the groundwater system. This kind of stream is situated such that the water table is always above the stream surface.

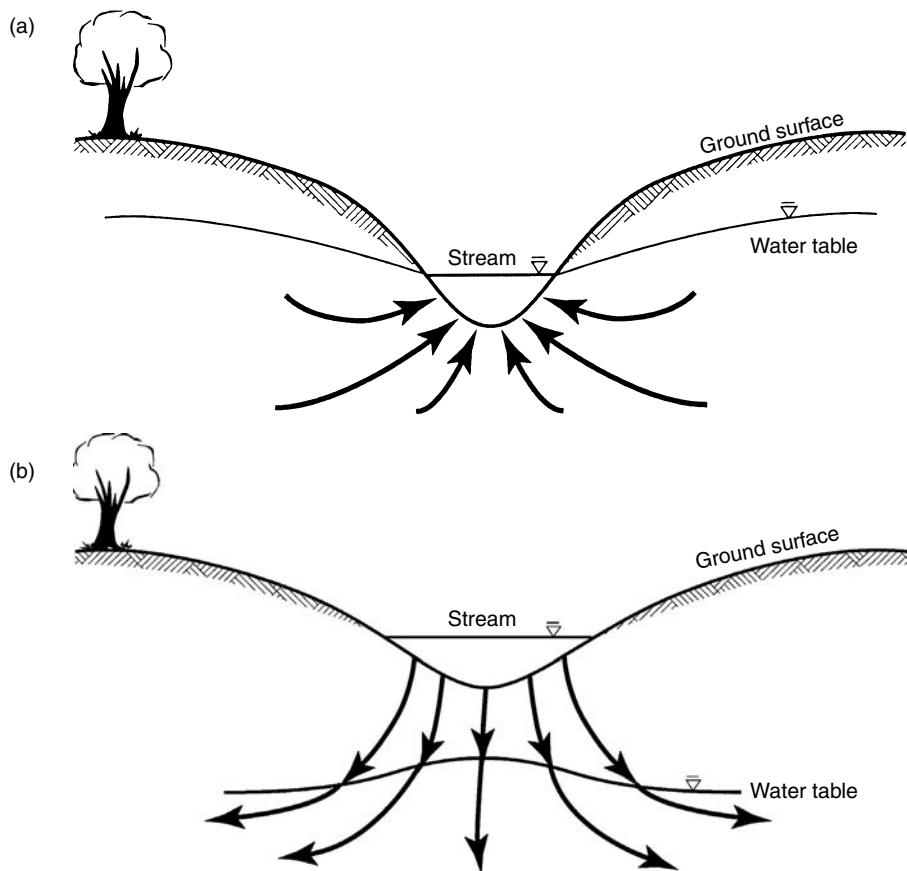


FIGURE 2.30 Gaining and losing streams. (a) Gaining stream; (b) losing stream.

Gaining streams in the cross section (Figure 2.30a) will usually show a stagnation point in the subsurface similar to that of lakes. However, in three dimensions, it is in reality a stagnation line that runs under the stream along its length.

Most streams are gaining streams; losing streams (Figure 2.30b) are generally those *intermittent streams* which flow only after significant precipitation and during runoff periods. They lose water to the subsurface because the water table is below the stream surface. Therefore, losing streams are generally found in mountainous areas on alluvial fans debouching onto pediments, on sand and gravel surfaces where the water table is low, or on steep slopes. In nearly all cases, the precipitation rate is insufficient and/or the subsurface is too permeable to maintain a water table high enough to support gaining streams.

Streams in certain geological settings may contain stretches which are gaining for some distance where the water table is above the stream surface, and then become losing stretches when they flow over a low-water-table area; or vice-versa. These streams are common in karst terraines where a stream may be losing in its upstream reaches, and may even disappear beneath the ground surface into a cave; downstream, it may later emerge to flow on the surface as a gaining stream. In glaciated areas of the Midwestern United States, streams may flow over till with high water tables as gaining streams, but after crossing a contact between till and outwash sand and gravel, the streams may become losing streams as their water infiltrates downward into the permeable outwash.

In areas where there are heavy demands for water for irrigation, industry, and large cities, damming of streams and pumping of aquifers has dramatically altered natural stream/groundwater interactions. Damming often raises the water table upstream of the dam, causing some losing streams to become

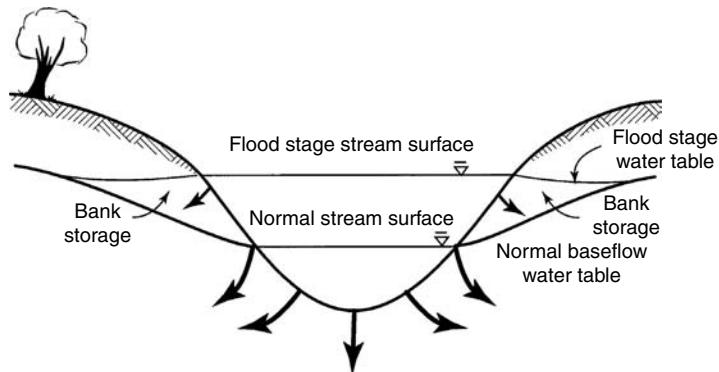


FIGURE 2.31 Bank storage.

gaining streams; and lowers the water table downstream, causing gaining streams to become losing. Heavy pumping has caused some small gaining streams to become losing because of drastic lowering of the water table. When a gaining stream does become losing, it can introduce surface contaminants into the groundwater system.

2.6.4 Bank Storage

Figure 2.30a shows a gaining stream during normal baseflow when the water table slopes toward the stream, and Figure 2.31 shows the same stream after a sudden rise in the surface elevation following an intense precipitation event. In the latter case, the water table near the stream has risen above the normal water-surface elevation and now slopes away from the stream to force water to move from the stream into the banks and beyond.

After the stream stage falls back to its normal position, the water table near the stream may not fall as fast and may remain unusually higher than the stream surface for several days or hours, depending on the permeability of the ground and the amount of water that was forced into the banks. This excess water in the stream banks is known as *bank storage*, and is common along rivers that flood.

In bottom lands along alluvial stretches, bank storage can slow natural drainage back into the stream, and thus, can cause farm fields along the stream to remain soggy for extended periods. If the stream is contaminated, the contaminants can be forced into the surrounding soil and they may not be flushed out for some time, depending on the soil permeability, the amount of precipitation, and the frequency of flooding.

2.7 Protection of Groundwater Supplies

With ever-increasing demands on our groundwater supplies, more and more thought is being given to their protection, but water demands of growing populations, especially in large cities, are often in conflict with established and traditional uses such as farming, ranching, and industrial uses. In the eastern United States where the precipitation may approach 180 cm/year, the problem is not as great as in the west where cities have been built in arid areas with precipitation of perhaps 15 cm/year. In these areas, not only has water from rivers in distant states been diverted (e.g., the Colorado River), but also, extensive pumpage of groundwater has been taking place that withdraws water faster than it can be replenished by inflow or natural recharge; this is known as *groundwater mining*.

The concept of *safe yield* was evidently suggested early in this century (Lee, 1915). The definition put forward by Todd (1959) is "... the basin draft on a groundwater supply which can be continued indefinitely without harming the supply or basin landowners." This definition is basically the same as that given by most hydrologists, but the concept of "harm" has changed over the years.

Early in this century, harm meant overdraft of the aquifer and adverse effects on the supplies of adjacent landowners who tapped the same aquifer. It came to include the breaking of water laws eventually. In the last twenty years, it has included adverse effects on ecosystems, especially if pumping would cause a lake or wetland to lose water and thereby endanger certain species, or if profitable and popular recreational sites and activities were adversely affected.

The added concerns complicate the “safe yield” definition beyond the rather simple concern of overdraft which can be calculated with the Equation of Hydrologic Equilibrium (Todd, 1959), which is given as

$$\begin{aligned}
 & [\text{Surface inflow} + \text{Subsurface inflow} + \text{Precipitation} + \text{Imported water} \\
 & \quad + \text{Decrease in surface storage} + \text{Decrease in groundwater storage}] \\
 = & [\text{Surface outflow} + \text{Subsurface outflow} + \text{Consumptive use} + \text{Exported water} \\
 & \quad + \text{Increase in surface storage} + \text{Increase in groundwater storage}]. \tag{2.59}
 \end{aligned}$$

Another complicating factor is that the added concerns above are subject to temporal and locational interpretations and values; what may be politically and socially acceptable definition today may change in ten years with changes in demographics and land use. An example is a recent problem in Park County, Colorado where a proposed *conjunctive use* plan by the city of Aurora, Colorado met stiff opposition from residents of the county.

Conjunctive use is defined simply as the use of both groundwater and surface water conjunctively. The city of Aurora, approximately 100 miles (150 km) east of Park County, bought water rights from several ranchers in the county some years ago when ranching started to become less profitable. The City is expanding and now has proposed to recharge aquifers with surface water during periods of higher runoff, and then to pump the water out later and use it conjunctively with surface water from the same area. It appears that the amount to be pumped would exceed the recharge and groundwater mining will occur. Although the greater population of Aurora was in favor of the plan, the much smaller population of Park County was opposed.

This situation is only one of many such conflicts that have occurred and are still occurring in the western United States. Similar plans have been put forward by Colorado Springs, Colorado and Las Vegas, Nevada. Again, the mining of vast quantities of water from remote areas would be necessary, even with conjunctive use to supply ever expanding cities with no practical limits to growth. Therefore, safe yield may be defined differently by developers and city planners than by landowners from whose area the water is to be taken.

2.7.1 Water Laws

The brief discussion of the subject in the following paragraphs can be supplemented by the information in Chapter 32, *Legal Framework for Groundwater Protection in the United States* by Veil, Elcock, Ranek, and Puder.

Numerous laws have been passed since the founding of the United States of America that regulate the actual use of water. These vary from state to state and from east to west. A good overview of these laws can be found in Domenico and Schwartz (1990), Fetter (1994), Todd (1959), and Walton (1991).

Within the past two decades, the U.S. Congress has passed several laws aimed at protecting water supplies, making them safer to drink and more available to everyone; they include the *National Environmental Policy Act of 1969* [P.L. 91–190], *Federal Water Pollution Control Act of 1972* [P.L. 92–500], *Clean Water Act Amendments of 1977* [P.L. 95–217], *Safe Drinking Water Act of 1974* [P.L. 93–523] and Amendments, *Resource Conservation and Recovery Act of 1976* [RCA — P.L. 94–580], *Comprehensive Environmental Response, Compensation and Liability Act of 1980* [CERCLA — P.L. 96–510], *Superfund Amendments and Reauthorization Act of 1986* [SARA], *Surface Mining Control and Reclamation Act* [SMCRA — P.L. 95–87], *Uranium Mill Tailings and Control Act of 1978* [UMTRCA — P.L. 95–604 and later amendments], *Toxic*

Substances Control Act [TOSCA — P.L. 94–469 and amendment], and Federal Insecticide, Fungicide and Rodenticide Act [FIFRA — P.L. 92–516 and later amendments].

To many individuals and industries, this number of recently-passed environmental laws seems onerous and restrictive. Yet, the lack of such laws earlier in the history of the United States resulted in severe contamination of both surface and groundwater from mining, industry, agriculture, and municipal and domestic sewage. For the most part, the laws are accomplishing a good deal of the protection and remediation for which they were written. For further details see the Chapter 32 on Legal Framework.

2.7.2 Natural Protective Barriers and Waste Disposal

Groundwater contamination is now a world-wide problem, especially in rapidly urbanizing areas, not only from trash and other wastes, but also from fertilization of lawns and gardens. For example, Cox et al. (1996) discuss the effects of urban growth on groundwater supplies of Australia; and Howard et al. (1996) discuss similar problems in Canada.

Much progress is being made today in the ability to map land for groundwater pollution vulnerability, and also to determine the areas that are protected by natural barriers, such as till and clay (Vrband and Zaporozec, 1994).

Today, there is a greater emphasis on utilization of natural protective barriers for groundwater protection than ever before. Impermeable or low-permeability natural barriers are normally clay (glacial till and lacustrine sediments), shale, salt deposits, and unfractured granitic rocks and basalt.

If waste, hazardous or otherwise, can be placed in proper containers and placed so that a natural barrier is between it and a groundwater supply, the chances are good that the water will be protected, providing proper engineering methods are followed. Unfortunately, such practices have not always been the case. In the past, abandoned gravel pits and rock quarries were often used as disposal places of convenience. Today, many of them leak and have seriously contaminated groundwater resources under and around them.

An example is the Tippecanoe County Landfill north of Lafayette, Indiana which was constructed approximately 35 years ago in a gravel pit in the glacial outwash sand and gravel bordering the Wabash River. A 3 m thick layer of clay on the bottom of the pit was removed to deepen the depression to allow more volume for refuse. Thus, the natural barrier was removed and leachate seeped into the aquifer below.

The landfill was used to dispose of not only domestic trash, but also industrial chemicals, including polychlorinated biphenyls (PCBs). The EPA and the Indiana Department of Environmental Management forced the landfill to close and cease operations and declared it a CERCLA (Superfund) site. Remediation was expected to cost the Principally Responsible Parties (PRPs) and perhaps the citizens of the county at least \$8-million, but it finally cost about \$18 million.

Another example in the same county, also on the outwash along the Wabash River, is the disposal of waste on the permeable surface. Sludge from municipal sewage processing combined with substrate sludge from mycin drug manufacturing has been disposed of for many years on the surface of the outwash. The outwash is perhaps 30 m thick and possesses a hydraulic conductivity of more than 200 m per day. The applied mixture is exceedingly rich in nitrate, phosphate, and potassium, which is welcomed by farmers as free fertilizer.

Since the soil is very permeable and retains little water, leachate from the sludge not only rapidly enters the soil, but also, center-pivot irrigation from wells in the outwash is necessary to maintain adequate soil moisture. As a result, the nitrate level in the groundwater below the site and downgradient from the site is high, exceeding by a factor of four or five the maximum permissible concentration (MPC) of 45 mg/l for drinking water set by the EPA. Not only is this water unhealthy to drink, but in addition, the property values of homes in the immediate area have dropped significantly, and perhaps as much as eight square kilometers of valuable and productive aquifer area is now seriously contaminated.

If the sludge were spread a few kilometers away on till 10 m thick with a hydraulic conductivity of 10^{-7} cm/sec, not only would it serve as inexpensive fertilizer, but the natural barrier properties of the till would prevent groundwater contamination.

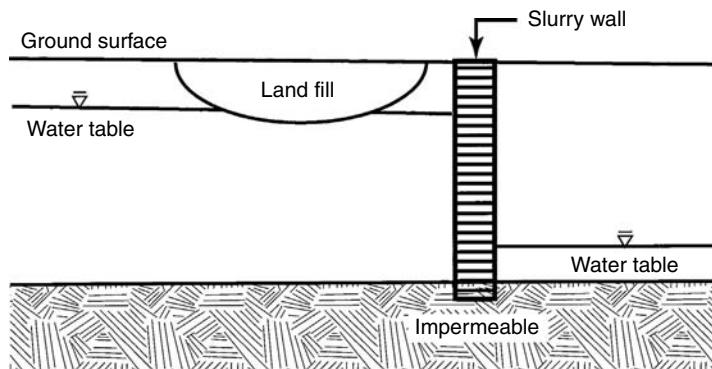


FIGURE 2.32 Slurry wall to prevent contaminant migration.

2.7.3 Artificial Protective Barriers

When natural barriers are not present, two kinds of artificial barriers can be constructed — solid and hydraulic. Extensive information about the use of such barriers can be found in EPA (1977, 1984).

Solid barriers may consist of grout curtains, slurry walls, or sheet piles constructed around a contaminated site such as a landfill to prevent the spread of contaminants to the surrounding groundwater. Figure 2.32 shows a slurry wall downgradient from a landfill which is blocking the movement of contaminants from moving downgradient at that location. This same figure can also demonstrate the utility of a sheet pile with interlocking pieces, or a grout curtain constructed at the same location.

A slurry wall is constructed by digging a trench down to an impermeable stratum and then filling it with a mixture of water, clay, and *bentonite*. The latter material is a clay mineral formed from the chemical alteration of volcanic ash and is often used as a sealant because when wet, it will expand eight to ten times its dry volume and is quite impermeable. If the depth to an impermeable stratum is too deep for construction of a slurry wall, a grout curtain can be constructed by drilling several wells close together in a line and then injecting a variety of sealing compounds into them which will infiltrate the soil pores to form an impermeable wall around the source of contamination.

A hydraulic barrier can be implemented by injecting water from a line of wells as illustrated in Figure 2.33. This kind of barrier is utilized to prevent the encroachment of saline water into fresh water aquifers along a seacoast and has been used successfully in Los Angeles, California and in Long Island, New York. It is also employed to prevent the downgradient migration of contaminants from landfills.

2.7.4 Well-Head Protection Programs

Several state environmental agencies have implemented Well-Head Protection Programs under directives from the U.S. EPA. These programs are designed to protect the recharge areas of municipal and public wells from surface contamination by limiting the kinds of activities that can take place within a certain distance from a well or well field within the *well-head protection zone*. Generally, this zone is the same as the *capture zone* for a well or well field in an unconfined aquifer. The distance is computed as the distance of travel of water to the well within a given time limit. For example, the Indiana Well-Head Protection Program, currently under revision, requires that the minimum distance limit be set for a travel time of five years, but strongly encourages a ten-year limit.

The argument for setting travel-time limits is two-fold. In the first place, an arbitrary distance limit would not make much sense without taking into account the geology, conductivity, and gradients specific to each site. Second, it is assumed that a five- or ten-year travel time will more likely guarantee that under most situations, any contaminants will have traveled far enough through the aquifer and enough water from outside the contaminated zone will have been pumped that sorption and dilution, respectively, will have rendered them innocuous.

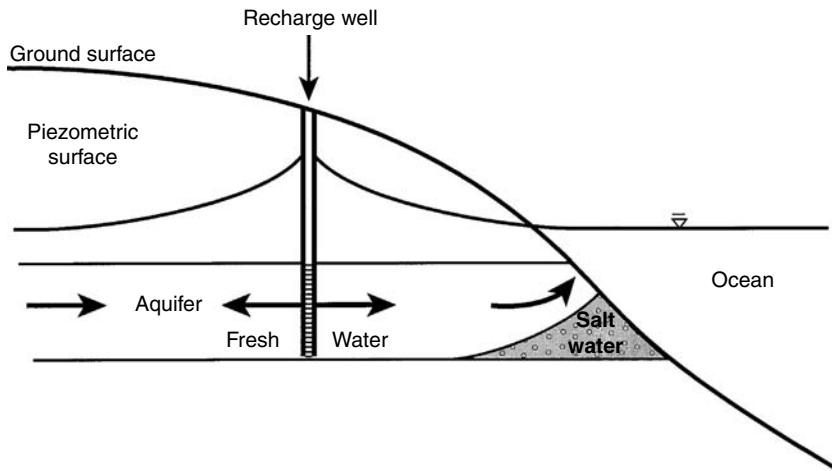


FIGURE 2.33 A hydraulic barrier.

The consequence of this law will be restriction from the recharge areas, or at least the strong regulation, of many activities that generate waste or require chemical applications. These will include livestock feedlots, fertilization of farm fields and golf courses, spreading of sludge or nitrate-rich water from sewage processing, landfills, oil refineries, chemical plants, and others.

A simple capture zone is illustrated in Figure 2.34 and is computed analytically as (Todd, 1959),

$$x = -y / \tan[2\pi Kbiy/Q] \quad (2.60)$$

where x and y are coordinate directions, Q is the pumping rate, K is the hydraulic conductivity, b is the initial saturated thickness of the unconfined aquifer, i is the hydraulic gradient, and the argument of the tangent is in radians. The units have been chosen so that the argument is dimensionless.

Bakker and Strack (1996) present a sophisticated analytical-element modeling approach to determining the capture zone under conditions of irregular and more complicated geology and boundary conditions.

Well-head protection plans have been in effect in The Netherlands for several years. The implementation of such rules in the United States has met with considerable opposition from some municipalities and well owners who realize that it will necessitate an extra expense. In order to adequately define the capture zone, there may be a need in some cases to perform test drilling and surface geophysical studies to determine the geology of the area. In addition, modeling studies may be necessary to completely understand the hydraulics of the aquifer. Small public water operations may find these studies to be financially burdensome. However, well-head protection plans, when implemented, are expected to provide a high level of protection for public water supplies as the population grows and more demands are put on groundwater resources.

For further details see the chapters on Site Remediation and Geosynthetics.

2.8 Groundwater and Modern Environmental Concerns

2.8.1 Possible Effects of Climate Change on Groundwater Flow

It is known that the world's climate has changed significantly over geologic time. Evidence from fossil plants, pollen, rodent middens, tree rings, and even clay-mineral formation has shown that precipitation and temperature have varied widely (Rasmussen et al., 1993). For example, near the end of the Pleistocene Epoch (the "Ice Age"), the climate was much wetter in the American southwest than now, and the water table in southern Nevada was 100 m higher than present (Winograd and Doty, 1980). It is not entirely clear

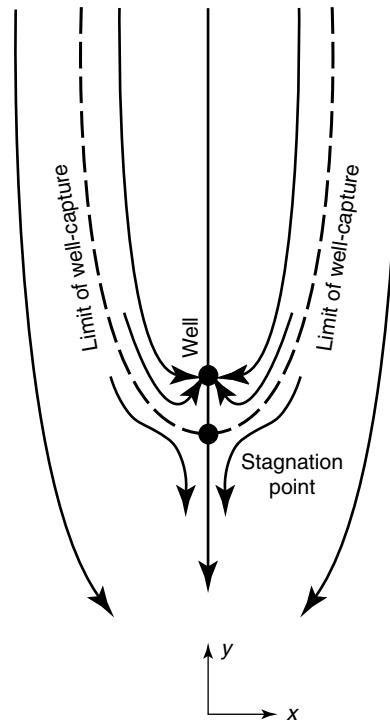


FIGURE 2.34 A capture zone around a pumping well.

what caused climate change in the past, although volcanic eruptions and tilting of the earth's rotational axis have been suggested.

Many climatologists believe that anthropogenic factors are now effecting a slow, but perceptible change in the atmosphere that is causing gradually warming temperatures. These factors include changes in land use with overgrazing and burning of forests, and the introduction into the atmosphere of abnormal amounts of certain gasses.

These gases, commonly known as *greenhouse gases*, include carbon dioxide (CO_2), and methane (CH_4) as the prime suspects in maintaining the *greenhouse effect*. This effect is the atmospheric entrapment of heat from the sun which, in turn, is expected to cause greater evapotranspiration in some areas and greater precipitation in others. As a result, the climate is expected to change significantly within the next century causing a reduction of precipitation in the Midwest and a great increase of precipitation in the southwestern United States.

Although much research has been accomplished in estimating the potential effects of climate change on surface water supplies, little attention has been directed toward the possible effects on groundwater except to estimate the increase or decrease in recharge rates and amounts that can be expected. Recent work by Leap and Belmonte (1992), Leap (1993), Reichard (1995), and Reichard and Leap (1995) illustrate the role of increased pore pressure in the opening of fractures in rocks with consequent increase in hydraulic conductivity.

It has been known for some time that a slight increase in fracture width (aperture) can cause the flow rate through a fracture to increase by a power of three; this is known as the "*Cubic Law*" (Witherspoon et al., 1980), and is given as

$$Q = (\rho g b^2 / 12\mu)(bw)(\partial h / \partial l) \quad (2.61)$$

where Q is the volumetric flow rate, ρ is the density of water, g is acceleration of gravity, b is aperture width, μ is dynamic viscosity of water, w is fracture width perpendicular to flow direction, h is the hydraulic head, and l is the length of the flow path through the fracture. The fracture is assumed to possess smooth sides and laminar flow exists. A consistent set of units must be used.

Most rocks of aquifer quality may contain tens, hundreds, or even thousands of fractures per cubic meter, and in order to quantify flow through such rocks for practical purposes, it is more convenient to assume that the fracture density is great enough that the flow system can be treated as a porous medium. Jones (1975) performed a series of experiments measuring flow through cores of fractured carbonate rocks under different pressures in order to discover the relationship between intrinsic permeability and effective stress. The rock contained enough fractures that an equivalent porous medium could be assumed and the relationship sought was determined to be

$$k/k_i = \{[\log S_e - 4.602]/[\log S_{ei} - 4.602]\}^3 \quad (2.62)$$

where k is intrinsic permeability at any pressure, k_i is initial intrinsic permeability, S_e is effective stress at any pressure, and S_{ei} is initial effective stress.

Leap (1993) postulated that these relationships between permeability, pore pressure, and effective stress could be significant in understanding and predicting changes in permeability if the water table rose during a period of increased recharge consequent to increased precipitation.

Reichard (1995) and Reichard and Leap (1995) used this equation as a basis for a modeling study to determine what the effect of a significant recharge-induced water-table rise would be on the hydraulic conductivity of an aquifer after a climate change which would cause significantly-increased recharge. Their research showed that under such conditions, the hydraulic conductivity of fractured aquifers could increase as much as 15 to 20%. More research is needed in this area, but the results thus far are significant because it suggests that not only will flow rates be expected to increase, but transport rates of contaminants could also increase as well.

Finally, changes in precipitation consequent to climate change can be expected to produce significant changes in the chemical quality of groundwater. At present, the total dissolved solids (TDS) content of groundwater generally increases from east to west across the United States as the result of two primary factors – precipitation and evapotranspiration.

In areas where precipitation amounts are large, infiltration from the surface to the subsurface is generally higher in permeable soils. The greater precipitation causes the formation of higher vertical hydraulic gradients and higher vertical groundwater flow rates than in areas of lower precipitation. As a result, increased amounts of fresher water are recharged to the groundwater system and the groundwater does not remain in place for as long as it does in areas with lower precipitation and lower infiltration rates.

If future climatic changes produce increased precipitation in certain areas, then it can be expected that groundwater quality should improve. In areas where precipitation will decrease, the TDS in groundwater should increase with a consequent decrease in quality.

2.8.2 Groundwater and Wetlands

“Wetlands consist of areas that are inundated or saturated by surface or ground water at a frequency and duration sufficient to support, and that under normal circumstances do support, a prevalence of vegetation typically adapted for life in saturated soil conditions” (USCIR, 1987).

Wetlands are important for intercepting and filtering surface water and groundwater as it discharges to the surface. Intense microbial activity and the uptake of nutrients by wetland plants significantly cleans the water of contaminants. For example, in the Atlantic Coastal Plain of South Carolina, the Congaree Bottomland Hardwood Swamp along the Congaree River, now the Congaree National Park, is a very old wetland of 22,200 acres (AAA, 2004). It was determined in 1990 that this wetland’s natural filtering and nutrient absorbing capabilities made it unnecessary to build a \$5-million waste-water treatment plant

for a nearby community (USEPA, 2005). Today, wetlands are being created artificially for treating waste water. Hammer (1989) thoroughly discusses this subject.

Wetlands are being reclaimed today, as in the Kankakee River Valley of northwestern Indiana. The main purpose of this project is for flood control by slowing runoff to main streams. Additionally, such reclamation greatly enhances aquatic species and water fowl.

In turn, water recharged through wetlands to groundwater systems generally tends to be cleaner and less contaminated than that recharged from non-wetland areas.

2.8.3 Groundwater and Land Subsidence

With increasing population and consequently, increasing numbers of domestic wells, irrigation wells and public water-supply wells, land subsidence is a serious problem. This phenomenon is primarily due to (1) lowering of the unconfined water table with resulting lowering of pore pressure and consequent soil compaction, and (2) lowering of the confined piezometric head which in turn reduces the upward pressure and buoyant force on overlying rocks and soil causing subsidence.

A thorough discussion of subsidence is given by J. Delleur in this handbook in Chapter 3, "Elementary Groundwater Flow and Transport Processes". In addition, Holzer et al. (1984) discuss many aspects of land subsidence with several examples.

Narasimhan and Goyal (1984) discuss land subsidence due to geothermal fluid withdrawal.

2.8.4 Groundwater and Soil Liquefaction

Liquefaction is the process by which normally solid and stable porous media becomes so saturated with water that it flows like a thick liquid or slurry like wet cement. The soil and water mixture is then called a thixotropic system.

Rearranging Equation 2.11 gives

$$\sigma_e = \sigma_t - P_p \quad (2.63)$$

which shows that if pore pressure (P_p) is equal to the total vertical stress (σ_t) then the effective stress (σ_e), that is, the stress exerted by the grains of the aquifer skeleton becomes zero and the soil then loses its strength. This condition is often created when there is excessive precipitation and infiltration that saturates the soil so much that pore pressure is equal to or exceeds total vertical stress and a mass of soil liquifies and flows like a slurry.

Liquefaction can also occur when the pore pressure is very high, but not equal to the vertical stress. The soil is then in a metastable condition and an earthquake can disturb the soil so much that the grains begin to move and slide over each other and the mass then flows.

Although such failures are often natural geological processes, their impact on society have historically been enormous. Homes and other buildings that have been built on filled land around bays and fiords have often been carried away and destroyed by such flows with significant loss of life and property.

Such disasters have been recorded around San Francisco Bay where there are large areas of fill deposited from dredging operations. Naturally-deposited silt and fine sand where rivers join fiords in Norway have also been the sites of such disasters. The Good Friday earthquake in the vicinity of Anchorage, Alaska in 1964 produced liquefaction of soil that destroyed many homes and caused significant loss of life due to liquefaction and flowage of soils around the edge of the bay. Coates (1990) discusses an approximate mobility index (AMI) that is given as

$$AMI = \theta_w / L_L \quad (2.64)$$

where θ_w is the saturated water content of in-place or undisturbed soil and L_L is its Atterberg liquid limit. At an AMI value of 0.85 debris flows may occur and liquification takes place at higher values.

2.8.5 Excess Evaporation of Groundwater and Consequent Salt Buildup

Although evaporation is commonly thought to be a surface phenomenon where heat and wind cause evaporation from surface water bodies such as lakes, water below the surface can also evaporate after it has been moved to the surface by capillary attraction. If the water table or the top of the tension-saturated capillary fringe are close enough to the ground surface, capillarity can attract the water to the surface and as it is evaporated, water continues to rise until the top of the capillary fringe or the water table drops to a low enough elevation that capillary attraction to the surface ceases. This will eventually happen unless precipitation and infiltration again raises the capillary fringe or the water table.

When ground water evaporates in this way, it usually leaves behind solid salts on and slightly beneath the ground surface. A classic example of this is found in the San Joaquin Valley of California where irrigation has been heavily used for many years to water many vegetable crops and orchards. Here, many acres of formerly arable land are covered with salt that causes these acres to be unsuitable for farming.

2.8.6 Artesian Pressure and Surface Effects

Artesian pressure in confined aquifers is in the vast majority of cases, natural. Water recharged at higher elevations flows to lower elevations beneath a confining layer of impermeable or semi-permeable rock or soil. This natural process can present problems for an expanding population that requires a greater number of new homes, wells, office buildings factories, schools, bridges, etc.

When excavating impermeable soils, engineers have to know beforehand the pore pressure beneath the material and this is generally accomplished by installing piezometers. If the pore pressure is too high, then the area must be dewatered before excavation can proceed. Otherwise, dangerous "blowouts" can occur because as excavation decreases the vertical confining stress, the pore pressure may cause the floor of the excavation fail and blow out. This has occurred in the past with loss of life and equipment.

Natural artesian pressure can force water of inferior quality from depth to the surface or to the areas of influence of pumping wells within which wells can pump it out. Such a scenario is not uncommon in the Appalachian Plateau or in the deeper river valleys of the Midwest where fractures in Paleozoic rocks can extend downward for several hundreds of feet. Saline water will often migrate upward to where they can be pumped out by municipal or irrigation wells.

On the other hand, high-quality artesian water is in demand as bottled drinking water. Artesian pressure in an aquifer also helps greatly to prevent or retard contamination from entering an aquifer from the surface, especially if the confining layer is leaky.

2.8.7 Groundwater Overdraft and Its Consequences

Overdraft simply means that (1) water is pumped out of aquifers faster than they are replenished (ground-water mining) or (2) pumped out at a rate that results in deterioration of the aquifer or causes undesirable effects on land (e.g., subsidence or dewatering of streams, lakes, and wetlands). Increasing population with consequent increasing groundwater withdrawals can result in overdraft and undesirable consequences.

The reader is encouraged to also read Chapter 27 entitled "Determining Groundwater Sustainability" by J. Bredehoeft for an in-depth discussion of the subject of overdraft.

2.8.8 Effects of Mining

Surface mining (strip mining) and deep mining (tunnel mining) both cause drastic changes to both surface water and groundwater. Surface mining can change the recharge process and produce vast chemical changes such as the production of acid-mine drainage that can run off to streams and infiltrate to groundwater systems.

Deep surface pit mining can create groundwater discharge areas that cause lowering of water tables around the pits. Deep tunnel mining also causes lowering of water tables above the tunnels because they constantly drain water or have to be pumped out to rid them of water. With the released water there is

also released to the environment acid from both coal and metals mining because of the conversion of iron sulfides to iron oxides and sulfuric acid. Metals mining also releases metal salts and some of them are toxic—especially heavy metals such as lead.

In the Appalachian Plateau of West Virginia, coal-mine tunnels are often dug into the sides of hills or deep valley walls and often these are below the water tables. As water from the mines flows or is pumped into the valleys and streams below, the water tables above the tunnels are drastically lowered and in some cases, wells go dry.

Keller (2002) presents a very good discussion of mining and its environmental and health consequences.

2.8.9 Effects of Artificial Drainage

Artificial drainage of soils for agriculture and/or buildings generally results in the lowering of water tables with often undesirable results. A good example is the drainage of natural wetlands. Construction of drainage ditches generally destroys their natural functions of wildlife production and flooding reduction. In certain types of soils, the lowering of water tables with consequent drying of the surface and the destruction of natural vegetation leaves the soil vulnerable to wind erosion.

Artificial drainage of sandy soils for farming can create additional problems. After the water table is lowered and the sandy soil is drained, maintaining adequate soil moisture for agriculture is difficult because of the low specific retention of the sand. In order to grow adequate crops, the soil must be irrigated heavily and this can result in drainage problems and contamination by fertilizers and pesticides applied to the crops directly or added to the irrigation water (e.g., “fertigation”). In addition, pumping enough water to properly irrigate can result in the lowering of water tables of productive aquifers.

2.8.10 Chemical Changes due to Groundwater and Surface-Water Interactions

In Section 2.6.2, surface discharge of groundwater to streams and lakes is discussed. There are also certain chemical reactions and processes that take place at the interface of groundwater and surface water. Winter et al. (1998) list the following: (1) acid-base reactions, (2) precipitation and dissolution of minerals, (3) sorption and ion exchange, (4) oxidation-reduction reactions, (5) biodegradation, and (6) dissolution and exsolution of gases.

As precipitation falls it dissolves from the atmosphere oxygen (O_2) and carbon dioxide (CO_2). The latter combines with water to form carbonic acid (H_2CO_3). As surface water is recharged into a groundwater system, it carries these two chemical species with it. Oxygen is used up to oxidize minerals such as iron and manganese and organic material in the near-surface soil. Consequently there is little if any dissolved oxygen in ground water. Carbon dioxide is increased in ground water. The carbon dioxide in turn, dissolves a variety of minerals (especially carbonates).

When ground water discharges to a stream valley there will often be precipitated at the unsaturated seepage face above the surface water level a variety of salts that attract wildlife. These areas are known as “salt licks.” Exsolution of carbon dioxide results in the deposition of calcium carbonate and other minerals. Iron and manganese are often precipitated in such situations as well. When ferrous iron is carried by groundwater to swamps iron hydroxides often form that are called bog iron ore. This material is used in certain undeveloped countries and societies as a source of iron for smelting and making tools. The reader is encouraged to examine Winter et al. (1998) for a more complete discussion.

2.8.11 Anthropogenic Changes in Recharge Areas

Increasing urbanization is causing significant changes to recharge areas. In large housing developments or cities, large area of pavement such as streets, sidewalks, playgrounds, and parking lots are increasing surface runoff that before development would be added to groundwater recharge. This runoff in turn contributes to increased flooding frequency and intensity. It is believed that severe flooding along the southern part of the Mississippi River has been exacerbated by increased runoff from pavement and roofs of growing urban centers over its entire drainage basin.

The use of pesticides, fertilizers, and sewage sludge causes problems of contamination of groundwater recharge areas. This is especially true in areas of sandy and therefore permeable soils as in Nebraska and northwestern Indiana. In areas where low-permeability materials such as clay-rich glacial till overlie aquifers, this problem is usually minimal or non-existent. Potentially contaminating activities are generally forbidden in well-head protection areas.

2.8.12 Groundwater Influence on Construction

Groundwater can present major challenges to construction. All too often, these potential effects are overlooked or ignored in the haste to construct homes and other buildings. Engineering soil tests are often not performed or performed in a cursory manner at best.

Soil tests for construction are generally concerned mostly with (1) water content and liquid and plastic limits (Atterberg tests), and (2) unconfined compressive strength which also is affected by the water content of soil.

The liquid limit of a soil is the water content at which the soil–water mixture will flow like a slurry such as wet cement. A soil with a water content lower than this will not flow, but will deform plastically.

The plastic limit is the water content of a soil above which a soil will deform plastically but not flow like a liquid. At still lower water contents, the soil will not even behave plastically, but will remain solid.

The unconfined compressive strength of a soil is the strength determined by compressive testing procedures and indicates the amount of soil deformation that can be expected for a given load of a building resting on footings.

Another very important consideration is selection of the correct kind of earth materials at first. Many disasters have occurred when houses were built on unstable soil close to a steep hill that resulted in landslides when the water content became too high. Swampy soils are not acceptable due to poor engineering properties and lack of sufficient support for buildings.

Descriptions of the engineering quality of specific soils are found in county soils reports published by the U.S. Soil Conservation Service.

2.8.13 Effects of Irrigation

In addition to an abundance of fertile soil, two other attributes help to make the United States and Canada major producers of grains, fruits, and vegetable crops. They are plentiful groundwater and modern irrigation equipment and methods. However, irrigation with groundwater is causing some undesirable side effects. Most notable among these are (1) groundwater overdraft and (2) contamination of groundwater and surface water.

Overdraft from pumping of irrigation water is very significant in Nebraska and in the High Plains Aquifer that stretches from Texas to South Dakota and includes nearly all of Nebraska, the western half of Kansas, parts of eastern Colorado, the panhandle of Oklahoma, the panhandle of Texas and parts of western Texas, and the extreme eastern part of New Mexico. The greatest area of this aquifer is composed of the Tertiary Ogallala Formation and the aquifer is therefore often referred to as the Ogallala Aquifer. Drawdowns in this aquifer as of 1980 exceeded 100 ft in parts of northwestern Texas and southwestern Kansas. In large parts of this aquifer in Nebraska, Colorado, Kansas, Oklahoma, Texas, and New Mexico, the drawdown has been as much as 50 ft (Gutentag et al., 1984).

Large drawdowns such as these not only deplete groundwater supplies, but also increase the vertical lift distance for pumps resulting in greater costs for pumping. In addition, Nebraska is the second of all the states in the number of acres irrigated and heavy irrigation with well water along the Platte River in Nebraska has reduced the depth of water in the stream (Winter et al., 1998).

Irrigation can cause contamination from over-pumping that causes deep water of inferior chemical quality to migrate into the area of influence of a pumping well. Contamination also results from “fertilization” where fertilizer is added to irrigation water. Additionally, pesticides added can cause contamination.

The additional hydraulic head at the surface from irrigation water causes contaminated water to migrate downward faster than clean water under natural conditions.

Contamination of soil and groundwater by salts can result from irrigation using groundwater. Overhead sprinkling and spraying can leave behind salts due to evaporation of some of the applied water. These salts can be introduced into groundwater systems by continued application of irrigation water.

Often, in order to prevent excess buildup of unwanted chemicals in the soil zone and in shallow aquifers, irrigation with clean water in excess of that required to support plant growth is employed to flush out the chemicals. This causes additional pumping and often depletion of groundwater (Winter et al., 1998).

Glossary

Adhesion: The attraction of water to solid grains by electrostatic forces.

Aeolian Deposits: Materials deposited from the air and moved about by wind (e.g., dune sand).

Alluvium: Any lithology or grains size of sediment deposited by flowing streams.

Anisotropy: The condition of a porous medium where the permeability is different in different directions.

Aquiclude: An impermeable geological formation or body that will prevent the flow of water through it.

Aquifer: A permeable geological formation or body that will yield water in economical amounts.

Aquitard: A semi-permeable geological formation or body that will retard the flow of water through it.

Artesian Aquifer: A confined aquifer in which the piezometric surface rises above the top of the aquifer.

Bank Storage: Water stored in a stream bank when the stream water level is higher than the adjacent water table, as during a period of high water.

Baseflow: The flow of water from the groundwater system to a surface stream or lake that maintains water in the surface body, even between precipitation events.

Blowout: A failure of overlying soil or rock due to high water pressure below that causes cratering of the soil by rapid expulsion of soil or rock from a site.

Cable-Tool or Churn Drilling: Drilling with a bit like a large chisel on the end of a cable. It is rapidly raised and lowered to chip a hole in the ground or in rock.

Carbonic Acid: H_2CO_3 that is formed when carbon dioxide (CO_2) is dissolved in water.

Capillary Fringe: A zone above the water table where the soil is saturated but under tension (pressure less than atmospheric), as opposed to the zone below the water table where water is under pressure.

Capillary Water or Pellecular Water: The water actually held in the soil or rock pores after drainage.

Capture Zone: The area around a well or well field from which water will move to the discharge point(s).

Characteristic Curves: Curves showing hydraulic conductivity and soil moisture content as functions of the soil pressure head; a set of such curves is "characteristic" of a particular soil.

Confined Aquifer: An aquifer overlain by an impermeable layer such that the piezometric head rises above the top of the aquifer.

Darcy's Law: The relationship discovered by Henri Darcy between flow rate, hydraulic conductivity, and gradient in a porous medium.

Dynamic Viscosity: The measure of the resistance of a fluid to shear.

Effective Porosity: The porosity of a rock or soil that is actually connected to provide flow through the rock or soil.

Fertigation: Irrigation with water that has soluble fertilizer added to it.

Fluid Potential: The potential energy per unit mass of water.

Gaining Stream: A stream in which the water surface is lower than the adjacent water table, thus causing water to flow from the groundwater system into the stream.

Glacial Drift: The general term for any and all earth materials deposited from the action of glacial ice or their meltwaters.

Glacial Outwash: Material deposited from meltwater from glaciers.

Glacial Spillway: Generally a river valley which served as a major outlet for sediment-laden water from melting glaciers (e.g., the Mississippi, Missouri, Big Sioux, and Wabash River valleys).

Glacial Till: A mixture of clay, silt, sand, gravel, and boulders deposited in a heterogeneous mass directly from glacial ice through melting or plastering.

Gravitational Water: In draining soil, it is the water that acutally drains out.

Greenhouse Gases: Gases that are believed to trap the escape of heat from the earth and therefore cause warming of the world's climate. The most important of these is carbon dioxide (CO_2), but methane (CH_4) is also of concern.

Groundwater: Water at or below the water table in earth materials.

Groundwater Mining: Removal of groundwater at a rate exceeding recharge.

Heterogeneous: The condition of a porous medium where the isotropic or anisotropic permeability is different in different parts of the medium.

Homogeneous: The condition of a porous medium where the isotropic or anisotropic permeability is the same in all parts of the medium.

Hydraulic Conductivity: The ease with which a fluid will flow through a porous medium. It is a function of the pore size and fluid properties of viscosity and density.

Hydraulic Head: An expression of potential energy of a fluid at a point expressed as a length. It includes the height of the point above a datum plane (normally mean sea level) plus the pressure of the column of water above the point expressed as the pressure-equivalent height of that water column.

Hydraulic-Rotary Drilling Method: A drilling method in which a bit on a long shaft, the drill stem, is rotated to drill a hole in the ground. Water, often mixed with clay or other substances, is pumped down the inside of the stem and forced out at the bit to lubricate the bit and to flush cuttings to the surface.

Hydrograph: A graph of flow rate of a stream vs. time. It also refers to a graph of the water level in a well vs. time.

Hygroscopic Coefficient: The maximum water content that can be held hygroscopically.

Hygroscopic Water: That water held in a thin film around grains. It is not available to plants.

Igneous Rock: A rock type that has solidified from the molten state (e.g., basalt, granite).

Induration: The process by which mineral grains become fused or cemented together to form a solid rock mass; as opposed to unconsolidated materials.

Infiltration Capacity: The maximum rate at which a soil will allow water to infiltrate it from the surface. It is dependent upon the initial moisture content of the soil, the vertical hydraulic gradient, and the rate of precipitation.

Interflow or Subsurface Stormflow: The flow of water between the surface permeable zone and a zone of low permeability below it.

Isotropic: The condition of a porous medium where the permeability is the same in all directions.

Juvenile Water: Water created by chemical reactions in volcanic and other geothermal activities.

Karst: A term describing an area of carbonate rocks which contains solution tunnels and caves. This is named after the "Karst" region of Yugoslavia.

Lacustrine Sediments: Generally fine-grained silts and clays deposited from still waters in lakes or ponds.

Liquid Limit: The water content of a soil at which the soil liquifies.

Liquefaction: The process by which a soil becomes so saturated with water that it can flow like a slurry.

Loess: "Rock flour," or fine silt deposited originally in wide outwash deposits, but later deposited downwind by the prevailing winds.

Losing Stream: A stream in which the water surface is above the water table, thus causing water to flow from the stream to the groundwater system below.

Metamorphic Rock: A rock type that was formed from either sedimentary rocks (e.g., limestone) or igneous rocks (e.g., granite) when they were subjected to heat and pressure to produce a totally different rock from the parent, but with similar chemical composition (e.g., marble or gneiss, respectively).

Mineral: A naturally occurring, non-organic substance, generally of crystalline form, that has a definite chemical composition or a narrow range of compositions.

Permeability: The property of a porous medium to transmit water. It is a function of pore diameter.

Phreatic Water (or soil water) That water above the capillary fringe that is under tension.

Plastic Limit: The water content of a soil at which the soil no longer behaves as a solid, but as a deformable plastic.

Pleistocene Epoch: A geological span of time beginning approximately 2,000,000 years ago during which four major glacial advances (Nebraskan, Kansan, Illinoian, and Wisconsinian) advanced over North America. The latest, the Wisconsinian, is contemporaneous with the Wurm advance of Europe.

Pore Velocity or Average Linear Velocity: The actual average velocity of flow through a porous medium. It is calculated as the specific discharge divided by the effective porosity.

Porosity: The percent ratio of void volume to total volume of a rock or soil.

Precipitates: Minerals such as calcite (CaCO_3) that are deposited in water from direct precipitation from the water.

Primary Porosity: The porosity characteristic of a rock or soil when first formed.

Pyroclastic: A rock type formed by the settling of sediment which in turn is the ash and cinders of igneous material from volcanic eruptions.

Reynold's Number: The ratio of accelerative forces to viscous forces which determines if a flow is laminar or turbulent.

Rock: A naturally-occurring stony entity made up of one or more minerals.

Safe Yield: The rate of withdrawal of groundwater from an aquifer such that no harm will be caused. The term, "harm" is relative to the economics of the area and time.

Saturation Ratio: The ratio of contained water in pores to the volume of the pores.

Secondary Porosity: The porosity that is added to a rock or soil after formation; it can include faults, fissures, fractures, etc.

Sedimentary Rock: A rock formed from the weathered products (detritus) of igneous, metamorphic or other sedimentary rocks (e.g., sandstone, siltstone, shale).

Seepage Face: The zone of wet soil just above the surface of a stream or lake which marks the atmospheric discharge of water from the capillary fringe.

Semi-Confining Aquifer: An aquifer overlain by a semi-permeable aquitard which will allow limited flow of water to pass through.

Specific Discharge: The volumetric flow rate per unit area of porous media perpendicular to the flow direction. It is sometimes called "Darcian Velocity."

Specific Retention: The ratio of the volume of water held by the soil to the total matrix volume.

Specific Storage: The volume of water that will be obtained from a unit volume of aquifer upon release of a unit value of head.

Specific Yield: The ratio of the volume of water drained from a soil to the total matrix volume.

Spring: A natural discharge point for groundwater to the surface.

Storage Coefficient: The specific storage time the saturated thickness of the aquifer.

Total Dissolved Solids (TDS): The mass of all dissolved solid material in a water sample divided by the volume of the sample.

Unconfined Aquifer: An aquifer that has no overlying confining impermeable layer.

Void Ratio: The volume of pores divided by the volume of solids in a sample of soil or rock.

Well-Head Protection Areas: Surface areas surrounding wells where groundwater recharge takes place or below which groundwater travels from recharge areas to the wells.

Wetlands: Areas that are continuously or temporarily saturated and support aquatic ecology.

Wilting Point: The moisture content of a soil at which capillary forces become greater than osmotic suction of plant roots, causing the plants to wilt.

References

- AAA, 2004. *American Automobile Association Tour Book of Georgia, North Carolina and South Carolina*, AAA Publishing, Heathrow, FL.
- Andrews, R.J., Barker, R., and Heng, L.M., 1995. The application of electrical tomography in the study of the unsaturated zone in chalk at three sites in Cambridgeshire, United Kingdom. *Hydrogeology Journal*, 3: 17–31.
- Back, W. Rosenshein, J.S., Seaber, P.R., eds., 1988. Hydrogeology, in *The Geology of North America*, Vol. O-2, The Geological Society of America, Bouldev, Co.
- Bakker, M., and Strack, O.D.L., 1996. Capture zone delineation in two-dimensional groundwater flow models. *Water Resources Research*, 32: 1309–1315.
- Bear, J., 1972. *Dynamics of Fluids in Porous Media*. American Elsevier Publishing Co., New York.
- Bennet, G.D., 1979. Regional groundwater systems analysis. *Water Spectrum*, 11: 36–42.
- Brooks, R.H. and Corey, A.T., 1964. Hydraulic properties of porous media, *Hydrology Paper 3*, Colorado State University, Fort Collins, CO.
- Boussinesq, J., 1904. Recherches théoriques sur l'écoulement des nappes d'eau infiltrées dans le sol et sur le débit des sources. *Journal de Mathématiques Pures et Appliquées*, 10: 5–78.
- Campbell, G.S., 1974. A simple method for determining unsaturated conductivity from moisture retention data. *Soil Science*, 117: 311–314.
- Chamberlin, T.C., 1885. *The Requisite and Qualifying Conditions of Artesian Wells*. U.S. Geological Survey 5th Annual Report, pp. 131–173.
- Coates, D.R., 1990. Relation of subsurface water to downslope movement and failure, in *Groundwater Geomorphology*, Special Paper 252, Higgins, C.G. and Coates, D.R. eds., Geological Society of America, Boulder, CO.
- Cox, M.E., Hillier, J., Foster, L., and Ellis, R., 1996. Effects of a rapidly urbanising environment on groundwater, Brisbane, Queensland, Australia. *Hydrogeology Journal*. 4: 30–47.
- Daniels, D.P., Fritz, S.J., and Leap, D.I., 1991. Estimating recharge rates through unsaturated glacial till by tritium tracing. *Ground Water*, 29: 26–34.
- Darcy, H., 1856. *Les Fontaines Publiques de la Ville de Dijon*. Dalmont, Paris, France.
- Darton, N.H., 1909. *Geology and Underground Waters of South Dakota*. U.S. Geological Survey Water-Supply Paper 227.
- Davis, S.N. and DeWiest, J.M., 1966. *Hydrogeology*. John Wiley & Sons, Inc., New York, NY.
- Domenico, P.A. and Schwartz, F.W., 1990. *Physical and Chemical Hydrogeology*. John Wiley & Sons, Inc., New York, NY.
- Dupuit, J., 1863. *Études théorétiques et pratiques sur le mouvement des eaux dans les canaux découverts et à travers les terrains perméables*, 2nd ed. Dunoud, Paris.
- Eggleston, J.R., Rojstaczer, S.A., and Peirce, J.J., 1996. Identification of hydraulic conductivity structure in sand and gravel aquifers: Cape Cod data set. *Water Resources Research*, 32: 1209–1222.
- EPA, 1977. *The Report to Congress; Waste Disposal Practices and their Effects on Ground Water*. U.S. Environmental Protection Agency, Washington, DC.
- EPA, 1984. *A Ground Water Protection Strategy for the Environmental Protection Agency*. U.S. Environmental Protection Agency, Washington, DC.
- Ferris, J.G., Knowles, D.B., Browne, R.H., and Stallman, R.W., 1962. *Theory of Aquifer Tests*. U.S. Geological Survey Water-Supply Paper 1536 E.
- Fetter, C.W., 1994. *Applied Hydrogeology*, 3rd ed. Macmillan College Publishing Co. Inc., New York, NY.
- Forchheimer, P., 1914. *Hydraulik*. B.G. Teubner, Leipzig.
- Freeze, R.A. and Cherry, J.A., 1979. *Groundwater*. Prentice-Hall, Inc., Englewood Cliffs, NJ.

- Gutentag, E.D., Heimes, F.J., Krothe, N.C., Luckey, R.R., and Weeks, J.B., 1984. *Geohydrology of the High Plains Aquifer in Parts of Colorado, Kansas, Nebraska, New Mexico, Oklahoma, South Dakota, Texas and Wyoming*. U.S. Geological Survey Prof. Paper 1400-B.
- Hammer, D.A., ed., 1989. *Constructed Wetlands for Wastewater Treatment*. Lewis Publishers, Inc., Chelsea, MI.
- Hantush, M.S., 1956. Analysis of data from pumping tests in leaky aquifers. *Transactions of American Geophysical Union*, 37: 702–714.
- Hazen, A., 1911. Discussion of “Dams on Soil Foundations.” *Transactions of American Society of Civil Engineers*, 73: 199.
- Heath, R.C., 1984. *Ground-Water Regions of the United States*. U.S. Geological Survey Water-Supply Paper 2242, pp. 1–78. U.S. Government Printing Office, Washington, DC.
- Heath, R.C., 1988. Hydrologic setting of regions, in *Hydrogeology, The Geology of North America*, Back, W., Rosenschein, J.S., and Seaber, P.R., eds., Vol. O-2, pp. 15–23, The Geological Society of America, Inc., Boulder, CO.
- Hem, J.D., 1985. *Study and Interpretation of the Chemical Characteristics of Natural Water*. U.S. Geological Survey Water-Supply Paper 2254, U.S. Government Printing Office, Washington, DC.
- Holzer, T.L., 1984. ed., Man-induced land subsidence, in *Reviews in Engineering Geology*, Vol. VI, The Geological Society of America, Boulder, CO.
- Horton, R.E., 1933. The role of infiltration in the hydrologic cycle. *Transactions of American Geophysical Union*, 14: 446–460.
- Howard, K.W.F., Eyles, N., and Livingstone, S., 1996. Municipal landfilling practice and its impact on groundwater resources in and around urban Toronto, Canada. *Hydrogeology Journal*, 4: 64–79.
- Hubbert, M.K., 1940. The theory of ground-water movement. *Journal of Geology*, 48: 785–944.
- Hurlbut, C.S., 1970. *Minerals and Man*. Random House Inc., New York, NY.
- Jacob, C.E., 1940. On the flow of water in an elastic artesian aquifer. *Transactions of American Geophysical Union*, 22: 574–586.
- Jones, F.O., 1975. A laboratory study of the effects of confining pressure on fracture flow and storage capacity in carbonate rocks. *Journal of Petroleum Technology*, 27: 21–27.
- Kanivetsky, R. and Rumynin, V.G., 1993. Determination of recharge rates to a glacial aquifer system in Minnesota using environmental tritium. *Hydrological Science and Technology*, Vol. 1, 62–73.
- Keller, E.A., 2002. *Introduction to Environmental Geology*, 2nd ed. Prentice-Hall, Englewood Cliffs, NJ.
- King, F.H., 1899. *Principles and Conditions of the Movement of Groundwater*. U.S. Geological Survey 19th Annual Report, part 2, pp. 59–294.
- Leap, D.I., 1988. *The Geology and Hydrology of Day County, South Dakota*. Bulletin 24, South Dakota State Geological Survey.
- Leap, D.I., 1993. Potential increases in contaminant-transport rates from increased aquifer pore pressure following increased precipitation, in *Proceedings of 2nd UAS/CIS Joint Conference on Environmental Hydrology and Hydrogeology, “Impact of Environmental and Climatic Change on Global and Regional Hydrology”*, pp. 73–91. Washington, DC, American Institute of Hydrology, St. Paul, MN.
- Leap, D.I. and Belmonte, P.M., 1992. Influence of pore pressure on apparent dispersivity of a fissured dolomitic aquifer. *Ground Water*, 30: 87–95.
- Lee, C.H., 1915. The determination of safe yield of underground reservoirs of the closed basin type. *Transactions of American Society of Civil Engineers*, 78: 148–151.
- Linsley, R.K., Jr., Kohler, M.A., and Paulhus, J.L.H., 1975. *Hydrology for Engineers*. McGraw-Hill, New York.
- Lohman, S.W., 1972. *Ground-Water Hydraulics*. U.S. Geological Survey Prof. Paper 708, U.S. Government Printing Office, Washington, DC.
- Maidment, D.R., 1993. Hydrology, in *Handbook of Hydrology*, Maidment, D.R., ed., pp. 1.1–1.15, McGraw-Hill, Inc., New York, NY.
- Marsily, de, G., 1986. *Quantitative Hydrogeology*. Academic Press, New York.
- Mavis, F.T. and Tsui, T.P., 1939. Percolation and capillary movements of water through sand prisms. *Bull. 18, University of Iowa Studies in Engineering*, Iowa City, Iowa.

- Meinzer, O.E., 1923. *The Occurrence of Groundwater in the United States with a Discussion of Principles*. U.S. Geological Survey Water Supply Paper 489.
- Meinzer, O.E., 1939. *Groundwater in the United States*. U.S. Geological Survey Water-Supply Paper 836-D, pp. 157–232.
- Mendenhall, W.C., 1905. *The Hydrology of San Bernardino Valley, California*. U.S. Geological Survey Water-Supply and Irrigation Paper 142, p. 124.
- Meyboom, P., 1967. Mass-transfer studies to determine the ground-water regime of permanent lakes in hummocky moraine of western Canada. *Journal of Hydrology*, 5: 117–142.
- Narasimhan, T.N. and Goyal, K.P., 1984. Subsidence due to geothermal fluid withdrawal, in *Man-Induced Land Subsidence*, Reviews in Engineering Geology, Vol. VI, pp. 35–66, Geological Society of America, T.L. Holzer, ed., Boulder, Co.
- Pilgrim, D.H. and Cordery, I., 1993. Flood runoff, in *Handbook of Hydrology*, Maidment, D.R., ed., pp. 9.1–9.42, McGraw-Hill, Inc., New York, NY.
- Rasmussen, E.M., Dickinson, R.E., Kutzbach, J.E., and Cleaveland, M.K., 1993. Climatology, in *Handbook of Hydrology*, Maidment, D.R., ed., pp. 2.1–2.44, McGraw-Hill, Inc., New York, NY.
- Rawls, W.J., Ahuja, L.R., Brakensiek, Shirmohammadi, A. 1993. Infiltration and soil water movement, in *Handbook of Hydrology*, Maidment, D.R., ed., pp. 5.1–5.51, McGraw-Hill, Inc., New York, NY.
- Reichard, J.S., 1995. *Modeling the Effects of Climatic Change on Groundwater Flow and Solute Transport*. Ph.D. Dissertation, Department of Earth and Atmospheric Sciences, Purdue University, West Lafayette, IN.
- Reichard, J.S. and Leap, D.I., 1998. The effects of pore pressure on the conductivity of fractured aquifers. *Ground Water*, 36: 450–456.
- Sun, R.J., ed., 1986. *Regional Aquifer Systems Analysis Program of the U.S. Geological Survey, Summary of projects, 1978–84*. U.S. Geological Survey Circular 1002, p. 264.
- Theis, C.V., 1935. The relation between the lowering of the piezometric surface and rate and duration of discharge of a well using groundwater storage. *Transactions of American Geophysical Union*, 2: 519–524.
- Thiem, G., 1906. *Hydrologische Methoden*, p. 56. Gebhardt, Leipzig.
- Todd, D.K., 1959. *Ground Water Hydrology*, John Wiley & Sons, Inc., New York, NY.
- Tolman, C.F., 1937. *Ground Water*. McGraw-Hill Book Co., Inc., New York, NY.
- Tóth, J.A., 1962. A theory of ground-water motion in small drainage basins in central Alberta, Canada. *Journal of Geophysical Research*, 67: 4375–4387.
- USCFSR, 1987. *United State Code of Federal Regulations, 1987*. U.S. Government Printing Office, Washington, DC.
- USDA, 1955. *United States Department of Agriculture Year Book, 1955*. U.S. Government Printing Office, Washington, DC.
- USEPA, 2005. *Water Quality and Hydrology*. www.epa.gov/0w0w/wetlands/wqhydrology.html
- Van Genuchten, M.Th., 1980. A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. *Soil Society of America Journal*, 44: 892–898.
- Vrband, J. and Zaporozec, A., eds., 1994. *Guidebook on Mapping Groundwater Vulnerability*. International Contributions to Hydrogeology, International Association of Hydrogeologists, p. 16.
- Walton, W.C., 1991. *Principles of Groundwater Engineering*. Lewis Publishers, Inc., Chelsea, MI.
- White, W.N., 1932. *A Method of Estimating Ground-Water Supplies Based on Discharge by Plants and Evaporation from Soil*. U.S. Geological Survey Water-Supply Paper 659, pp. 1–105, U.S. Government Printing Office, Washington, DC.
- Winograd, I.J. and Doty, G.C., 1980. *Paleohydrology of the Southern Great Basin, with Special Reference to Water Table Fluctuations Beneath the Nevada Test Site During the Late(?) Pleistocene*. U.S. Geological Survey Open File Report 89–569. U.S. Government Printing Office, Washington, DC.
- Winter, T.C., 1976. *Numerical Simulation Analysis of the Interaction of Lakes and Ground Water*. U.S. Geological Survey Prof. Paper 1001. U.S. Government Printing Office, Washington, DC.

- Winter, T.C., Harvey, J.W., Franke, O.L., and Alley, W.M., 1998. *Groundwater and Surface Water, A Single Resource*. U.S. Geological Survey Circular 1139.
- Witherspoon, P.A., Wang, J.S.Y., Iwai, K., and Gale, J.E., 1980. Validity of the cubic law for fluid flow in a deformable rock fracture. *Water Resources Research*, 16: 1016–1024.

Further Information

There are many journals and books about hydrology in general, including geology and engineering sources. However, the list below pertains mostly to hydrogeology.

Several very good college-level texts on the broad subject of hydrogeology exist, such as those listed in the *References* by Fetter, Freeze and Cherry, Todd, and Domenico and Schwartz. These books cover a wide variety of subjects and do not specialize in any one subdiscipline exclusively.

Water Supply Papers and Professional Papers by the U.S. Geological Survey are continuously produced on more specialized subjects in hydrogeology and are available from the USGS National Center in Reston, Virginia.

Similar publications are produced by the Canadian Geological Survey in Ottawa, and the Surveys of each of the Provinces.

Geotimes, published monthly by the American Geological Institute covers a wide range of geological and geo-environmental topics and news items.

Soil survey reports are excellent sources of information about agricultural and engineering properties of specific soils. They have been published for the majorities of counties in the United States and are generally available in university libraries and from state offices of the U.S. Department of Agriculture's Soil Conservation Service.

Various publications of the American Society of Civil Engineers and the American Society for Testing and Materials are also excellent sources.

Other excellent journals include:

Water Resources Research, American Geophysical Union, Washington, DC.

Groundwater, National Association of Groundwater Scientists and Engineers, Dublin, Ohio.

Hydrogeology Journal, co-sponsored by the Geological Society of America in Boulder, Colorado and the International Association of Hydrogeologists in Hanover, Germany.

Hydrological Science and Technology, American Institute of Hydrology, Minneapolis, Minnesota.

Bulletin of the Geological Society of America and *Geology*, both sponsored by the GSA, often contain papers on hydrogeology.

Special publications by the Geological Society of America and the American Geophysical Union are produced from time to time, specifically on hydrogeological subjects.

State geological surveys also produce results of studies of hydrogeology in each state.

Environmental Protection

Environmental Science and Technology

Scientific American

American Scientist

Soil Science of America Journals

Ground Water Monitoring Review

Journal of Irrigation

Engineering and Environmental Geology

National Speleological Society Bulletin

Journal of Hydrology

Soil Science Society of America Proceedings

Ground Water Monitoring and Remediation

Karst Hydrogeology

Journal, Irrigation and Drainage Division, American Society of Civil Engineers

3

Elementary Groundwater Flow and Transport Processes

3.1	Introduction.....	3-2
3.2	Pressure, Suction, Piezometric Head, and Hydraulic Gradient	3-2
3.3	The Motion of Groundwater.....	3-6
3.4	Flow Through Porous Media — Darcy's Law..... Similarity of Darcy's Law and Other Laws of Physics • Limitations of Darcy's Law • Laboratory Measurement of Hydraulic Conductivity • Field Measurement of the Hydraulic Conductivity	3-7
3.5	One-Dimensional Hydraulics.....	3-12
3.6	One-Dimensional Flow Through Porous Media — Leaky Aquifers.....	3-14
3.7	Dupuit-Forchheimer Assumptions	3-15
	Steady Flow over a Horizontal Aquiclude • Seepage from Open Channels • Recharge Basins • Steady Flow Toward a Well in Confined and Unconfined Aquifers	
3.8	Velocity Potential, Force Potential, and Flow Nets.....	3-20
3.9	Laplace's Equation	3-21
3.10	Land Subsidence	3-25
	Calculation of Subsidence • Seepage Force	
3.11	Salt Water Interfaces	3-28
3.12	Groundwater Quality	3-29
3.13	Transport Mechanisms of Dissolved Contaminants .. Advection • Diffusion • Dispersion • Sorption • Radioactive Decay and Degradation	3-32
3.14	Monitoring, Site Remediation, and Landfills	3-39
3.15	Parameter Values.....	3-39

Glossary.....	3-39
References	3-41
Further Information	3-43
Appendix 3.1 Values of the Error Function and Complementary Error Function	3-45

3.1 Introduction

In Chapter 2, porosity n was defined as the volume V_p of the pores of a rock or soil sample divided by the total volume V_t of both pores and solid material so that $n = V_p/V_t$. But porosity at a geometric point cannot be defined since a porous medium is a conglomerate of solid grains and voids. It is then necessary to introduce the concept of representative elementary volume (REV). The REV is sufficiently large to define a *space-averaged* porosity, but it is small enough that the variation from one REV to the next may be approximated by a continuous function on the scale of the measuring instruments. Thus one could take, for example, an REV of 1 cm³ for a fine sand, but it could be quite larger for a fractured rock. An alternate approach is to consider the porous medium as a realization of a random process. The porosity at a geometrical point is then an *ensemble average* of an infinite number of realizations (de Marsily, 1986; Charbeneau, 2000). Either of these constructs allows us to use the infinitesimal calculus and thus to apply the concepts of fluid mechanics.

Essentially all natural groundwater flows are three dimensional (3D). That is, the average velocity of a percolating water particle is represented by a vector that has three components. A simple example is the 3D radial flow toward a well (see Chapter 10 and Chapter 11). However, there are many situations in which the velocities are nearly coplanar or there is radial symmetry. In these cases the flow can be analyzed as two dimensional (2D) with sufficient accuracy for many engineering problems. An example is the infiltration of water into a series of long, parallel horizontal tile drains. Away from the pipe extremities, the shape of the water table is independent of the location along the pipe drain. In some cases the flow problem can be further reduced to one dimension. For example, the flow in karst conduit can be regarded as approximately one dimensional (1D). The existence of symmetry or special assumptions permits the simplification of many problems. It is important, however, to recognize the size of the errors that such simplifications can entail.

Groundwater flow variables, such as velocity and pressure, can vary in time or can be independent of time. For example, when a well is pumped, the water table drawdown increases with time. This is an unsteady or transient flow problem. If the flow variables do not change with time, the flow is steady.

This chapter is concerned with elementary 1D and 2D steady groundwater flow and transport problems. Advanced 3D flows, transient flows, transport processes, and modeling are treated in subsequent chapters. The geological setting is discussed in Chapter 2. Chapter 2 and Chapter 3 can be regarded as an introduction to groundwater hydrogeology and engineering. The following chapters deal with a number of advanced or specialized subjects in more detail.

The notions of saturated zone, confined aquifer, unconfined aquifer, water table, aquitard, aquiclude, perched aquifer, unsaturated zone or vadose zone and the physical properties of aquifers are discussed in Chapter 2. This chapter starts directly with the discussion of the hydraulics of groundwater or the motion of water below ground. The second part of this chapter deals with the transport of contaminants by groundwater.

3.2 Pressure, Suction, Piezometric Head, and Hydraulic Gradient

The water pressures are not only expressed as *gage pressures* but they are also expressed as *absolute pressures*. These pressures are related by the equation (Figure 3.1)

$$\text{Absolute pressure} = \text{Local atmospheric pressure} + \text{Gage pressure} \quad (3.1)$$

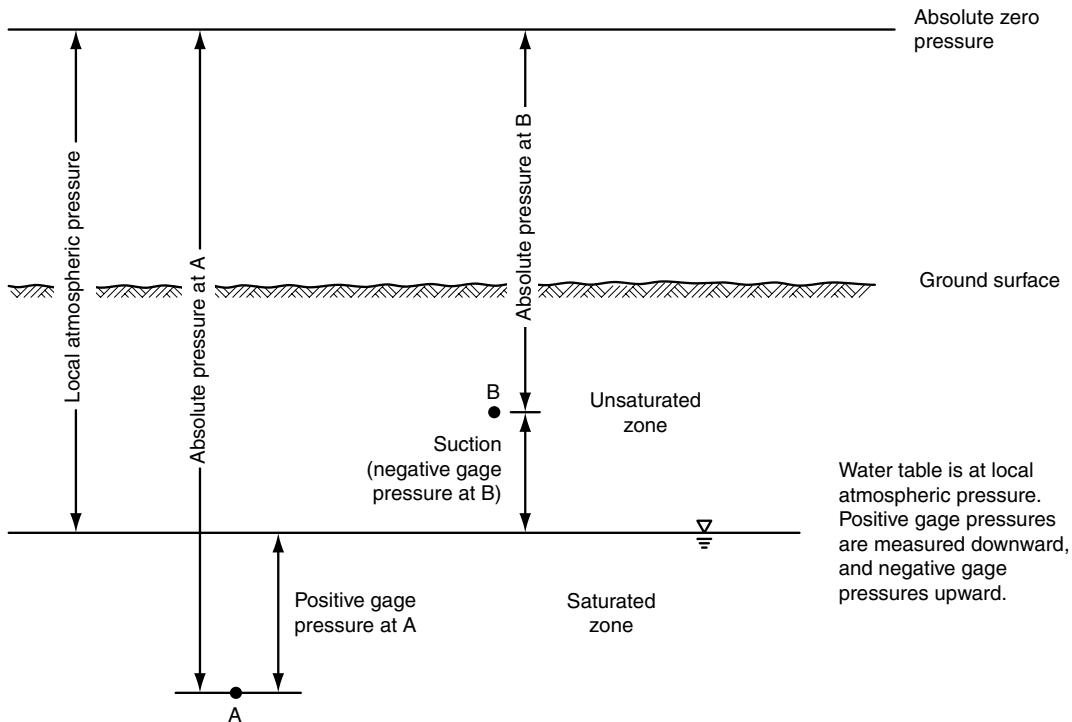


FIGURE 3.1 Absolute and gage pressures.

Gage pressures are used in the following discussion. The water table is at the local atmospheric pressure and serves as a datum for gage pressures. Point A in Figure 3.1 is in the saturated zone and the gage pressure is positive and is called the *pore pressure*. Point B is in the unsaturated zone and the gage pressure is negative. This negative pressure is referred to as a *suction* or *tension*. The suction is expressed as a positive number. Thus a positive suction corresponds to a negative gage pressure. The dimensions of pressure are F/L^2 , that is Newton per square meter or pascal (Pa), kiloNewton per square meter or kilopascal (kPa) in SI units and pound per square inch and pound per square feet in U.S. units. The law of hydrostatics states that pressure p can be expressed in terms of height of liquid h_p measured from the water table (assuming that groundwater is at rest or moving horizontally). This height is called the *pressure head*:

$$h_p = p/\gamma_w = p/\rho g \quad (3.2)$$

where $\gamma_w = \rho g$ is the specific weight, ρ is the density of water, and g is the gravitational acceleration. For point A the quantity h_p is positive whereas it is negative for point B. The pressure head is generally expressed in meters of water but it can also be expressed in centimeters of mercury. Some conversion factors for pressure and pressure heads are listed in Table 3.1 and some approximate equivalents of atmospheric pressure are listed in Table 3.2.

If the medium is saturated, pore pressure, p , can be measured by the pressure head, $h_p = p/\gamma_w$, in a *piezometer*, a nonflowing well. The difference between the altitude of the well, H (Figure 3.2), and the depth to the water inside the well is the *total head*, h_t , at the well.

In fluid mechanics the total head is defined as the sum of the *elevation head*, z , the *pressure head*, p/γ_w , and the *velocity head*, $v^2/2g$, where v is the flow velocity. For groundwater flow the velocity head can generally be ignored because the water moves very slowly. Therefore, the total head at an observation well is taken to be equal to the *piezometric head*, or the sum of the elevation head and the pressure head. The

TABLE 3.1 Conversion Factors for Pressures and Related Terms

To Convert From	To	Multiply by
Atmosphere	Pascal	1.013 E + 5
Bar	Pascal	1.000 E + 5
Dyne/cm ²	Pascal	0.100
Feet of water (39.4°F)	Pascal	2.989 E + 3
Inch of mercury (32°F)	Pascal	3.386 E + 3
Inches of water (39.4°F)	Pascal	249.1
Millibar	Pascal	100
Millimeters of mercury (0°C)	Pascal	133.3
Pound per square foot	Pascal	47.88

Source: Système International d'Unités, Universities Council on Water Resources, 1976.

TABLE 3.2 Approximate Equivalents of Atmospheric Pressure

U.S.	Metric
14.7 psi abs.	101.3 kN/m ² , abs.
	1.013 mbars, abs.
29.9 in Hg	760 mm Hg
	0.76 m Hg
33.9 ft H ₂ O	10.3 m H ₂ O

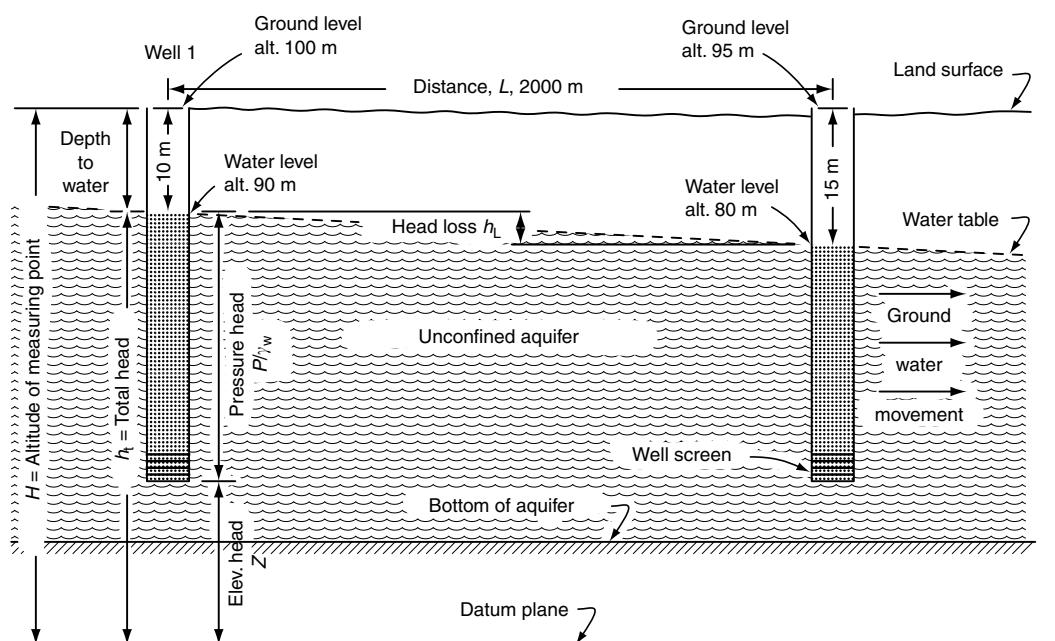


FIGURE 3.2 Heads and gradients. (From Heath, R.C. 1995. *Basic Ground-Water Hydrology*. U.S. Geological Survey, Water-Supply Paper 2200, 7th printing, Denver, CO.)

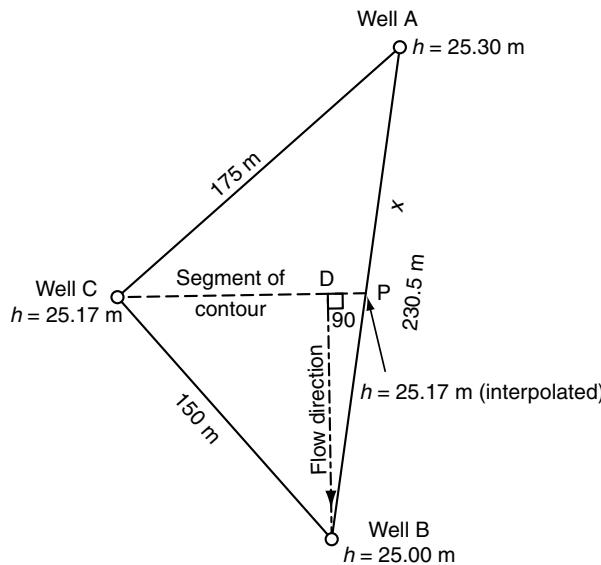


FIGURE 3.3 Finding flow direction and gradient from three observation wells. (From Heath, R.C. 1995. *Basic Ground-Water Hydrology*. U.S. Geological Survey, Water-Supply Paper 2200, 7th printing, Denver, CO.)

symbol ψ is sometimes used to designate the pressure head.

$$h_t = z + p/\gamma_w = z + \psi \quad (3.3)$$

The piezometric head is also referred to as the *piezometric potential*. The change of piezometric head per unit distance in a given direction is the *hydraulic gradient*. If the direction is not specified it is assumed to be in the direction of the maximum gradient. The hydraulic gradient is a dimensionless quantity (L/L) when consistent units are used (see also Chapter 2, Sections 2.4.7 and 2.4.11).

Example 3.1

With the data of Figure 3.2, find the average hydraulic gradient.

Solution. The average hydraulic gradient is $h_L/L = [(100 - 10) - (95 - 15)]/2000 = (90 - 80)/2000 = 0.005$.

If the piezometric head is known at three observation wells A, B, and C that are not in a straight line, then both the direction of the groundwater flow and the hydraulic gradient can be calculated approximately. For this purpose, with reference to Figure 3.3, (1) select observation well C with the intermediate head. (2) By linear interpolation, find the location of the point P having the intermediate head on the line connecting the observation wells A and B with the maximum and minimum heads. (3) Connect point P with the intermediate well, this line is a segment of a piezometric contour line. (4) Draw a line perpendicular to this contour that passes either through the maximum head or through the minimum head observation well. This line is in the direction of groundwater movement. (5) The ratio of the head differential between the ends of the perpendicular line and the length of this line is the hydraulic gradient.

Example 3.2

Using the data of Figure 3.3, find the flow direction and the hydraulic gradient.

Solution. (1) Select well C with an intermediate head of 25.17 m. (2) Find the distance x from well A to point P: $x = (25.30 - 25.17) * 230.5 / (25.30 - 25.00) = 99.88$ m. (3) Draw the line BD perpendicular to

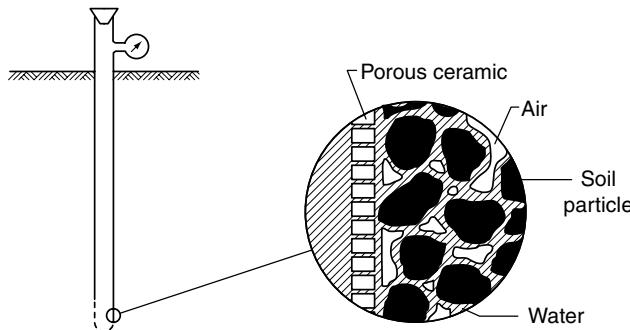


FIGURE 3.4 Tensiometer. (From Bouwer, H. 1978. *Groundwater Hydrology*. McGraw-Hill, New York. With permission.)

CP, this is the direction of groundwater movement. Measure the distance $BD = 114.29$ m. (4) Divide the head differential by the length BD to obtain the hydraulic gradient $(25.17 - 25.00)/114.29 = 0.0015$.

If the medium is unsaturated, the negative pressure or *suction* or *tension* is measured by a *tensiometer*. This instrument is composed of a vertical tube closed at the top and a porous ceramic cup at the bottom (Figure 3.4). The tensiometer is initially filled and the ceramic cup is saturated with water. When the instrument is placed in the soil, the water in the tensiometer is generally at atmospheric pressure. The soil water, which is at negative pressure, produces a suction that drains water from the tensiometer causing a pressure drop. When equilibrium is reached, the pressure inside the tensiometer is equal to that in the soil and can be measured with a vacuum gage. The practical suction range of a tensiometer is 0 to 0.8 bar or approximately 0 to 8 m of water. Relationships between the soil moisture content θ and the pressure head ψ are presented in Chapter 2 and in Chapter 6 and Chapter 7.

The *total pressure* at a point in a porous medium is the weight per unit area of the overburden above this point. This total pressure is the sum of the pore pressure and the *intergranular stress*, that is, the stress due to forces transmitted from grain to grain of the rock matrix.

3.3 The Motion of Groundwater

The motion of water requires energy. This energy can be expressed as a *head* above a *datum*. The elevation of this datum is arbitrary. This is because the difference in energy or the difference in head is the concern. It is therefore important that the energies be measured with respect to the *same* datum. In groundwater engineering the mean sea level (MSL) is usually taken as the datum. The *hydraulic head* is defined as the energy per unit weight measured relative to the datum.

Water can possess several forms of energy. Perhaps the most obvious is the energy that water possesses by virtue of its elevation above the datum. This is the *potential energy*. A mass m of water at an elevation z above the datum has a potential energy mgz , where g is the acceleration due to gravity. This is the work necessary to move the mass m from the datum to the elevation z . If ρ is the density of the water, a unit volume of water has a mass ρ and a weight ρg and a potential energy ρgz . The potential energy per unit weight, that is the *elevation head*, is thus $\rho gz/\rho g = z$. Note that the head has the unit of length.

The energy that water possesses by virtue of its motion is the *kinetic energy*. A mass m of water that moves with a velocity v has a kinetic energy $\frac{1}{2}mv^2$. Thus the kinetic energy per unit volume is $\frac{1}{2}\rho v^2$ and the kinetic energy per unit weight or *velocity head* is $\frac{1}{2}\rho v^2/\rho g = v^2/2g$. The velocity head has the dimension of length. When groundwater is flowing through the pores of the rock or soil formation, the velocity is very small, perhaps of the order of centimeters per year, and the velocity head is usually negligible with respect to the other forms of energy. One exception is near wells where the velocity increases significantly.

Another exception is in certain karst conduits where groundwater can flow fast enough that the velocity head is important.

The energy that water possesses by virtue of its pressure is the *pressure energy*. The pressure intensity of the fluid, p , acting on an area dA produces a force $p dA$. If the area is displaced by a distance ds , in the flow direction, then the force produces an amount of work $p dA ds$ known as *flow work*. The volume $dA ds$ has a weight $\rho g dA ds$ and the flow work per unit weight is $p dA ds / \rho g dA ds = p / \rho g$ known as the *pressure head*. The sum of the elevation head and the pressure head is known as the *piezometric head* $h = z + p / \rho g$.

3.4 Flow Through Porous Media — Darcy's Law

The French engineer Henry Darcy performed experiments on the filtration of water through sand columns. His finding that the rate of flow through a sand column is proportional to the loss of head appeared in an appendix to his treatise on the public fountains of the city of Dijon (Darcy, 1856). For further historic details, see Chapter 1. Figure 3.5 shows the original set up utilized by Darcy and Figure 3.6 shows some of his experimental results as plotted by Hubbert (1956) from Darcy's data. Darcy's law states that the volumetric flow rate, Q [$L^3 T^{-1}$], across a gross area A of a formation with a hydraulic conductivity K [LT^{-1}] (definition and values are given in Chapter 2), under a hydraulic gradient $i = -\partial h / \partial s$ in the

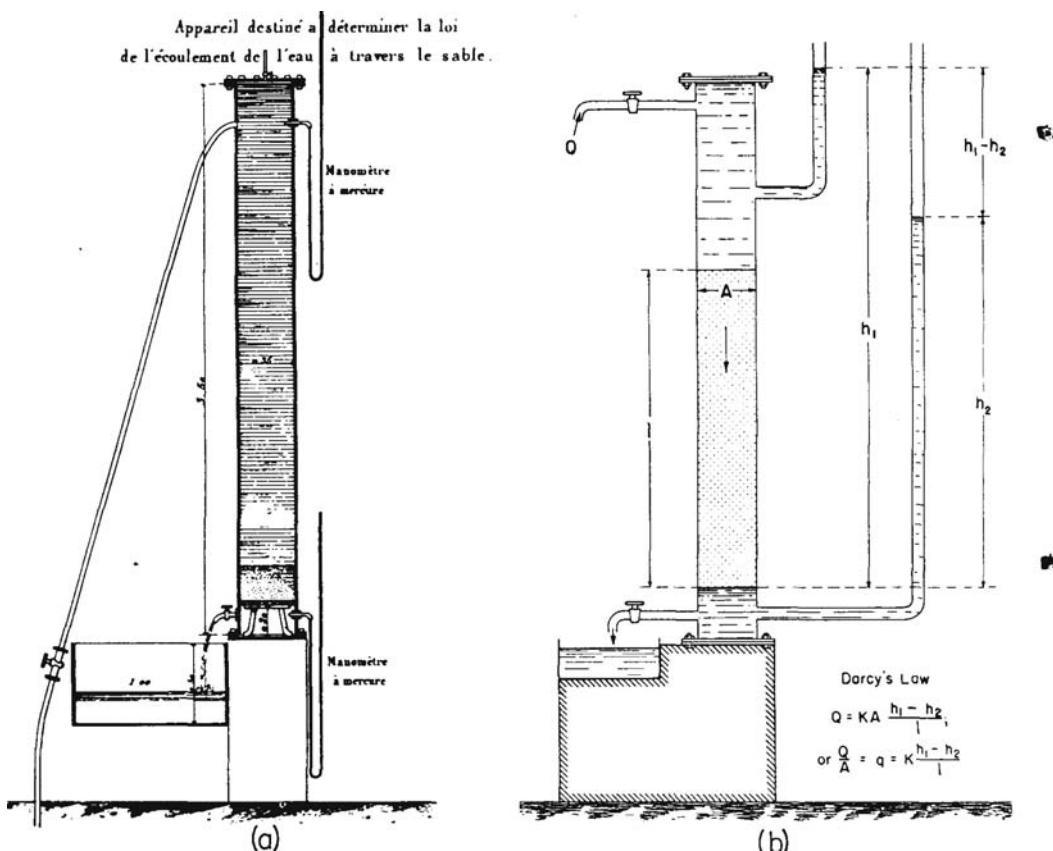


FIGURE 3.5 (a) Darcy's original apparatus with mercury manometer and (b) equivalent apparatus with water manometers. (From Hubbert, M.K. 1953. Entrapment of Petroleum under hydrodynamic conditions, *Bulletin of American Association of Petroleum Geologists*, 37, 1954–2026. With permission.)

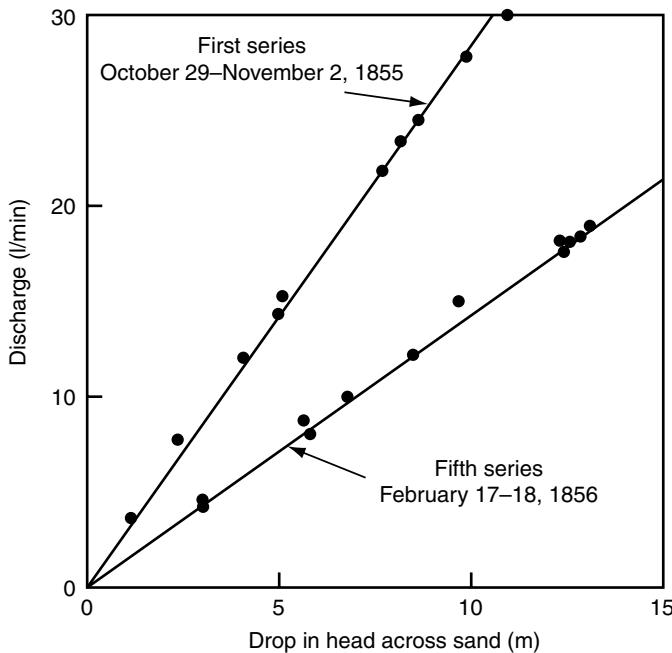


FIGURE 3.6 Darcy's data plotted by Hubbert. (From Hubbert, M.K. 1956. Darcy's law and the field equations of the flow of underground fluids. *Transactions of the American Institute of Mining, Metallurgical, and Petroleum Engineers*, 207, 229–239. With permission.)

s direction is given by

$$Q = qA = -KA\partial h/\partial s = KAi = \frac{k\rho g}{\mu} Ai \quad (3.4)$$

where q is a conceptual velocity called the *specific discharge* or flow rate per unit area [LT^{-1}] also known as the *Darcy velocity*, μ is the dynamic viscosity, and k is the *intrinsic permeability*. The hydraulic head, h , is the sum of the elevation head z and the pressure head p/γ_w . The minus sign in Equation 3.4 indicates that the flow takes place from high to low head, namely in the direction of decreasing head. The *pore velocity* is given by $v = q/n_e$, where n_e is the *effective porosity*, namely the porosity available for the fluid flow, and v is the average flow velocity in the pores, usually called the *seepage velocity*. It is the average velocity for the transport of solutes that are nonreacting. In the boundary determination of contributing areas for *well head protection zones* (see also Sect. 2.7.3) it is often necessary to find the time it takes for the water to move from a point to a bore hole. This can be done using the pore velocity. Pollutants that travel primarily by advection would move at the same velocity as the water, but those that are subject to the effects of diffusion and adsorption move more slowly. These effects are discussed in Section 3.13 and in more detail in Chapters 18 and 19, which are concerned with transport processes of nonreactive and reactive contaminants, respectively.

The one-dimensional form of Darcy's law is

$$q = K \frac{(p_1/\gamma_w + z_1) - (p_2/\gamma_w + z_2)}{L} \quad (3.5)$$

where subscripts 1 and 2 refer to the points at which the pressure heads and the elevation heads are considered, respectively, and L is the distance between these points.

Example 3.3

Find the hydraulic conductivity of the sands used in Darcy's first series experiments (Figure 3.6), assuming that the height of the sand column is 3 m and the diameter of the stand pipe is 0.35 m.

Solution. Take the flow rate $Q = 30 \text{ l/min} = 0.03 \text{ m}^3/\text{min}$. The specific discharge is $q = Q/A = 0.03/(\pi * 0.35^2/4) = 0.312 \text{ m/min}$. From Equation 3.5 $K = qL/\Delta h$. From the graph $\Delta h = 10.5 \text{ m}$. Thus $K = 0.312 * 3.0/10.5 = 0.089 \text{ m/min}$ or 0.0015 m/sec . This corresponds to a coarse sand.

Example 3.4

Find the time it takes for a molecule of water to move from a factory to a bore hole located 4 km away in a homogeneous silty sand unconfined aquifer with a hydraulic conductivity of $K = 5 \times 10^{-5} \text{ m/sec}$ or 4.32 m/d , an effective porosity of 0.4 and observing that the water table drops 12 m from the factory to the bore hole.

Solution. As a simple approximation $v = q/n_e = Ki/n_e$ and the pore velocity is calculated as

$$v = \frac{4.32 * (12/4000)}{0.4} = 0.0324 \text{ m/d}$$

It would take $4000/(0.0324 * 365) \approx 338 \text{ yr}$. If instead, the aquifer was a fractured limestone with a porosity of 0.01 and the hydraulic conductivity the same, the pore velocity would be approximately 1.3 m/d and the time to travel the 4 km would be 8.5 yr . With a porosity of 0.001 the travel time would reduce to 0.85 yr or about 10 months. Pumping at the bore hole will increase the hydraulic gradient and increase the pore velocity and thus decrease the travel time.

3.4.1 Similarity of Darcy's Law and Other Laws of Physics

Darcy's law is similar to Fourier's law of heat transfer, Ohm's law of electricity, and Fick's law of solute diffusion. *Fourier's law* governs the conduction of heat from high temperatures to low temperatures. It states that the heat flux is proportional to the temperature gradient and the constant of proportionality is the thermal conductivity. A number of flow through porous media problems have been solved using the heat conduction analogy. The Theis equation for transient flow toward wells (see Chapter 10 on "Well Hydraulics and Aquifer Tests") was obtained using an analogous problem in heat flow. *Ohm's law* can be stated as $I = V/R$, where I is the electric current, R is the resistance, and V is the voltage or potential difference across the resistor. For a cylindrical wire of length L , cross-sectional area A , and conductivity c , the resistance is $R = L/(cA)$ so that the expression for the current becomes

$$I = cA \frac{V}{L} \quad (3.6)$$

The similarity between Equation 3.4 and Equation 3.6 is now obvious. The flow rate is analogous to the current and the drop of head to the voltage drop. This analogy forms the basis for the resistance network models of aquifers (Karplus, 1958). *Fick's law* states that diffusion of a solute takes place along the concentration gradient from zones of high concentration to zones of low concentration, and the coefficient of proportionality is the diffusion coefficient. Fick's law is used in Section 3.13 and in Chapter 18 and Chapter 19, which are concerned with pollutant transport.

3.4.2 Limitations of Darcy's Law

Darcy's law implies that the flow is laminar, as is generally the case in porous media. The limit of validity can be stated in terms of the Reynolds number, N_R ,

$$N_R = qD/v \quad (3.7)$$

TABLE 3.3 Density and Viscosity of Water as a Function of Temperature

Temperature (°C)	Density ρ kg/cm ³	Viscosity $\mu \times 10^6$ Nsec/m ²	Kinematic Viscosity $v \times 10^6$ m ² /sec
0	999.8	1.781	1.785
5	1000.0	1.518	1.519
10	999.7	1.307	1.306
15	999.1	1.139	1.139
20	998.2	1.002	1.003
25	997.0	0.890	0.893
30	995.7	0.798	0.800
40	992.2	0.653	0.658
50	988.0	0.547	0.553
60	983.2	0.466	0.474
70	977.8	0.404	0.413
80	971.8	0.354	0.364
90	965.3	0.315	0.326
100	958.4	0.282	0.294

Source: Daugherty, L., Franzini, J.B., and Finnemore, E.J. 1985. *Fluid Mechanics with Engineering Applications*, 8th ed., McGraw-Hill Inc., New York. With permission.

where q is the velocity, v is the kinematic viscosity of the fluid, defined as its dynamic viscosity μ divided by its density ρ , and D is a representative length. For flow in porous media, q is taken equal to the specific discharge, and the representative length, D , is often taken equal to the pore size or the effective grain diameter, d_{10} (the grain size such that 10% of the material is larger by weight). The Reynolds number measures the importance of the inertia forces relative to the viscous forces. It depends on the viscosity that varies with temperature as shown in Table 3.3. As a result N_R also varies with temperature. Likewise the hydraulic conductivity $K = k\rho g/\mu$, where k is the intrinsic permeability of the porous medium (See Chapter 2), also varies with temperature through μ . Schneebeli (1955), using spheres of uniform diameter, found that deviations from Darcy's law start at $N_R \cong 5$ as inertia forces become effective and that turbulent flow started around $N_R \cong 60$. For flows in which the dimension D is large, such as in rocks with large fractures (see Chapter 20) or in karstic limestones (see Chapter 21), the flow can be turbulent and Darcy's law does not apply.

Darcy's law, as given above, applies to *isotropic* media, that is, where the hydraulic conductivity is independent of direction. It also applies to flows where the direction of the hydraulic conductivity corresponds to the direction of the hydraulic gradient. In *anisotropic*, media the hydraulic conductivity depends upon the direction of measurement. Then a hydraulic conductivity *tensor* is used and Darcy's law is expressed as a *tensorial equation* (see Chapter 4 on "Two- and Three-Dimensional Flow of Groundwater").

3.4.3 Laboratory Measurement of Hydraulic Conductivity

Samples of the aquifer material must be obtained and returned to the laboratory in undisturbed condition. This is generally possible for consolidated materials but usually impossible for unconsolidated material and rarely possible for fissured aquifers. The samples must be representative of the aquifer. Where the aquifer has horizontally bedded strata, the samples can be collected from bore holes that intersect the several strata. They can also be obtained from cliffs or quarry faces. Where the strata are inclined the samples likewise can be obtained from bore holes as well as from the outcrop. Bore hole samples are preferred because the outcrop material may be weathered and consequently unrepresentative of the aquifer. Consolidated aquifer samples typically are cylinders or cubes with diameter and length of 25 to 50 mm. Cubes have the advantage that they permit the determination of the permeability in three directions.

Permeameters are used for the laboratory determination of hydraulic conductivity making use of Darcy's law. It is best to use groundwater from the formation in the permeameter test because this water will be

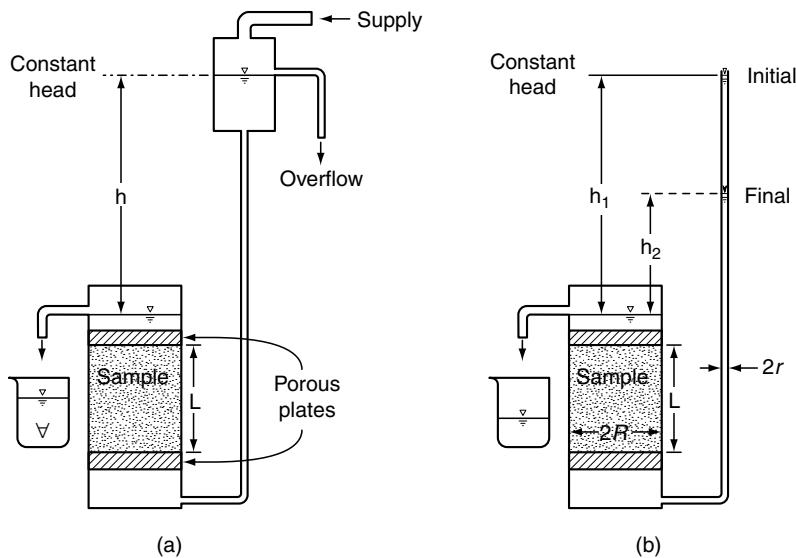


FIGURE 3.7 Permeameters. (a) Constant head and (b) falling head.

in chemical equilibrium with the aquifer material. In particular, clays can swell or shrink with changes in the water chemistry. There are two types of permeameter: the *constant head* permeameter (Figure 3.7a) used for noncohesive soils such as sands and gravels and the *falling head* permeameter (Figure 3.7b) used for materials with lower hydraulic conductivity. For the constant head permeameter, the hydraulic conductivity obtained from Darcy's law is

$$K = \frac{\forall L}{Ath} \quad (3.8)$$

where \forall is the volume of water collected in time t , A and L are the cross-section area and length of the sample, respectively, and h is the constant head. For the falling head permeameter, the hydraulic conductivity is obtained by equating the flow rate through the sample to the flow rate obtained from the observed head drop as

$$K = \frac{r^2 L}{R^2 t} \ln \frac{h_1}{h_2} \quad (3.9)$$

where R and r are the radii of the sample and of the tube, respectively, L is the length of the sample, h_1 and h_2 are the heads at the beginning and at a later time t .

3.4.4 Field Measurement of the Hydraulic Conductivity

A dependable method of field determination of hydraulic conductivity is by pumping tests. The hydraulic conductivity is obtained from observations of the water levels near pumping wells. It yields an integrated value of K (on a scale of 10 to 10^3 m) rather than the almost punctual information (on a scale of 10^{-2} to 10^{-1} m) obtained by laboratory tests. It also has the advantages that the aquifer is not disturbed and formation water is used. The pumping tests are described in detail in Chapter 10 on “Hydraulics of Wells and Aquifer Tests”.

Tracer tests using a dye such as fluorescein or a salt such as calcium chloride can also be used. If there is a drop of water table h in a distance L between the injection test hole and the observation bore hole and t is the observed travel time between the two bore holes, the hydraulic conductivity is obtained by

equating the pore velocity obtained by Darcy's law and that obtained by dividing the distance by the time. This results in

$$K = \frac{nL^2}{ht} \quad (3.10)$$

where n is the porosity of the material. In practice this test is difficult to accomplish because the flow direction must be known exactly, the distance between the bore holes should be small enough that the travel time does not become too long, and it is assumed that the aquifer is not stratified.

Other tests such as the *slug test*, the *auger-hole test*, and the *piezometer test* are based on the observation of the rate of recovery of the water level in a bore hole after the water level has been abruptly raised or lowered by addition or removal of water with a bailer or bucket or by insertion or withdrawal of a solid piece of pipe or slug into the well. These tests tend to give a more localized value of the hydraulic conductivity than the well tests but are less expensive to conduct. For more detail about these tests, the reader is referred to Bouwer (1978), Schwartz and Zhang (2003), or Butler (1997). Chapter 14 and Chapter 15 on aquifer characterization give a summary of several methods employed for the determination of aquifer properties.

3.5 One-Dimensional Hydraulics

The two basic laws of hydraulics are the continuity equation and the energy equation. They are discussed here for the case of *steady incompressible* flow. The *continuity equation* is a statement of the conservation of mass. For an incompressible fluid such as water, the equation becomes a conservation of volumes:

$$Q = A_1 V_1 = A_2 V_2 \quad (3.11)$$

where Q is the flow rate ($L^3 T^{-1}$), A_1 and A_2 are the cross-sectional areas, and V_1 and V_2 are the mean velocities at Sections 1 and 2, respectively. This assumes that there is no inflow or outflow between Sections 1 and 2. Figure 3.8 illustrates the continuity in a karst conduit.

The *conservation of energy* for steady flow is

$$z_1 + \frac{p_1}{\gamma} + \frac{V_1^2}{2g} = z_2 + \frac{p_2}{\gamma} + \frac{V_2^2}{2g} + h_L \quad (3.12)$$

where z is the elevation head or the elevation above a datum or reference plane, p/γ is the pressure head (where γ is the specific weight of the liquid), $V^2/2g$ is the velocity head, and h_L is the head loss. Each term has the dimension of length (L), and represents a form of energy per unit weight. As discussed earlier, the elevation head is the potential energy per unit weight, the pressure head is the flow work per unit weight,

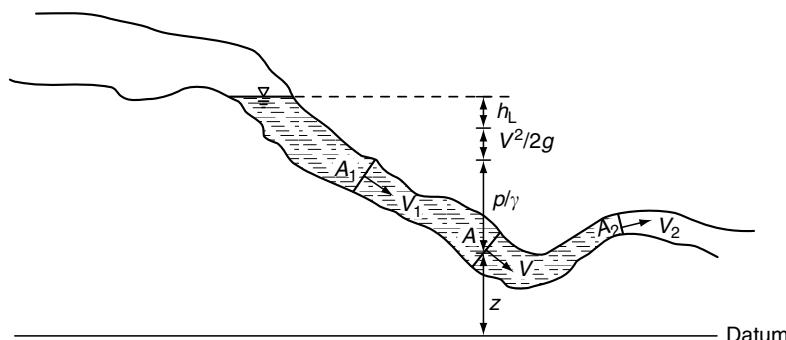


FIGURE 3.8 Continuity velocity head, pressure head, elevation head, and head loss in a karst system.

the velocity head is the kinetic energy per unit weight, and the head loss is the energy loss due to friction or other causes per unit weight. The equation can also be written as energy per unit mass

$$g z_1 + \frac{p_1}{\rho} + \frac{1}{2} V_1^2 = g z_2 + \frac{p_2}{\rho} + \frac{1}{2} V_2^2 + g h_L \quad (3.13)$$

where ρ is the fluid density. If the head loss is negligible and is made equal to zero, then the equation is known as the *Bernoulli equation*. It states that in an unbranched conduit the sum of the potential, pressure, and kinetic energies remains constant if the losses are neglected. In flow through porous media the velocity is very small and the velocity head is generally neglected but the friction losses due to viscosity, however, are important. Figure 3.8 illustrates the terms of Equation 3.12 in a karst system.

In a karst terrain, flow usually occurs in larger conduits that hydraulically act as pipes or fissures. If the Reynolds number is sufficiently small, the flow is laminar, otherwise it is turbulent. For circular pipes the *Reynolds number* is defined as

$$N_R = \frac{\rho V D}{\mu} = \frac{V D}{\nu} \quad (3.14)$$

where D is the diameter, $V = Q/A$ is the average flow velocity, that is, the flow rate Q divided by the cross-sectional area A , ρ is the density of the fluid, μ is the absolute viscosity, and ν is the kinematic viscosity. With a consistent set of units, the Reynolds number is a dimensionless quantity. The flow is laminar if the Reynolds number is less than 2000 and it is generally turbulent if the Reynolds number is larger than 4000. For values between 2000 and 4000 the flow is in a transitional regime.

For noncircular conduits such as fractures, the Reynolds number is defined as

$$N_R = \frac{\rho V (4R_h)}{\mu} \quad (3.15)$$

where $R_h = A/P$ is the *hydraulic radius* which is the ratio of the cross-sectional area, A , to the wetted perimeter, P . For a circular cross-section, $D = 4R_h$.

Friction head losses in conduits are calculated using the *Darcy–Weisbach formula* that is generally written as

$$h_L = f \frac{L}{D} \frac{V^2}{2g} \quad (3.16)$$

in which f is a dimensionless quantity called the *friction coefficient*. For laminar flow in circular conduits (also known as *Poiseuille flow*), $f = 64/N_R$ and the flow rate is thus expressed as

$$Q = \frac{\rho g D^2}{32\mu} A \frac{h_L}{L} \quad (3.17)$$

This equation is seen to be of the same form as Darcy's law (Equation 3.4) with $K = \rho g D^2 / (32\mu)$ or $k = D^2 / 32$ and $i = h_L / L$.

The Darcy–Weisbach formula (Equation 3.16) also holds for turbulent flows. Methods of calculation of the friction coefficient f for turbulent flows are described in fluid mechanics textbooks such as Finnemore and Franzini (2002).

Free surface *open channel flow* can occur in large karst cavities. Free surface streams in karst aquifers are discussed in Chapter 21.

3.6 One-Dimensional Flow through Porous Media — Leaky Aquifers

Simple 1D flows in porous media can be analyzed using Darcy's law as expressed in Equation 3.5. As an example, consider the case of a *leaky aquifer*. Many aquifers are either not fully confined or unconfined. One aquifer may be overlain by another and the stratum separating them is not fully impervious: its hydraulic conductivity is much less than that of either top or bottom aquifer but it is not zero. The rate of leakage q (volume per unit area per unit time) can be calculated using Darcy's law as

$$q = K' \frac{(h' - h)}{b} \quad (3.18)$$

where K' and b are the vertical hydraulic conductivity and the thickness of the aquitard; h' and h are either the heads at the top of the aquitard (bottom of upper aquifer) and at the top of the lower aquifer, just below the aquitard (Figure 3.9). Observe that the level in the right piezometer is slightly lower than the water table due to the downward motion of the flow. This can be verified by writing Darcy's law between point A at the water table and point B at the bottom of the unconfined aquifer: $q = K(H - h')/(y + z)$, from which $H - h' = q(y + z)/K$.

Example 3.5

As an example of a 1D flow problem, consider the case of the confined aquifer of Figure 3.9 that is recharged from an unconfined aquifer through an aquitard. The recharge rate is 0.3 m/yr or $8.22 \times 10^{-4} \text{ m/d}$. The water table is at $H = 30 \text{ m}$ above the datum. The aquitard is 2 m thick and its vertical hydraulic conductivity is $K' = 10^{-3} \text{ m/d}$. The unconfined aquifer is 20 m thick and has a hydraulic conductivity $K = 10^{-1} \text{ m/d}$. Find the piezometric head h' at the bottom of the unconfined aquifer and the difference in elevation between the water table and the piezometric surface of the confined aquifer.

Solution. Let y be the height of the piezometric surface over the top of the aquitard and z the difference in elevation between the water table and the piezometric surface (Figure 3.9). Applying Darcy's law (Equation 3.5) between points A and B yields $8.22 \times 10^{-4} = 10^{-1}(H - h')/(y + z)$. Thus $h' = H - 8.22 \times 10^{-3}(y + z) = 30 - 8.22 \times 10^{-3}(20) = 29.84 \text{ m}$. Writing Darcy's equation between the top

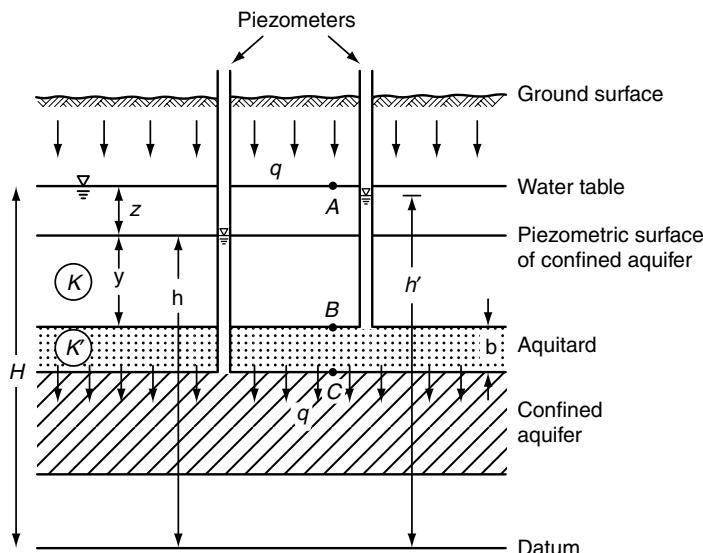


FIGURE 3.9 Leaky aquifer.

and the bottom of the aquitard yields $8.22 \times 10^{-4} = 10^{-3}(h' - h)/b$, thus $h = h' - 8.22 \times 10^{-1}b = 29.84 - 8.22 \times 10^{-1} \times 2 = 28.20$. Hence, $z = H - h = 30 - 28.20 = 1.80$ m.

When a well discharges from a leaky confined aquifer, the piezometric surface is lowered throughout a wide circular area. This lowering, called *drawdown*, is largest near the well and decreases outward. This variable drawdown changes the head differential between the confined and unconfined aquifers and alters the rate of leakage through the aquitard. From Darcy's law it follows that, at any point, the downward flow is proportional to the difference of elevation between the water table of the unconfined aquifer and the piezometric surface of the confined aquifer. Steady and unsteady flows toward wells in leaky aquifers are discussed in Chapter 10.

3.7 Dupuit–Forchheimer Assumptions

For some of the 2D flow problems, one component of the flow can be neglected with respect to the other. In particular, in some unconfined flows with a free surface, the vertical component of the flow can be neglected. This approximation pioneered by Dupuit (1863) and utilized later by Forchheimer (1930) is known as the Dupuit–Forchheimer assumption. It gives reasonable results when the depth of the unconfined flow is shallow and the slope of the free surface is small. These assumptions are summarized as follows:

1. The flow is horizontal at any vertical cross-section.
2. The velocity is constant over the depth.
3. The velocity is calculated using the slope of the free surface as the hydraulic gradient.
4. The slope of the water table is relatively small.

3.7.1 Steady Flow over a Horizontal Aquiclude

A simple application of the Dupuit–Forchheimer approximations is the analysis of steady flow through an unconfined aquifer overlying an impervious horizontal aquiclude (Figure 3.10). The discharge per unit width is $q_x = qh$, where q is the Darcy velocity or specific discharge and h is the depth of flow. From

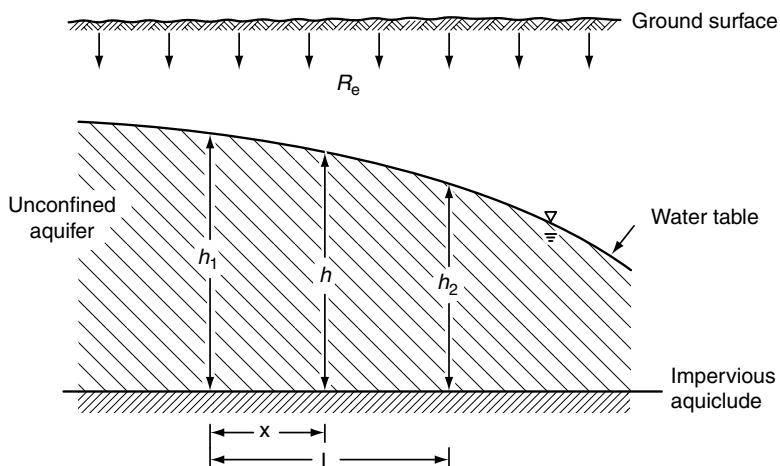


FIGURE 3.10 Steady flow through an unconfined aquifer overlying an impervious horizontal aquiclude.

Darcy's law,

$$q_x = -Kh \frac{dh}{dx} \quad (3.19)$$

Equation 3.19 is integrated from $x = 0$ (where $h = h_1$) to $x = L$ (where $h = h_2$) to obtain the *Dupuit equation*:

$$q_x = K \frac{h_1^2 - h_2^2}{2L} \quad (3.20)$$

If there is a uniform recharge with a rate R_e , then $q_x = R_e x$ with $x = 0$ at the groundwater divide and

$$R_e x = -Kh \frac{dh}{dx} \quad (3.21)$$

from which

$$K(h_1^2 - h_2^2) = R_e L^2 \quad (3.22)$$

Example 3.6

For an unconfined aquifer with a hydraulic conductivity $K = 1.75$ m/d, an effective porosity of 0.3, and water depths of 10 m and 8 m at two observation wells 200 m apart, calculate the discharge per unit width, the specific discharge, and the pore velocity.

Solution. From Equation 3.20, the discharge per unit width is $q_x = 1.75(10^2 - 8^2)/(2 * 200) = 0.1575$ m²/d. The Darcy velocity or specific discharge at the observation well with a 10 m depth is $q = q_x/h_1 = 0.1575/10 = 0.01575$ m/d. The pore velocity is $v = 0.01575/0.3 = 0.0525$ m/d. Thus, the pore water velocity changes with distance.

3.7.2 Seepage from Open Channels

The Dupuit–Forchheimer assumptions can be used to analyze the seepage from an open channel embedded in a homogeneous soil underlain by a material of much lower hydraulic conductivity, assumed impervious in the analysis (Figure 3.11). The average slope of the water table is $D_w/(L - 0.5W_s)$. The specific discharge

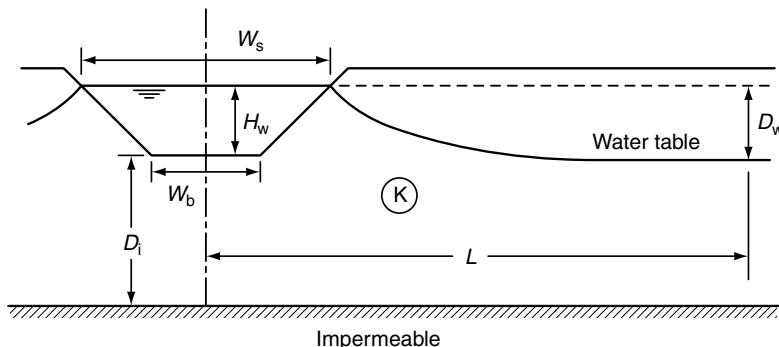


FIGURE 3.11 Seepage from a channel in a soil underlain by impervious material. (From Bouwer, H. 1978. *Groundwater Hydrology*, McGraw-Hill, New York. With permission.)

is given by Darcy's law (Equation 3.5) as $q = KD_w/(L - 0.5W_s)$. The average flow depth is $D_i + H_w - 0.5D_w$. The total seepage per unit length of channel (on both sides) is thus

$$Q = 2KD_w \frac{D_i + H_w - 0.5D_w}{L - 0.5W_s} \quad (3.23)$$

Bouwer (1969) states that this type of analysis gives reasonable results for $D_i < 3W_s$. General solutions of the canal seepage problem have been given by Bouwer (1969, 1978), Harr (1962), Polubarnova-Kochina (1962), and Yussuff *et al.* (1994). The case of partially lined channels has been examined by Subramanya *et al.* (1973). Integrated groundwater-surface water models have been developed and applied to the Imperial County in southern California by Saquib *et al.* (1995) and Taghavi *et al.* (1995).

Example 3.7

Estimate the seepage from a canal with a depth $H_w = 1$ m, dug in a soil with a hydraulic conductivity $K = 2$ m/d so that the distance from the bottom of the channel to the impervious stratum is $D_i = 10$ m, given a drop in the water table $D_w = 0.5$ m is observed at a distance $L - 0.5W_s = 400$ m from the shore. Is this seepage a significant portion of the flow?

Solution. Replacement in Equation 3.23 yields $Q = 0.05375 \text{ m}^3/\text{d}$ per meter of channel length or $53.75 \text{ m}^3/\text{d}$ per kilometer. Assume the channel had a bottom width $W_b = 4$ m, side slopes of 45° and a longitudinal slope of $S_o = 0.0004$, and a Manning roughness coefficient $n = 0.022$ (straight channel excavated in clean earth, after weathering) (Chow, 1959). The discharge capacity can be calculated by Manning's equation $Q = (1/n)AR_h^{2/3}S^{1/2}$ (see Equation 21.12, Chapter 21 on "Karst Hydrology"). The cross-sectional area is $A = 5 \text{ m}^2$, and the wet perimeter is $P = 6.828 \text{ m}$. The hydraulic radius is $R_h = 5/6.828 = 0.732 \text{ m}$. The flow rate is $Q = 5 * 0.732^{2/3} * 0.0004^{1/2}/0.022 = 3.692 \text{ m}^3/\text{sec}$ or $318,988 \text{ m}^3/\text{d}$. In a distance of 40 km, the seepage loss would amount to $40 * 53.75 * 100 / 318,988 = 0.674\%$ or less than 1%. This does not appear to be significant.

3.7.3 Recharge Basins

A third application of the Dupuit-Forchheimer assumptions is the analysis of the recharge of an unconfined aquifer from a recharge basin. A long rectangular basin of width W is considered with the assumption that the flow is horizontal and steady (Figure 3.12). Thus the depth of the aquifer, H , should not be large compared with the width, W , of the recharge basin, that is, $H \leq W$, otherwise a more detailed analysis, such as a digital model, should be used. Letting R_e be the infiltration rate from the recharge area, x the horizontal distance from the centerline of the recharge area, h the height of the groundwater mound above the static water table, and approximating the average transmissivity $T = K(H + h_c/2)$ by KH (Figure 3.12), the Dupuit-Forchheimer assumptions yield

$$R_e x = -T \frac{dh}{dx} \quad (3.24)$$

Integration of Equation 3.24 yields $h_c - h_e = R_e W^2 / (8T)$, where h_c and h_e are the height of the groundwater mound at the center and at the edge of the recharge basin, respectively. If h_c and h_e are measured in the field, then the aquifer transmissivity can be calculated.

Further discussion can be found in Bouwer (1970, 1978). An analytical solution has been given for rectangular recharging areas by Mariño (1975). Bouwer *et al.* (1999), and Bouwer (2002) developed additional equations to estimate heights of perched groundwater mounds above restricting layers in the vadose zone for narrow rectangular and circular recharge areas and gave design procedures for the design of recharge projects.

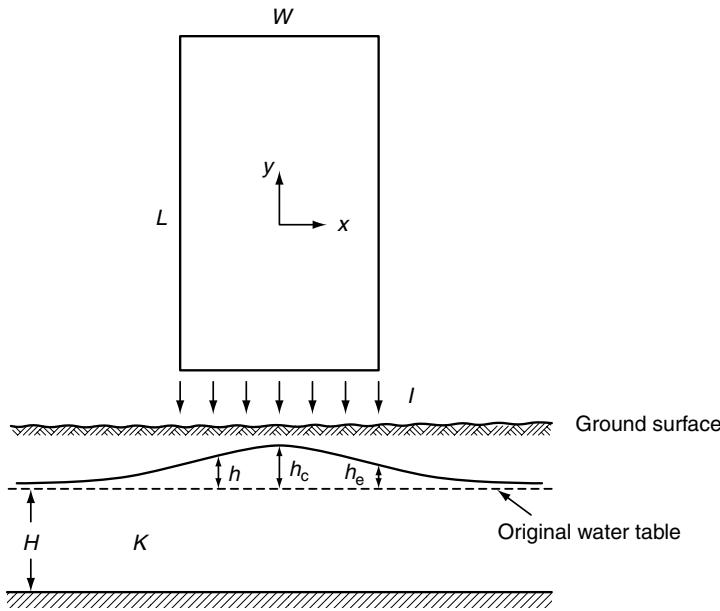


FIGURE 3.12 Seepage from a recharge basin.

Example 3.8

A long recharge basin has a width $W = 70$ m and recharge rate $R_e = 0.6$ m/d. The observed heights of the mound are $h_c = 1.2$ m and $h_e = 0.7$ m. Find the aquifer transmissivity.

Solution. From the integration of Equation 3.24, $T = 0.6(70)^2/[8(1.2 - 0.7)] = 735$ m²/d. Note that if $H > 70$ m, the transmissivity obtained is an “effective” transmissivity, as the lower layers of the aquifer are not fully contributing to the flow.

3.7.4 Steady Flow toward a Well in Confined and Unconfined Aquifers

Consider a well that fully penetrates an isotropic confined aquifer of hydraulic conductivity K . The initial piezometric level is assumed to be horizontal so that originally there is no motion of the groundwater. As water is being pumped, it flows from the aquifer toward the well lowering the piezometric surface and creating a drawdown (Figure 3.13).

For a constant pumping rate, a steady state is eventually achieved. Consider two imaginary cylinders around the well with radii r_1 and r_2 . The flows through each of these cylinders of height b are horizontal. Furthermore the flows must be equal to the discharge at the well for steady state conditions. Thus, from Darcy's law

$$Q = 2\pi r_1 b K i_1 = 2\pi r_2 b K i_2 \quad (3.25)$$

where i_1 and i_2 are the values of the hydraulic gradient at radii r_1 and r_2 , respectively. Since $2\pi b K$ is constant, and since $r_1 < r_2$, then $i_1 > i_2$. Thus the hydraulic gradient becomes steeper as the water approaches the well creating the *cone of depression*. As the same flow occurs through the two cylinders, the gross velocity, or specific discharge, increases as the well is approached. Writing Equation 3.25 as $Q = 2\pi r b K d h / dr$ and integrating yields the *Thiem equation* for confined aquifers:

$$Q = \frac{2\pi K b (h_2 - h_1)}{\ln(r_2/r_1)} = \frac{2\pi T (h_2 - h_1)}{\ln(r_2/r_1)} \quad (3.26a)$$

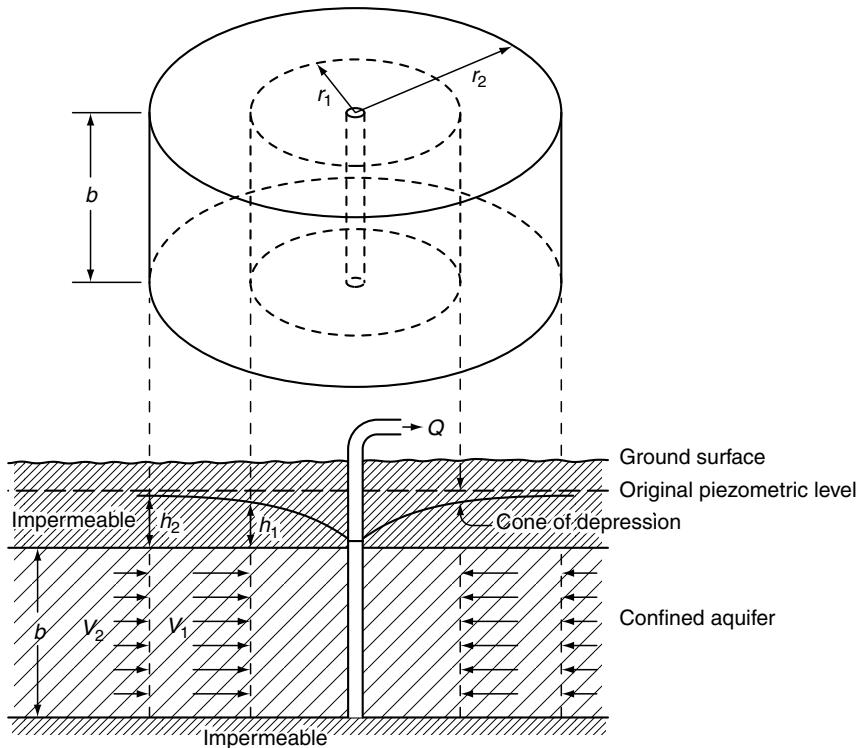


FIGURE 3.13 Flow toward a fully penetrating well in a confined aquifer.

where $T = Kb$ is the *transmissivity* of the aquifer. This solution is applicable between $r_w \leq r_1, r_2 < R$, where R is the *radius of influence*. This radius delineates the region inside of which the effect of pumping is felt and outside of which the drawdown is taken to be zero.

In the case of an *unconfined aquifer*, the saturated depth b is not constant and decreases toward the well. Perhaps the most important application of the Dupuit assumption is the computation of steady flow toward a well in an unconfined aquifer as done by Forchheimer. Assuming essentially horizontal flow, the discharge may be expressed as

$$Q = \frac{\pi K(h_2^2 - h_1^2)}{\ln(r_2/r_1)} \quad (3.26b)$$

where h_2 and h_1 are interpreted as the saturated thicknesses at locations r_1 and r_2 for $r_w \leq r_1, r_2 < R$. Equation 3.26b is known as the *Dupuit–Forchheimer formula*. Close to the well the vertical component of the flow comes into effect and the Dupuit–Forchheimer assumptions are not fully satisfied. Equation 3.26a can still be used with reasonable results for unconfined aquifers if T is interpreted as the average transmissivity $K(h_1 + h_2)/2$ and the drawdowns are small compared with the saturated thickness of the aquifer. Equation 3.26a and Equation 3.26b can also be used with one observation well and considering the pumping well as the other observation point. Both can be solved for the hydraulic conductivity K given the discharge and the heads h_1 and h_2 at radii r_1 and r_2 .

Example 3.9

A well with a radius of 50 cm completely penetrates an unconfined aquifer. It has been pumped for a long time at the rate of $15,000 \text{ m}^3/\text{d}$. The drawdown in the well is 10 m. Find the hydraulic conductivity of the

formation given that the well essentially does not affect the water table at a distance of 600 m = R , where the depth of the water table is 50 m. Well losses are neglected.

Solution. Using Equation 3.26b with $T = K(h_1 + h_2)/2$ and $h_1 = 50 - 10 = 40$ m, and solving for K gives

$$K = \frac{Q \ln(r_2/r_1)}{\pi(h_2^2 - h_1^2)} = \frac{15,000 \ln(600/0.50)}{\pi(50^2 - 40^2)} = 37.6 \text{ m/d}$$

Solving the well equation for the aquifer properties is called the *inverse problem*. The method employed in Example 3.9 using the Dupuit–Forchheimer equation only gives a rough approximation, as in practice steady state rarely exists and the transient flow formulas must be used instead. The transient flow methods (discussed in Chapter 10 on “Hydraulics of Wells and Aquifer Tests”) also yield the storage coefficient S in addition to the hydraulic conductivity K .

3.8 Velocity Potential, Force Potential, and Flow Nets

The piezometric head $h = z + p/\gamma$ is interpreted as an energy or potential per unit weight. The quantity

$$\Phi = Kh + C = K \left(z + \frac{p}{\gamma} \right) + C \quad (3.27)$$

where C is an arbitrary constant, is defined as the *velocity potential*, Φ . By virtue of Darcy’s law, the negative of its derivative in the flow direction, for constant K , is the Darcy velocity, q . As a more general term, the negative of the *gradient* of the velocity potential is a velocity. This is the definition of the velocity potential in classical hydrodynamics. (*Note:* In Chapter 5 a minus sign is introduced in the right-hand side of Equation 3.27, then the velocity is the derivative of the velocity potential.)

Hubbert (1953, 1987) introduced the concept of the *force potential* Ψ (with $C = 0$ in Equation 3.27)

$$\Psi = gh = \frac{g}{K} \Phi \quad (3.28)$$

(*Note:* In Chapter 5 the symbol Ψ is used for the “stream function,” Equation 5.16.) The force potential at a point is the work that is required to move a unit mass of fluid from a reference elevation and pressure to the elevation and pressure at the given point. In vector form, for K constant,

$$\mathbf{q} = -\nabla \Phi \quad (3.29a)$$

or

$$\mathbf{q} = -\frac{K}{g} \nabla \Psi \quad (3.29b)$$

The gradient of the force potential is the force per unit of mass acting upon the water at a given point. Formula 3.29b is a generalized Darcy’s law in three dimensions (Hubbert, 1953, 1987).

From Equation 3.29b \mathbf{q} is seen to have the same direction as $-\nabla \Psi$ as long as K is constant. Thus, the *streamlines*, which are lines everywhere tangent to the velocity vector, are perpendicular to lines of $\Psi = \text{constant}$ or *equipotential lines*. The streamlines and the equipotential lines are orthogonal. A network of streamlines and equipotential lines form the *flow net*, which is a useful tool in the analysis of 2D flows. When the hydraulic conductivity is not constant, then Equation 3.29a must be used. With some practice, good flow nets can be drawn by hand. The equipotential lines are drawn so that the drop of head Δh or potential drop $\Delta \Phi$ between adjacent lines is the same. The streamlines are drawn so that the same fraction

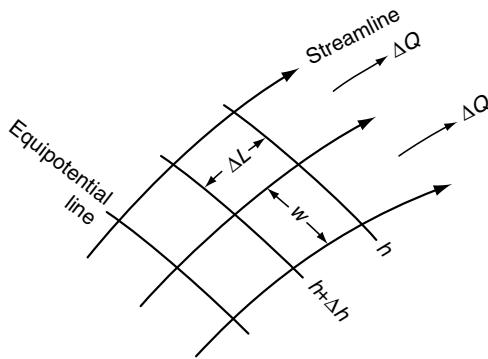


FIGURE 3.14 Streamlines and equipotential lines in a flow net.

of the total flow ΔQ takes place between adjacent streamlines. They are normal to the equipotential lines (Figure 3.14) forming “square” shapes. The flow between two streamlines, ΔQ , is obtained from Darcy’s equation as

$$\Delta Q = Kbw \frac{\Delta h}{\Delta L} \quad (3.30)$$

where b is the average depth of flow, Δh and ΔL are the difference in head and the distance between adjacent equipotential lines, respectively, and w is the width between adjacent streamlines for the square considered. The total flow through a group of n flow paths is $Q = n\Delta Q$. Figure 3.15 shows a flow net for a recharge area in an unconfined aquifer.

The mathematics, construction, and application of flow nets are discussed in Chapter 5 on Groundwater and Seepage.

3.9 Laplace’s Equation

Laplace’s equation is fundamental to the analysis of many groundwater flow problems. It arises from the combination of Darcy’s law and the equation of continuity or conservation of mass for a homogeneous isotropic aquifer. Consider a steady flow of an incompressible fluid through an elementary cube (Figure 3.16) of a porous medium of porosity n . Let u, v, w be the velocity components in the x, y, z directions, respectively. The inflow through the vertical face near the origin is $nu dy dz$. The outflow through the vertical face away from the origin is

$$\left[nu + \frac{\partial}{\partial x}(nu) dx \right] dy dz$$

and the net change of volume in the x -direction (inflow rate–outflow rate) is

$$-\frac{\partial}{\partial x}(nu) dx dy dz.$$

The sum of the net changes of volume in the x, y, z directions must be equal to zero. Thus

$$-\left[\frac{\partial}{\partial x}(nu) + \frac{\partial}{\partial y}(nv) + \frac{\partial}{\partial z}(nw) \right] = 0$$

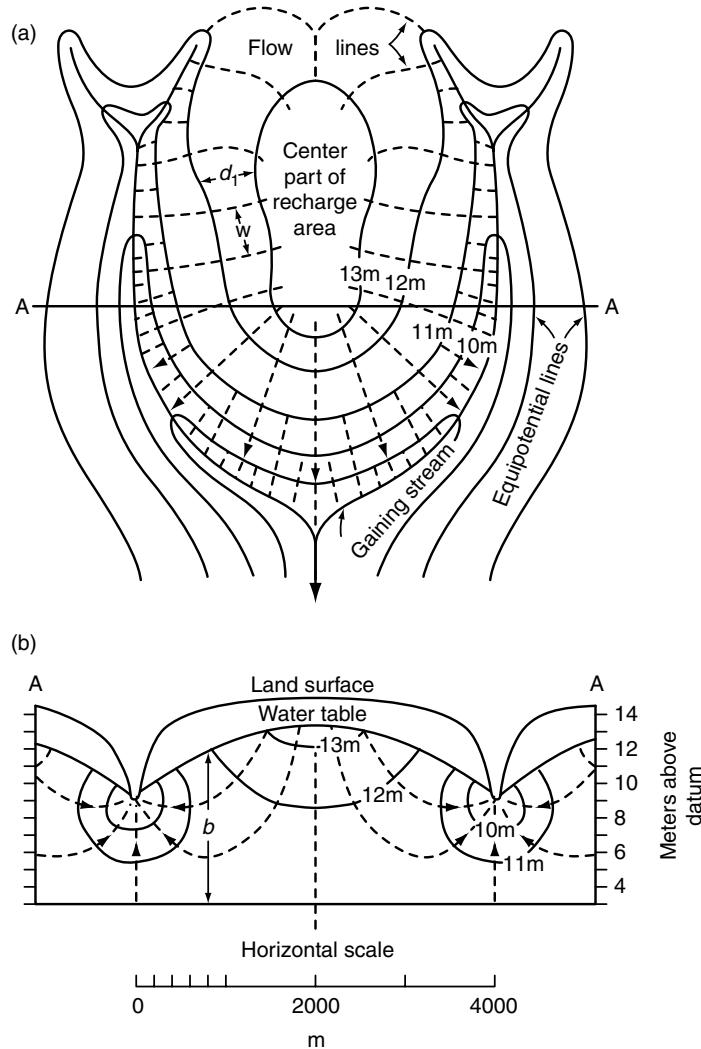


FIGURE 3.15 Flow net of a recharge area and gaining stream. (a) Plan view and (b) cross-section. (From Heath, R.C. 1995. *Basic Ground-Water Hydrology*. U.S. Geological Survey, Water-Supply Paper 2200, 7th printing, Denver, CO.)

For an incompressible isotropic homogeneous porous medium, the equation of continuity for steady incompressible flow is thus

$$\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z} = 0 \quad (3.31)$$

Using Darcy's law for an isotropic homogeneous medium

$$u = -K \frac{\partial h}{\partial x}, \quad v = -K \frac{\partial h}{\partial y}, \quad w = -K \frac{\partial h}{\partial z}$$

Laplace's equation is obtained in terms of the head h :

$$\nabla^2 h = \frac{\partial^2 h}{\partial x^2} + \frac{\partial^2 h}{\partial y^2} + \frac{\partial^2 h}{\partial z^2} = 0 \quad (3.32)$$

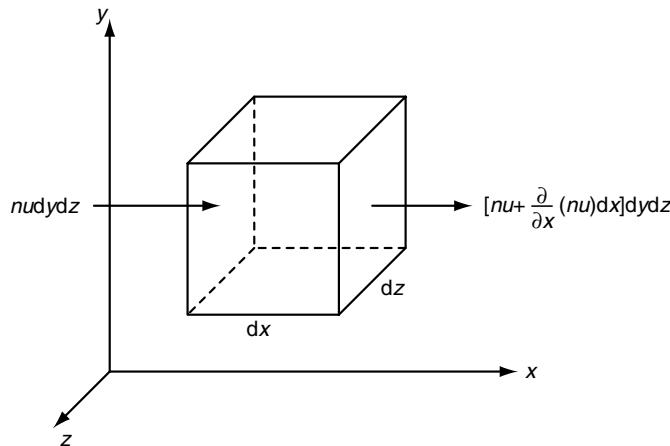


FIGURE 3.16 Elementary cube of porous media.

If in Equation 3.31 the velocity components are expressed in terms of the velocity potential (see Equation 3.29a)

$$u = -\frac{\partial \Phi}{\partial x}, \quad v = -\frac{\partial \Phi}{\partial y}, \quad w = -\frac{\partial \Phi}{\partial z}$$

then Laplace's equation is expressed in terms of the velocity potential Φ

$$\nabla^2 \Phi = \frac{\partial^2 \Phi}{\partial x^2} + \frac{\partial^2 \Phi}{\partial y^2} + \frac{\partial^2 \Phi}{\partial z^2} = 0 \quad (3.33)$$

Equation 3.33 is fundamental to the analysis of flow nets as developed in Chapter 5.

If the porous medium of a confined aquifer is assumed to be compressible with a compressibility α (the reciprocal of the modulus of elasticity,) and the water with a compressibility β , then a *specific storage coefficient* S_s is introduced as (see Equation 4.57)

$$S_s = \rho g(\alpha + n\beta) \quad (3.34)$$

S_s has the dimension L^{-1} . The compressibility α has a range of values from 10^{-6} to 10^{-8} Pa^{-1} for clays and a range of 10^{-9} to 10^{-11} Pa^{-1} for sound rock with intermediate values for sand, gravel, and jointed rock, and the compressibility of water, β , is $4.4 \times 10^{-10} \text{ Pa}^{-1}$ (Freeze and Cherry, 1979). The term $\rho g \alpha$ represents the water yield from storage due to the compression of the porous medium and $\beta \rho g n$ is the water yield resulting from the expansion of water storage. Integrating S_s over the thickness of the confined aquifer one obtains the *storativity* or *storage coefficient* S . Thus, $S \approx S_s b$, where b is the thickness of the aquifer. The storativity S is dimensionless and for confined aquifers it is of the order of 5×10^{-2} to 10^{-5} (de Marsily, 1986). For a horizontal confined aquifer, the storage coefficient is interpreted as the volume of water that the aquifer releases per unit surface area of aquifer per unit drop of the piezometric surface.

For an unconfined aquifer, S is the *drainage porosity* or *specific yield*, which is the volume of water released per unit drop of the water table per unit horizontal area. The storage coefficient can be determined in the field from pumping tests as described in Chapter 10.

Since, for the elementary cube of Figure 3.16, the difference between the inflow and the outflow is now equal to the rate of change of storage, the flow is now unsteady, thus requiring the time derivative in the

right-hand side of Equation 3.32. For 2D flow in a horizontal confined aquifer, Equation 3.32 becomes

$$\frac{\partial^2 h}{\partial x^2} + \frac{\partial^2 h}{\partial y^2} = \frac{S}{T} \frac{\partial h}{\partial t} \quad (3.35)$$

where $S (\approx S_s b)$ is the aquifer storativity, $T = Kb$ is the *transmissivity* of the aquifer, and b is the depth for a confined aquifer and is approximated as the average depth for an unconfined aquifer. If there is a leakage or inflow rate q in the aquifer per unit area, then Equation 3.35 becomes

$$\frac{\partial^2 h}{\partial x^2} + \frac{\partial^2 h}{\partial y^2} + \frac{q}{T} = \frac{S}{T} \frac{\partial h}{\partial t} \quad (3.36)$$

The leakage rate can be calculated by Darcy's law (see Equation 3.18).

The *Dupuit–Forchheimer* assumption that was applied to 1D unconfined flow (see Equation 3.19 and Equation 3.20) can be generalized to the 2D case. In this case the inflows and outflows into a small parallelepiped of unconfined aquifer can be calculated by Darcy's law and their sum must be equal to zero for a steady incompressible flow. This yields

$$\frac{\partial^2(h^2)}{\partial x^2} + \frac{\partial^2(h^2)}{\partial y^2} = 0 \quad (3.37)$$

It is seen that for the Dupuit–Forchheimer assumptions, the square of the head must satisfy Laplace's equation (Harr, 1962). If there is recharge at the rate R_e , then Equation 3.37 is replaced by (Fetter, 2004)

$$-K \left(\frac{\partial^2(h^2)}{\partial x^2} + \frac{\partial^2(h^2)}{\partial y^2} \right) = 2R_e \quad (3.38)$$

Example 3.10

A river and a canal run parallel to each other $L = 500$ m apart (Figure 3.17). They fully penetrate an unconfined aquifer with a hydraulic conductivity of 0.3 m/d. The elevation of the water surface in the river is 1.25 m lower than in the canal where the depth is 5 m. Assuming no recharge, find the water table elevation midway between the river and the canal and find the discharge into the river.

Solution. The 1D form of Equation 3.37 is $d^2(h^2)/dx^2 = 0$, which integrates as $h^2 = c_1 x + c_2$. Applying the boundary conditions at $x = 0$, $h = h_1$ and $x = L$, $h = h_2$, one obtains $h^2 = h_1^2 - (h_1^2 - h_2^2)x/L$. Thus, with $h_1 = 5$ m, $h_2 = 3.75$ m, and $L = 500$ m, $h = [5^2 - (5^2 - 3.75^2) * 0.5]^{1/2} = 4.42$ m. The seepage into the river is given by $q = -Kh(dh/dx) = K(h_1^2 - h_2^2)/2L = 0.3(5^2 - 3.75^2)/1000 = 0.0033$ m³/(m d) or 3.3 m³/(km d).

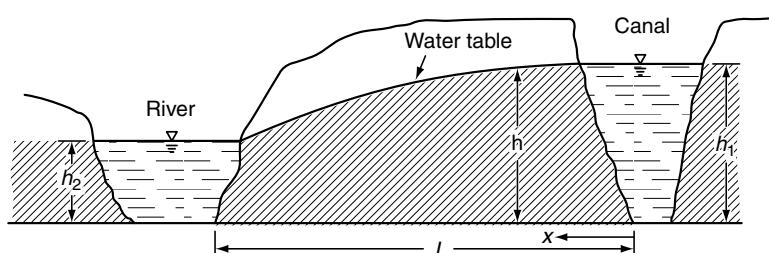


FIGURE 3.17 Example 3.10.

3.10 Land Subsidence

Pumping large volumes of water from confined aquifers at rates substantially larger than the natural recharge causes a contraction of the aquifer that can result into a downward movement of the land surface. In unconfined aquifers, groundwater pumping causes a downward movement of the water table that likewise can lead to a downward movement of the land surface. This downward movement of the land surface is called *subsidence* or *consolidation*. This movement can be a few centimeters to several meters. If the subsidence is not uniform, the differential settlement can produce severe damage to structures. Important subsidence has occurred in the San Joaquin Valley in California, in Mexico City, in Venice, around Shanghai, and in southern Taiwan.

Galloway et al. (1999) give a detailed description of subsidence in the United States where 17,000 square miles in 45 states have been affected. Man-induced subsidence is caused by three basic mechanisms (1) compaction of aquifer systems due to groundwater mining, (2) dewatering of organic soils, and (3) mass wasting through dissolution and collapse of rocks that are relatively soluble in water such as limestone, dolomite, and evaporites. Examples of these mechanisms are (1) the subsidence near Mendota in the San Joaquin Valley in California that in places, exceed 28 ft, (2) the extensive land subsidence caused by drainage and oxidation of peat soils in the Everglades in Florida, and (3) the dissolution and erosion of carbonate rocks (limestone and dolomite) in Missouri and solution of evaporites (salt and gypsum) in the Permian basin of Texas, New Mexico, Oklahoma, and Kansas.

Large subsidence that results from excessive pumping tends to occur in aquitards containing interbedded layers of silt and clay. As the drainage of the aquitard squeezes fluid from its interior, the internal stress increases. This causes the compressibility to increase by a factor of 20 to 100 resulting in compaction that is mostly not recoverable. Decrease of the groundwater pumpage can reduce and sometimes reverse the subsidence as is currently the case in the Harris-Galveston area in south central Texas (Hibbs, 1997). UNESCO has published a guidebook to studies of land subsidence due to groundwater withdrawals (Poland, 1984). The next section describes the calculation of subsidence by this mechanism.

Dewatering of organic soils accelerates the aerobic decomposition of the plant litter as compared to the anaerobic decomposition that occurs under undrained conditions. The natural rate of accumulation of organic soil is of the order of a few inches per century, whereas the rate of loss of drained organic soil can be 100 times greater.

Limestone dissolution and groundwater flow result in a geomorphology known as karst. It includes sinkholes, springs, caves, subsurface drainage networks. The frequency of occurrence of sinkholes can be accelerated by groundwater and land resources development as is the case in West Central Florida. Groundwater flow in karst aquifers is discussed in Chapter 21.

3.10.1 Calculation of Subsidence

Consider a unit area of a horizontal plane at a depth Z below the ground surface. The total downward pressure P_t due to the weight of the overburden on the plane is resisted partly by the upward hydrostatic pressure P_h and partly by the *intergranular pressure* P_i exerted between the grains of the material: $P_t = P_h + P_i$ or $P_i = P_t - P_h$.

A lowering of the water table results in a decrease of the hydrostatic pressure and a corresponding increase of the intergranular pressure. If P_{i1} and P_{i2} denote the intergranular pressures before and after a drop in the water table or piezometric surface, the vertical subsidence can be calculated as

$$S_u = Z \frac{P_{i2} - P_{i1}}{E} \quad (3.39)$$

where Z is the thickness of the soil layer and E is the modulus of elasticity of the soil. Typical ranges of values of E are given in Table 3.4. The reciprocal of the modulus of elasticity is the *compressibility*, generally designated by the symbol α and given in inverse Pascal or square meter per Newton. In general, the modulus of elasticity increases nonlinearly with the intergranular pressure. If there are layers of

TABLE 3.4 Modulus of Elasticity of Soils and Rocks

Material	Modulus of Elasticity N/cm ²
Peat	10–50
Clay, plastic	50–390
Clay, stiff	390–790
Clay, medium hard	790–1470
Sand, loose	980–1960
Sand, dense	4910–7850
Gravel, sandy gravel, dense	9806–19,620
Rock, fissured, jointed	14,710–294,200
Rock, sound	294,200–∞

Source: Jumikis, A.R. 1984. *Soil Mechanics*. Robert E. Krieger Co., Malabar, FL. Reproduced with permission.

different soil types, the subsidences are calculated separately for each layer and added to obtain the total subsidence. As the modulus of elasticity of clayey materials is much less than that of sand or gravel, most of the settlement occurs in the clayey layers.

The previous equation can also be used to calculate the rebound when the intergranular pressure decreases. Caution must be exercised because the modulus of elasticity is not the same for decompression as for compression. This is particularly the case for clays. For Boston blue clay, the rebound modulus of elasticity is only about 50% of that for compression (Bouwer, 1978, p. 323). If subsidence has occurred for a long time, complete rebound is unlikely.

Example 3.11

Consider a 60 m thick sand layer. The water table is located at a depth of 10 m below the ground surface. Calculate the total and the intergranular pressures at 10 m depth and at the bottom of the sand layer, given that the porosity of the sand is $n = 0.35$, its volumetric water content above the water table is $\theta = 0.08$, the specific weight of the solids is $\gamma_s = 25.5 \text{ kN/m}^3$, and the specific weight of the water is $\gamma = 9.81 \text{ kN/m}^3$.

Solution. At the water table the intergranular pressure, which is also the total pressure, is $P_t = 10[(1 - 0.35)25.5 + 0.08 * 9.81] = 173.6 \text{ kPa}$. The total pressure at the bottom of the sand layer is $P_t = 173.6 + 50[(1 - 0.35)25.5 + 0.35 * 9.81] = 1174.0 \text{ kPa}$. The hydrostatic pressure at the bottom of the sand layer is $P_h = 9.81 * 50 = 490.5 \text{ kPa}$. The intergranular pressure is thus $1174.0 - 490.5 = 683.5 \text{ kPa}$.

Example 3.12

If in Example 3.11 the water table drops 40 m, what is the change in intergranular pressure at the bottom of the sand layer? See Figure 3.18.

Solution. The depth to the water table would then be 50 m. The total pressure at the bottom of the sand layer is $P_t = 50[(1 - 0.35)25.5 + 0.08 * 9.81] + 10[(1 - 0.35)25.5 + 0.35 * 9.81] = 1068.1 \text{ kPa}$. The hydrostatic pressure is $P_h = 9.81 * 10 = 98.1 \text{ kPa}$. The intergranular pressure is $1068.1 - 98.1 = 970.0 \text{ kPa}$. The increase in intergranular pressure due to the 40 m drop in the water table is $970.0 - 683.5 = 286.5 \text{ kPa}$.

Example 3.13

Calculate the subsidence for the situation depicted in the previous problem if the modulus of elasticity of the sand is $10,000 \text{ N/cm}^2$ or 10^5 kN/m^2 .

Solution. The drop in the water table produces a linear increase in the intergranular pressure varying from 0.0 at the 10 m depth to 286.5 kPa at the 50 m depth. The average increment in intergranular pressure

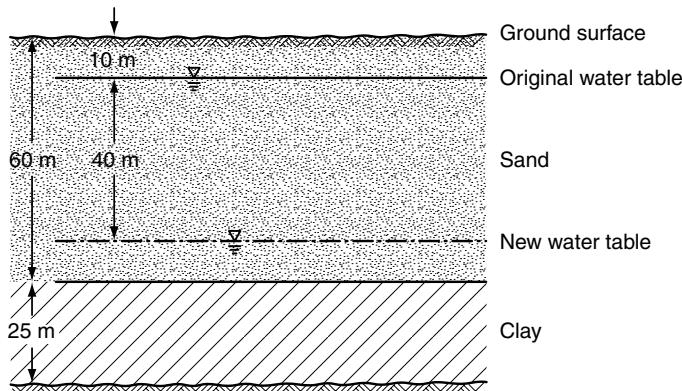


FIGURE 3.18 Subsidence problems.

is $(0 + 286.5)/2 = 143.25$ kPa and the settlement in the layer from 10 to 50 m is $S_{u1} = 40 * 143.25/10^5 = 0.0573$ m. The subsidence in the layer from 50 to 60 m is $S_{u2} = 10 * 286.5/10^5 = 0.0287$ m. The total subsidence is thus $0.0573 + 0.0287 = 0.086$ m.

Example 3.14

If the sand layer of the previous problem is underlain by a 25 m thick layer of clay with a modulus of elasticity of 10^3 N/cm 2 or 10^4 kN/m 2 , what is the total subsidence?

Solution. The increment of intergranular pressure of 286.5 kPa carries through the clay layer. The subsidence is thus $Z_3 = 25 * 286.5/10^4 = 0.716$ m. The total subsidence of the sand and clay layers is thus $0.086 + 0.716 = 0.802$ m. It is seen that 89% of the subsidence occurs in the clay layer because its modulus of elasticity is 10 times smaller.

3.10.2 Seepage Force

When water flows horizontally through an aquifer, the flow undergoes a reduction of pressure head because of friction. Thus the pressure on the upstream side of a small element is larger than on the downstream side. The water then exerts a net force on the aquifer element. The net force in the flow direction is the *seepage force*. This force can cause lateral displacements. If the drop of the water table in a length L is Δh , the horizontal movement is $S_h = \gamma_w \Delta h L / E_h$, where E_h is the modulus of elasticity in the horizontal direction. Letting i be the slope of the water table and $\Delta h = iL$, the horizontal movement, S_h , is calculated as

$$S_h = \gamma_w i \frac{L^2}{E_h} \quad (3.40)$$

If there is an upward vertical flow, the head loss due to friction, as the water flows into the pores, results in an increase in the hydrostatic pressure. This in turn results in a decrease of the intergranular pressure. A point can be reached when the upward seepage force is large enough to carry the weight of the sand grains so that the sand or silt behaves like a liquid. It has no strength to support any weight on it. This condition is known as *quicksand*. It is reached when the intergranular pressure vanishes and sand loses its bearing capacity. It can be shown that the upward hydraulic gradient necessary to produce quicksand is very close to one (Harr, 1962; Bauwer, 1978).

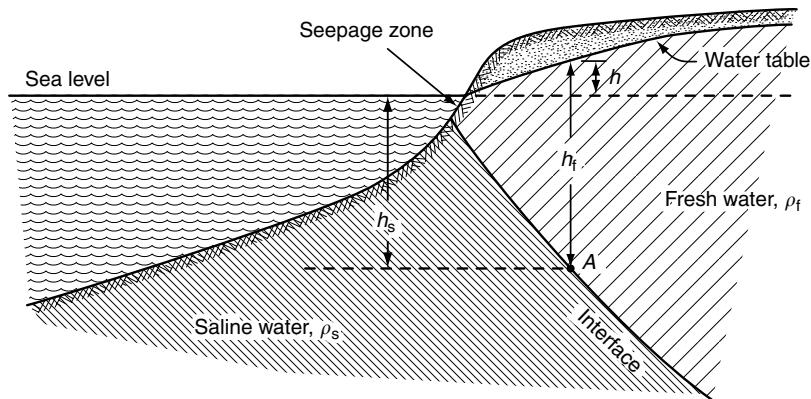


FIGURE 3.19 Coastal aquifer under natural conditions.

3.11 Salt Water Interfaces

The fresh and saline groundwaters have densities ρ_f and ρ_s . In coastal aquifers under natural conditions, the lighter fresh water lies over the heavier saline water and the flow is usually from the aquifer to the sea. Mixing of fresh water and salt water occurs only by molecular diffusion. Turbulent diffusion, the most effective mixing mechanism, is absent in aquifers. As a result, the mixing zone between salt water and fresh water is small compared with the thickness of the aquifer and an abrupt well-defined interface is usually assumed. At a point on the interface between the fresh and saline waters, the pressure of the fresh water, $\rho_f g h_f$, usually exceeds the pressure of the saline water, $\rho_s g h_s$, causing the flow from land to sea (Figure 3.19). But when pumping takes place in excess of replenishment, the drawdown of the water table creates a piezometric head in the fresh water that becomes less than in the adjacent salt water wedge. Then the saline water moves inland causing a *salt water intrusion*. The salt water may reach the well that becomes contaminated. Salt water intrusions have occurred in many coastal aquifers, for example, in the coastal aquifers of Florida, California, the Netherlands, Israel, and the South coast of England, mostly because of excessive pumping.

Assuming static conditions, the seepage zone is reduced to a point. The pressure at point A on the interface must be the same on the salt water side and on the fresh water side. Thus, the depth h_s of the interface below sea level is (Figure 3.19)

$$h_s = \frac{\rho_f}{\rho_s - \rho_f} h \quad (3.41)$$

where h is the height of the water table above sea level. With $\rho_s/\rho_f = 1.025$, the fraction in Equation 3.41 is equal to 40. For a confined aquifer, h is the fresh water piezometric head. The wedge of fresh water is known as the *Ghyben–Hertzberg lens* after the Dutch and German scientists who first obtained Equation 3.41.

When fresh water is pumped from an aquifer overlying a body of salt water, the drawdown of the fresh water table around the well causes a pressure reduction on the interface. This in turn causes the interface to rise below the well. This is called *upconing*. If the salt water cone reaches the well, it will discharge a mixture of salt and fresh water. For a water table well, assuming hydrostatic conditions, a rough approximation of the height δ of the cone is (Figure 3.20)

$$\delta = \frac{\rho_f}{\rho_s - \rho_f} s_w \quad (3.42)$$

where s_w is the drawdown at the well. More exact relationships are given in Chapter 12 on *Sea Water Intrusion into Coastal Aquifers* and, among others, in Bear and Dagan (1964), Dagan and Bear (1968), and Bear (1979).

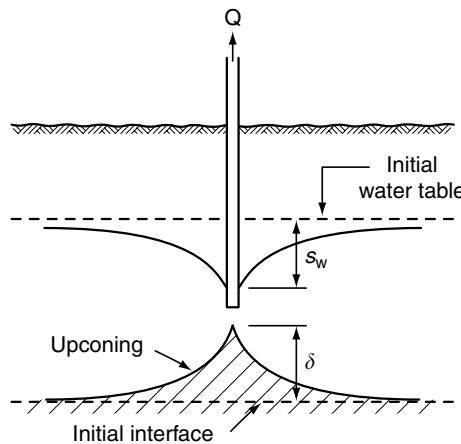


FIGURE 3.20 Upconing.

3.12 Groundwater Quality

The quality of groundwater is determined by the dissolved elements and gases and by the presence of suspended solids, bacteria, and viruses. The quality of the groundwater depends upon its natural and physical state and on the changes due to human activity. In its natural state, the dissolved elements and their concentrations depend on the chemical composition of the aquifer and on the travel time of the water through the rock formation. If the rock minerals are relatively soluble, slow water velocity and the ensuing long travel time result in a chemical equilibrium between the water and the rock medium. Because of the large range of flow velocities and of chemical compositions of the aquifers, there is a very wide range of compositions of the groundwater. Figure 3.21 shows the comparison of the range of concentrations of several constituents in groundwater to the concentrations of a 2.7 g sugar cube dissolved in a 2.7 l bottle, in a gasoline truck and in an oil tanker.

If the groundwater is no longer fit for a specific use, such as drinking, the water is said to be *contaminated*. If the water becomes heavily contaminated it is said to be *polluted*. Chapter 17 discusses in detail the types of groundwater contaminants, and a table of drinking water standards can be found in Table 32.2 in Chapter 32 on the legal framework for groundwater in the United States. Groundwater monitoring for water quality is discussed in Chapter 35.

Figure 3.22 illustrates the principal groundwater contaminant sources. These sources can be classified as *point*, *line*, or *nonpoint* (areal) sources. The geometry of the source affects the geometry of the contaminant plume. Contaminants can reach the groundwater in several manners. The contaminant may be *miscible*, that is, it can be dissolved in water, *immiscible*, that is, the water and the contaminant are in separate phases: lighter or heavier than water and the contaminant can be *adsorbed* on fine particles that are transported in suspension by the water. The miscible liquid and the water form a single phase. The movement of this single phase flow is governed by Darcy's law, discussed earlier in this chapter. Solutes are further classified as *conservative* and *nonconservative*. Conservative solutes remain stable in the groundwater: they do not react with the rock medium. Chloride solutions fall in this category. *Tracers* are conservative constituents that do not affect the viscosity and density of the water, such as fluoresceine in low concentration, for example. Conservative and nonconservative transport processes are discussed in detail in Chapter 18 and Chapter 19, respectively. Multiphase flow occurs primarily in the unsaturated zone, for example in the case of spills of hydrocarbons resulting in hydrocarbon and water phases. Multiphase flow is discussed in detail in Chapter 22 entitled "Contaminant Transport in the Unsaturated Zone—Theory and Modeling."

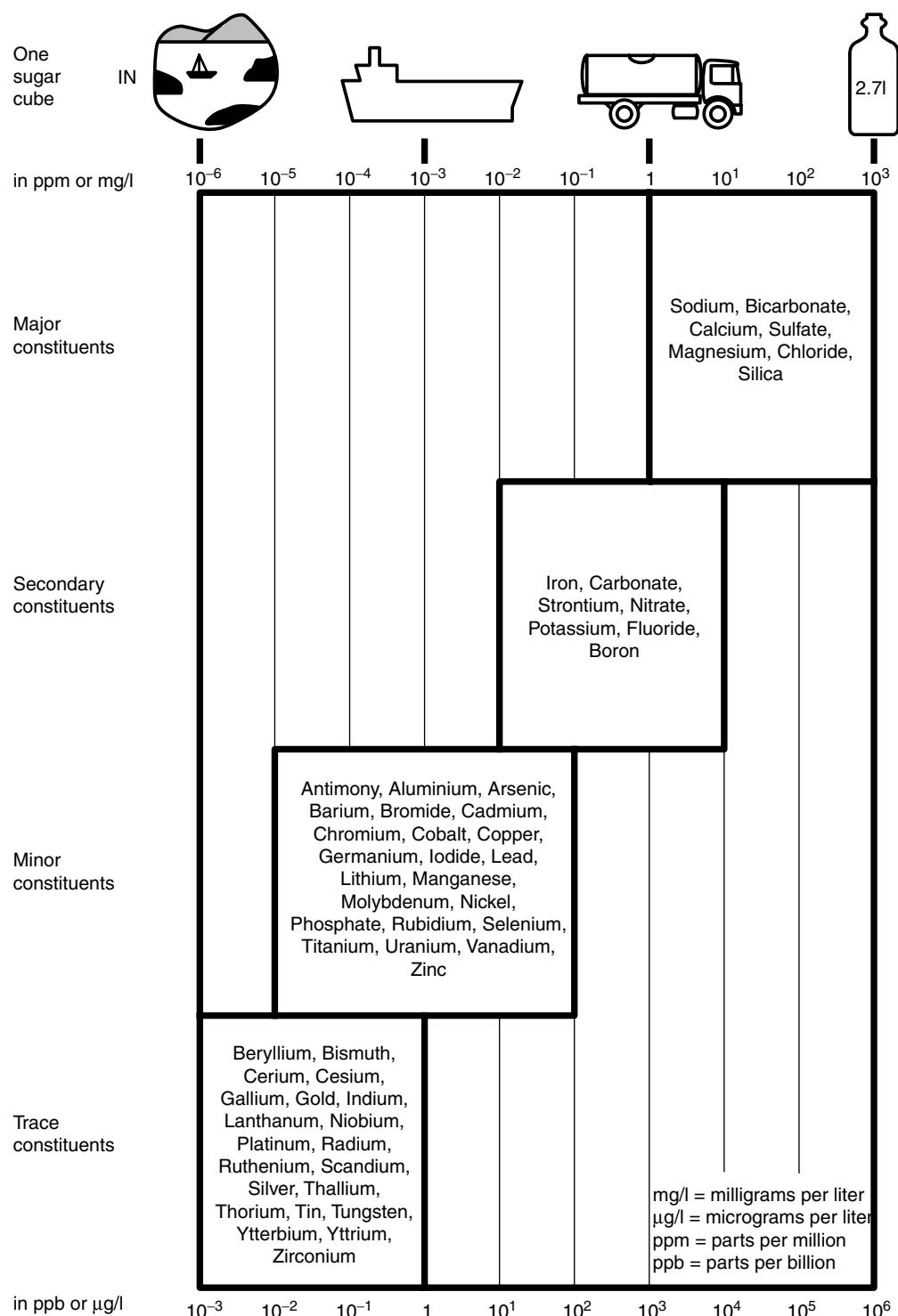


FIGURE 3.21 Examples of dissolved constituents in groundwater. (From Spitz, K. and Moreno, J. 1966. *Practical Guide to Groundwater and Solute Transport Modeling*, John Wiley & Sons, New York. With permission.)

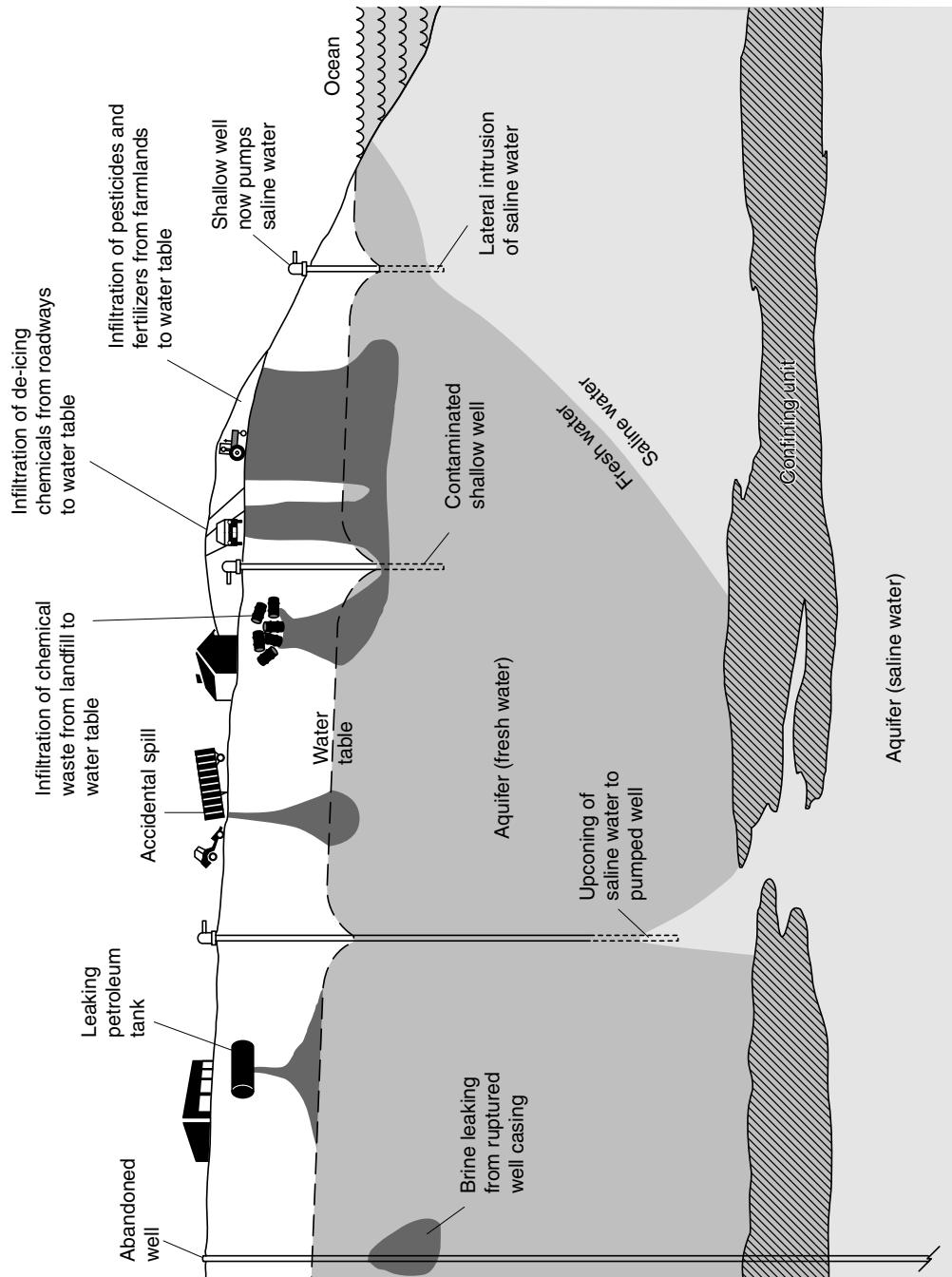


FIGURE 3.22 Sources of groundwater contamination. (From Johnston, R.H. 1986. *Water Quality Issues – Factors Affecting Ground-Water Quality*. U.S. Geological Survey, Water-Supply Paper 2325, Washington, D.C.)

3.13 Transport Mechanisms of Dissolved Contaminants

The principal fate and transport mechanisms are advection, diffusion, dispersion, sorption, and decay. These processes are discussed in their simplest 1D form in this section and in more detail in Chapter 18 and Chapter 19. Other processes such as hydrolysis, volatilization, and biotransformation are not discussed in this section. Biotransformation is discussed in Chapter 31. The combined effects of advection, dispersion, and biodegradation on the transport of contaminants were visualized in the Cape Cod experiments described in LeBlanc et al. (1991) (see also Section 18.10.2) as well as in the Experiments in the Columbus Air Force Base in Mississippi that are discussed in Chapter 26.

3.13.1 Advection

Advection is the transport of solute by the bulk groundwater flow. The average pore velocity, v , is obtained by dividing the Darcy flux q (see Example 3.4) by the effective porosity n_e

$$v = \frac{q}{n_e} \quad (3.43)$$

The 1D mass flux due to advection, F , is the product of the quantity of water flowing and the concentration of dissolved solids

$$F = v n_e C \quad (3.44)$$

The change of mass of contaminant over time in a control volume

$$n_e \frac{\partial C}{\partial t} dx dy dz$$

is equal to the balance between the mass inflow and outflow of contaminant

$$F dz dy - \left(F + \frac{\partial F}{\partial x} dx \right) dz dy = - \frac{\partial F}{\partial x} dx dy dz = - n_e \frac{\partial}{\partial x} (vC) dx dy dz$$

Thus, for a conservative solute in a homogeneous aquifer, the 1D advective transport equation is

$$\frac{\partial C}{\partial t} = -v \frac{\partial C}{\partial x} \quad (3.45)$$

where x is distance in the flow direction [L], t is time [T], C is the concentration [$M L^{-3}$], and v is the advective transport velocity [$L T^{-1}$]. It is observed that a precise estimation of the flow velocity is needed for an accurate estimation of the transport. In sand/gravel aquifers with significant groundwater, the plume movement is dominated by advection. However, it must be recalled that Equation 3.45 is not valid in karstic aquifers for which Darcy's law is not applicable.

3.13.2 Diffusion

Diffusion is the flux of solute from a zone of higher concentration to one of lower concentration due to the Brownian motion of ionic and molecular species. Under steady-state condition, the diffusion flux F is described by Fick's first law

$$F = -D \frac{\partial C}{\partial x} \quad (3.46)$$

where D is the *diffusion coefficient* [$\text{L}^2 \text{T}^{-1}$]. For diffusion in water, D for different cations and anions ranges approximately from 6×10^{-10} to $9 \times 10^{-9} \text{ m}^2/\text{s}$ (Fetter, 1999). For diffusion in porous media, Freeze and Cherry (1979) suggest taking an *effective diffusion coefficient* $D^* = \omega D$ to account for the tortuosity of the flow paths with ω ranging from 0.5 to 0.01 for laboratory studies of nonadsorbed ions in porous geological materials. A value of $\omega = 0.7$ was obtained by Perkins and Johnson (1963) for uniform sand columns. The change of concentration over time inside a control volume subject to diffusion flux is given by Fick's second law

$$\frac{\partial C}{\partial t} = D^* \frac{\partial^2 C}{\partial x^2} \quad (3.47)$$

Consider the case of two adjacent saturated strata. The first one is initially with zero concentration throughout its length $C(x, 0) = 0$ and the second stratum maintains constant concentration at the interface $C(0, t) = C_0$. Because of the concentration step at the interface, diffusion takes place from the interface into the first stratum. Far away in the first stratum, where the effect of diffusion has not yet reached, the concentration is still zero, namely $C(\infty, t) = 0$. Crank (1956) gives the following solution of Equation 3.47 subject to the above boundary and initial conditions for the evolution of the concentration as a function of time and distance from the interface:

$$C(x, t) = C_0 \operatorname{erfc} \left(\frac{x}{2\sqrt{D^* t}} \right) \quad (3.48)$$

where erfc is the complementary error function

$$\operatorname{erfc}(u) = 1 - \operatorname{erf}(u) \quad (3.49)$$

and erf is the error function

$$\operatorname{erf}(u) = \frac{2}{\pi} \int_0^u e^{-v^2} dv \quad (3.50)$$

Tables of the error function and of the complementary error function can be found, for example, in Appendix 3.1.

Example 3.15

Consider a substance diffusing at the boundary of a clay layer with an effective diffusion coefficient of $10^{-9} \text{ m}^2/\text{sec}$. Find the time it will take to obtain a concentration of 15% of the initial concentration at a distance of 20 m.

Solution. Equation 3.48 yields $\operatorname{erfc}[x/(2\sqrt{D^* t})] = 0.15$. Interpolating from a table of complementary error functions, one finds $x/(2\sqrt{D^* t}) = 1.02$. Thus

$$t = \frac{400}{4 * 10^{-9} (1.02)^2} = 9.61 \times 10^{10} \text{ sec} \approx 3047 \text{ yr}$$

This shows that diffusion is a very slow mechanism but over geological time it can become important.

The importance of diffusion increases as flow velocities decrease. Thus diffusion may be the governing transport mechanism in unfractured clays with low hydraulic conductivities. Diffusion can generally be neglected in gravel aquifers with high flow velocities. It can also be significant in fractured porous aquifers. Flow and transport processes in fractured rocks are discussed in Chapter 20.

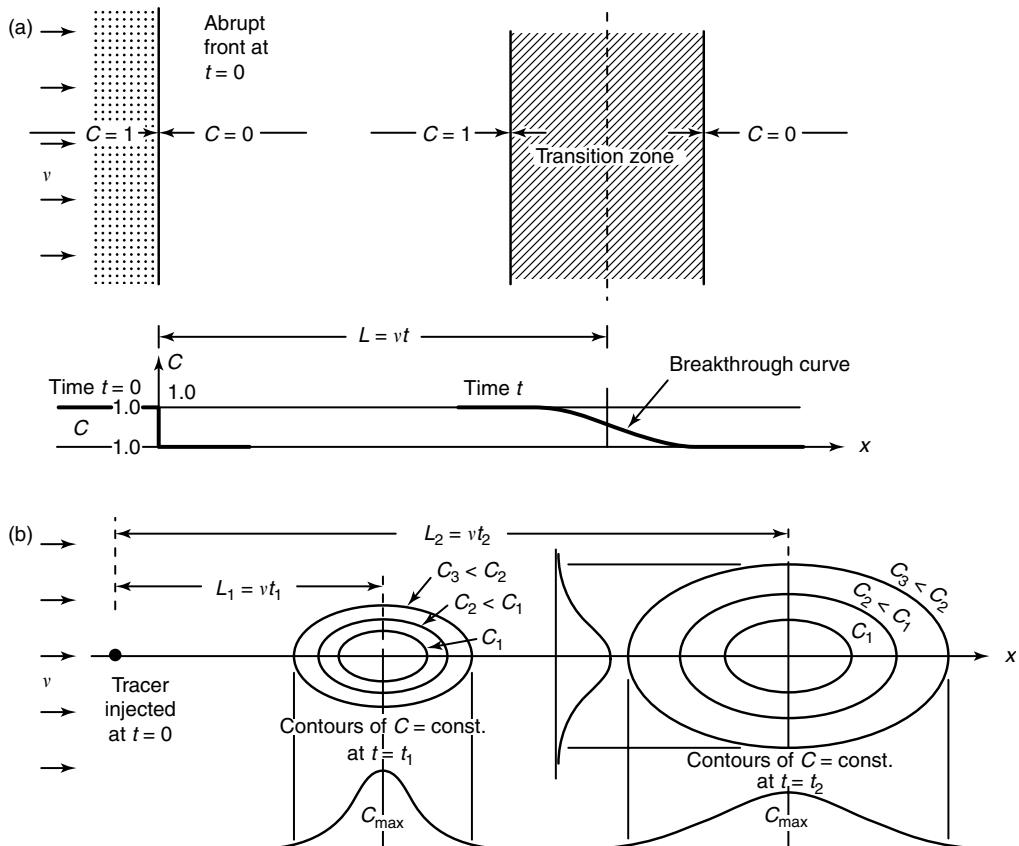


FIGURE 3.23 (a) Longitudinal and (b) transverse spreading due to mechanical dispersion. (From Bear, J. and Verruijt, A. 1987. *Modeling Groundwater Flow and Pollution*, Springer, Germany. With permission.)

3.13.3 Dispersion

Dispersion is the spreading of the plume that occurs along and across the main flow direction due to aquifer heterogeneities at both the small scale (pore scale) and at the macroscale (regional scale). Dispersion tends to increase the plume uniformity as it travels downstream. Factors that contribute to dispersion include faster flow at the center of the pores than at the edges, some pathways being longer than others, flow velocity larger in smaller pores than in larger ones. This is known as *mechanical dispersion*. The spreading due to both mechanical dispersion and molecular diffusion is known as *hydrodynamic dispersion*.

As a conceptual example, consider an aquifer with an abrupt concentration front at $t = 0$. At that time there is a *tracer*, such as sodium chloride, with a concentration $C = 1$ to the left of the front and there is no tracer to the right of the front, that is $C = 0$, as shown in Figure 3.23a. At a later time t , the *center* of the front has moved through a distance $L = vt$, where v is the pore flow velocity. But, due to dispersion, the tracer has spread around the center as shown in the lower part of Figure 3.23a. The plot of the S-shaped curve of the concentration at time t , $C(t)$, is called the *breakthrough curve*. In a second experiment, a quantity of tracer is injected at point $x = 0$ at time $t = 0$. At a later time t_1 , the center of the plume has moved a distance $L_1 = vt_1$ but, due to dispersion, the tracer has spread around the center with elliptical concentration contours as shown in Figure 3.23b. At time $t_2 > t_1$ the spreading has extended further as shown in the figure.

If one considers a *representative elementary volume* REV (see Chapter 4; Section 4.1.3), dispersion can be described by Fick's law (Equation 3.46). The *dispersion coefficient* D_L replaces the diffusion coefficient D^* and becomes a phenomenological coefficient that combines the effects of diffusion and dispersion. As mechanical dispersion is more pronounced in the longitudinal direction than in the transverse direction,

a longitudinal dispersion coefficient D_L and a transverse dispersion coefficient D_T are introduced. The longitudinal and the transverse dispersion–diffusion coefficients are defined as

$$D_L = \alpha_L v + D^* \quad (3.51a)$$

and

$$D_T = \alpha_T v + D^* \quad (3.51b)$$

where α_L is the longitudinal dispersivity [L], α_T is the transverse dispersivity [L], and v is the pore velocity [$L T^{-1}$].

The use of the Fickian theory to describe dispersion requires that the dispersion coefficients be travel-distance dependent or time dependent. A rough approximation based on averaging published data is (Gelhar, 1986, Gelhar et al., 1992)

$$\alpha_L \approx 0.1L \quad (3.52)$$

where L is the length of the flow path. Another estimate for flow lengths between 100 and 3500 m was given by Neuman (1990) as

$$\alpha_L \approx 0.0175L^{1.46} \quad (3.53a)$$

where α_L and L are in meters. Xu and Eckstein (1995) gave an estimate of α_L that does not have travel length limitation as

$$\alpha_L \approx 0.83(\log L)^{2.414} \quad (3.53b)$$

where α_L and L are in meters. This regression equation is based on weights of 1:2:3 to points of low, medium, and high reliability, respectively. When the scale is greater than 1 km, then the increase in longitudinal dispersivity becomes so small that it can be ignored without significant error.

The transverse dispersivity α_T is typically 1/10 to 1/100 of the longitudinal dispersivity α_L .

The combined advection–dispersion equation in 1D is thus

$$\frac{\partial C}{\partial t} = -v \frac{\partial C}{\partial x} + D_L \frac{\partial^2 C}{\partial x^2} \quad (3.54)$$

The solution of Equation 3.54 subject to the fixed step conditions:

$$\text{Initial condition} \quad C(x, 0) = 0 \quad x \geq 0$$

$$\text{Boundary condition} \quad C(0, t) = C_0 \quad t \geq 0$$

$$\text{Boundary condition} \quad C(\infty, t) = 0 \quad t \geq 0$$

has been given by Ogata and Banks (1961) as

$$C(x, t) = \frac{C_0}{2} \left[\operatorname{erfc} \left(\frac{x - vt}{2\sqrt{D_L t}} \right) + \exp \left(\frac{vx}{D_L} \right) \operatorname{erfc} \left(\frac{x + vt}{2\sqrt{D_L t}} \right) \right] \quad (3.55)$$

where x is the distance from the injection point. The argument of the exponential is the *Peclet number* ($P_e = vx/D_L$). The Peclet number is a measure of the ratio of the rate of transport by advection to the rate of transport by diffusion. For large Peclet numbers ($P_e > 100$), the advection dominates and the second term in the right-hand side becomes negligible. Figure 3.24 shows the behavior of the solution (3.55).

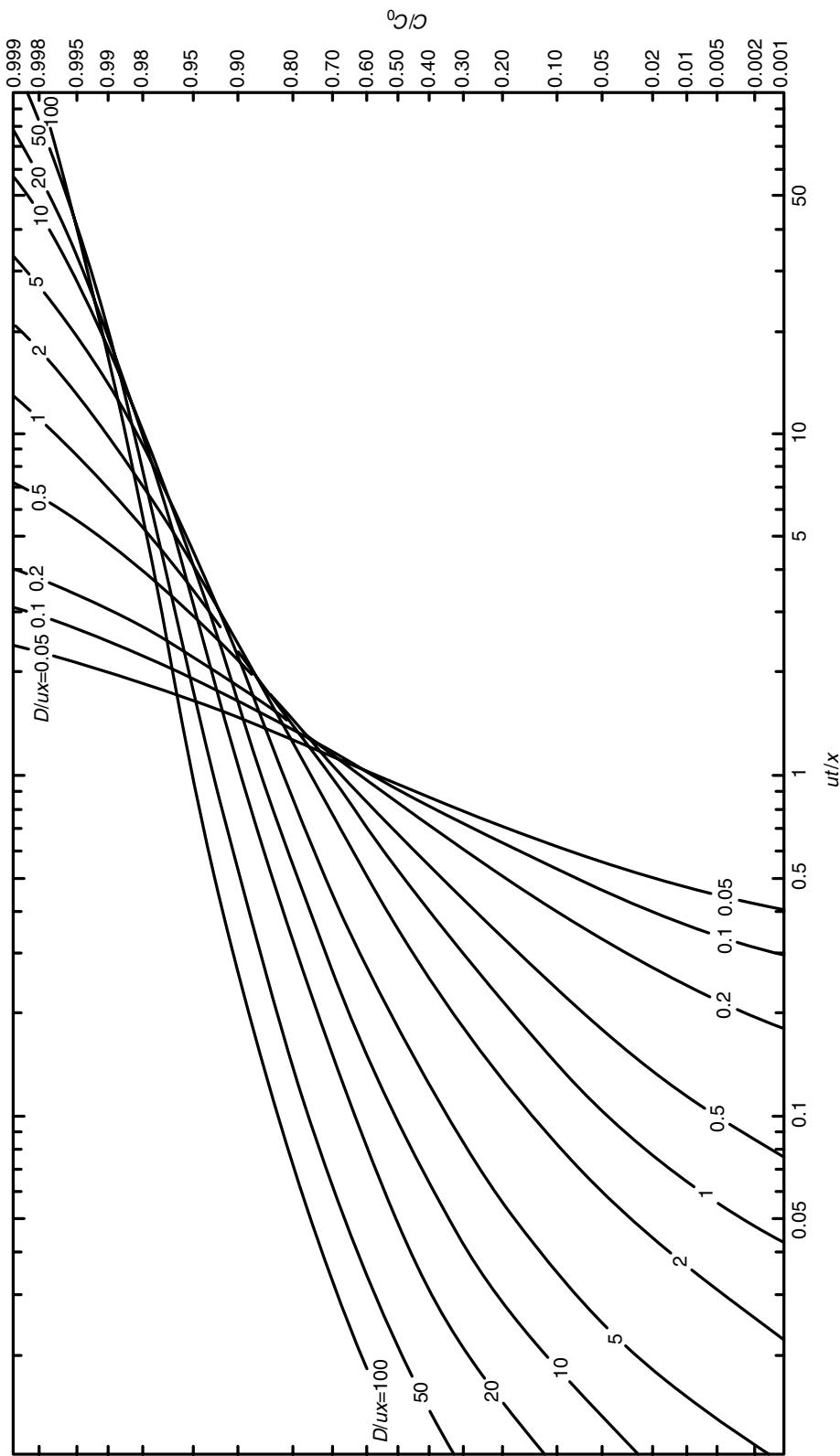


FIGURE 3.24 Solution of the 1D advection-dispersion equation. (From Ogata, A. and Banks, R.B. 1961. *A Solution of the Differential Equation of Longitudinal Dispersion in Porous Media*, U.S. Geological Survey, Professional Paper 411A, U.S. Government Printing Office, Washington, DC.)

If instead of a fixed step function a line source with continuous injection into the aquifer is considered as, for example, the leakage from a canal, Sauty (1978) found that in the solution (3.55) the + sign in the right hand side is replaced by a – sign. For large Peclet numbers the fixed step solution and the line source solution are essentially identical as the second term in the right hand side of Equation 3.55 is negligible. Further discussion and other boundary conditions are considered in Fetter (1999).

Example 3.16

An aquifer has a hydraulic conductivity of 2×10^{-5} m/sec, a hydraulic gradient of 0.003 m/m, an effective porosity $n_e = 0.2$, and an effective diffusion coefficient $D^* = 0.5 \times 10^{-9}$ m²/sec. A chloride solution with a concentration of 500 mg/l penetrates in the aquifer along a line source. Find the chloride concentration at a distance of 20 m from the point of entry, after a period of two years.

Solution. From Darcy's law, the pore flow velocity is

$$v = \frac{K}{n_e} dh/dx = \frac{2 \times 10^{-5} * 0.003}{0.2} = 3 \times 10^{-7} \text{ m/sec}$$

Using Neuman's approximation (3.53), the longitudinal dispersivity is

$$\alpha_L = 0.0175 * 20^{1.46} = 1.388 \text{ m}$$

The coefficient of longitudinal dispersion–diffusion, from Equation 3.51a is

$$D_L = 1.388 * 3.0 \times 10^{-7} + 0.5 \times 10^{-9} = 4.170 \times 10^{-7} \text{ m}^2/\text{sec}$$

The Peclet number is $P_e = vL/D_L = 3 \times 10^{-7} * 20 / 4.170 \times 10^{-7} = 14.387$. As $P_e < 100$ both terms in the right-hand side of Equation 3.55 need to be considered. Thus after 2 yr = $2 * 365 * 24 * 60 * 60 = 6.3072 \times 10^7$ sec and at a distance of 20 m the concentration is obtained from Equation 3.55 as

$$\begin{aligned} C &= \frac{500}{2} \left[\operatorname{erfc} \left(\frac{20 - 3 \times 10^{-7} * 6.3072 \times 10^7}{2\sqrt{4.170 \times 10^{-7} * 6.3072 \times 10^7}} \right) \right. \\ &\quad \left. + \exp 14.384 * \operatorname{erfc} \left(\frac{20 + 3 \times 10^{-7} * 6.3072 \times 10^7}{2\sqrt{4.170 \times 10^{-7} * 6.3072 \times 10^7}} \right) \right] \\ &= 250(\operatorname{erfc} 0.105 + 14.384 * \operatorname{erfc} 3.795) \approx 250 * 0.0.881984 = 220.5 \text{ mg/l} \end{aligned}$$

as the last $\operatorname{erfc}(\cdot)$ is negligible as seen in Appendix 3.1.

3.13.4 Sorption

Sorption refers to the exchange of molecules and ions between the solid phase and the liquid phase. It includes adsorption and desorption. *Adsorption* is the attachment of molecules and ions from the solute to the rock material. Adsorption produces a decrease of the concentration of the solute or, equivalently, causes a *retardation* of the contaminant transport compared with water movement. *Desorption* is the release of molecules and ions from the solid phase to the solute.

The relationship between the solute concentration in the adsorbed phase and in the water phase is called a *sorption isotherm*. The simplest expression is the *linear isotherm*

$$C_a = K_d C \quad (3.56)$$

where C_a is the sorbed concentration as mass of contaminant per mass of dry rock matrix [dimensionless], C is the dissolved concentration in mass of contaminant per volume of water [M L⁻³], and K_d is the

distribution coefficient [$\text{L}^3 \text{M}^{-1}$]. This expression implies that there is *equilibrium* between the adsorbed concentration and the dissolved concentration. This can be assumed when the adsorption process is fast compared with advection of contaminant.

The adsorption causes retardation in the migration of contaminants compared with advection. The contaminant transport gets more retarded as the fraction adsorbed increases. This effect can be described by a *retardation factor*, R_a , which for a linear isotherm, is

$$R_a = 1 + \frac{(1 - n)\rho_s}{n} K_d \quad (3.57)$$

where n is the porosity and ρ_s is the density of the solids. The retardation coefficient may take values from 1 to 10,000. The velocity of the solute front v_c (where the concentration is half that of the original concentration) is given by

$$v_c = \frac{v}{R_a} \quad (3.58)$$

The 1D advection–dispersion equation then becomes

$$\frac{\partial C}{\partial t} = -\frac{v}{R} \frac{\partial C}{\partial x} + \frac{D_L}{R_a} \frac{\partial^2 C}{\partial x^2} \quad (3.59)$$

where the term on the left side represents the time rate of change in storage of contaminant in the control volume, the first term on the right-hand side represents the retarded advective inflow–outflow, and the last term represents the retarded diffusion and dispersion. Solutions of the advection/dispersion/adsorption equation have been given by van Genuchten (1981).

The linear isotherm also applies to the case of organic compounds dissolved in groundwater. If the aquifer or soil contains at least 1% of organic carbon, then

$$K_d = K_{oc} f_{oc} \quad (3.60)$$

where K_{oc} is the organic carbon partition coefficient for the organic solute and f_{oc} is the fraction of organic carbon by weight in the soil. Methods for estimating K_{oc} from solubility data and from the octanol–water partition coefficient can be found, for example, in Fetter (1999) and in Spitz and Moreno (1996) (see Table 17.9 in Chapter 17 on “Groundwater Contaminants”; see also Section 3.15).

In this section it was assumed that the adsorption was fast compared with the advection of the contaminant. If instead the reaction is slow compared to the travel time and if chemical equilibrium cannot be attained, then it is necessary to describe the kinetics of the reaction. The reader is referred to Chapter 19 and to de Marsily (1986) or Fetter (1999).

3.13.5 Radioactive Decay and Degradation

The simplest model for decay of contaminants without transport is first order model:

$$\frac{\partial C}{\partial t} = -\lambda C \quad (3.61)$$

where λ is the first order decay rate constant [T^{-1}]. This relation also applies to *radioactive decay* and *degradation processes*. Equation 3.61 integrates as

$$C = C_0 e^{-\lambda t} \quad (3.62)$$

where C_0 is the concentration at time $t = 0$ and

$$\lambda = \frac{\ln 2}{T_{1/2}} \quad (3.63)$$

where $T_{1/2}$ is the *half-life* [T] of the radioactive isotope or of the degraded contaminant. Some values of radioactive and organic half-lives can be found in Spitz and Moreno (1996). The degradation causes a mass λC to disappear per unit volume per unit of time. The transport equation including decay is obtained by appending the quantity $-\lambda C$ to the right-hand side of Equation 3.59.

3.14 Monitoring, Site Remediation, and Landfills

Monitoring wells are necessary to measure the elevations of the water table or of the piezometric level, to collect samples of water for chemical analyses, to collect samples of nonaqueous phase liquids, to provide access for geophysical instruments, etc. The techniques used in groundwater monitoring and soil sampling are discussed in detail in Chapter 35 and the geophysical exploration techniques requiring wells and of other types are presented in Chapter 14 and Chapter 15.

Site remediation must consider at least source control and treatment of contaminated water or soil or both. Source control is necessary to prevent continuing discharge of contaminants to the subsurface or the groundwater. Treatment may be necessary to remove or substantially decrease the concentration of contaminants. Techniques for the proper design of landfills have been developed to eliminate or minimize the leakage to the vadose zone or the groundwater. These activities are regulated by legislation such as the Resource Conservation and Recovery Act (RCRA) and the Comprehensive Environmental Response, Compensation and Liability Act (CERCLA).

These topics of regulations, landfills, and site remediation are the subject of Chapters 32, 33, and 36, respectively.

3.15 Parameter Values

In the previous sections, a number of equations have been presented that describe the flow of groundwater and the transport of contaminants. The application of these equations and of the models that are based upon them requires the use of a number of parameters. In general, the hydrogeological parameters and the pollutant characteristics exhibit great variability and hence uncertainty exists in the model predictions. In many applications data may not be readily available. It may then be useful to have access to a database that brings together most of the previous experience, at least until field experiments can be conducted.

Spitz and Moreno (1996) present an extensive compilation of hydrogeological and pollutant transport parameters and list the literature sources of these data: porosity, pp. 342–344; specific yield, p. 345; horizontal hydraulic conductivity, pp. 346–350; vertical hydraulic conductivity, pp. 351–352; specific storage, p. 353; unsaturated hydraulic conductivity relationships, pp. 353–354; moisture content relationships, p. 354; molecular diffusion coefficients, pp. 368–369; mechanical dispersion coefficients, pp. 370–371; field-scale dispersivities, pp. 372–379; K_d values for metals, pp. 380–389; K_d values for organics, pp. 390–392; organic carbon content of sediments, f_{oc} , pp. 392; empirical K_{oc} relationships, pp. 393–394; radioactive half-lives, p. 395; half-lives for organics, pp. 396–409.

Glossary

Adsorption: Adhesion of solute molecules and ions to the rock or soil material.

Advection: Mass transport of solute by the gross movement of groundwater.

Conservative solute: Solute that remains stable and does not react with rock or soil material.

Contaminant: Substance that causes contamination.

Contamination: Degradation of groundwater quality that renders it unfit for domestic consumption.

Darcy's law: An equation that relates the gross flow velocity (i.e., the discharge divided by the gross cross-section of an aquifer segment) to the product of the hydraulic conductivity and the gradient of the total head.

Datum: An arbitrary reference elevation from which the hydraulic heads are measured.

Desorption: Removal of molecules or ions from the rock or the soil.

Diffusion: The flux of solute from areas of higher concentration to areas of lower concentration due to random molecular motion.

Dispersion: The spread of solute due to heterogeneities of the pore sizes and shapes (mechanical dispersion) and heterogeneities in the aquifer (macrodispersion).

Dispersion coefficient: A coefficient in Fick's law that relates mass flux to concentration gradient.

Dispersivity: A constant of dispersion that, when multiplied by the pore flow velocity, yields the dispersion coefficient.

Dupuit–Forchheimer assumption: The assumption of primarily horizontal flow, neglecting unimportant vertical flow.

Effective porosity: The part of the porosity that is available for the fluid flow.

Elevation head: Difference in elevation between a point in a flow field and an arbitrary reference datum, the latter often being taken as the mean sea level.

Equipotential line: Line of equal potential, used in the flow net.

Fick's law: Equation that describes the diffusion or dispersion of solutes.

Force potential: The product of the acceleration of gravity and the total head; it represents the total energy per unit of mass.

Flow net: A network of streamlines and equipotential lines that intersect at right angles.

Half-life: The time required for the concentration of a solute to be reduced to half its initial value by radioactive decay or biodegradation.

Head: See elevation head, pressure head, velocity head, and total head.

Hydraulic gradient: The rate of change of the piezometric head with displacement in a given direction.

Karst: Geological formation characterized by features associated with dissolution and collapse of carbonate rocks such as underground drainage, caves, sinkholes, and deep gullies. Named after the Karst plateau, a barren limestone plateau in West Slovenia.

Leakage: Seepage of water through a semipermeable layer called aquitard.

Leaky aquifer: An aquifer into which there is seepage from an overlying formation.

Linear isotherm: A chemical equilibrium relationship in which the concentration of the adsorbed solute in the solid phase is assumed to be proportional to the concentration in the water phase.

Longitudinal dispersion coefficient: Dispersion coefficient in the flow direction.

Monitoring well: A nonpumping well used to measure water levels or to obtain water samples for chemical analysis.

Organic: Pertaining to carbon compound usually in rings or chains and with other elements, not necessarily derived from a living organism.

Peclet number: A dimensionless quantity that expresses the relative importance of convection and dispersion of solutes.

Piezometric head: The sum of the elevation head and the pressure head.

Pollutant: Substance that causes pollution.

Pollution: Excessive contamination of the environment, in this chapter specifically contamination of groundwater that renders it unfit for human consumption as a result of human or natural activities.

Pressure head: The ratio of the fluid pressure intensity to the fluid specific weight; it has the dimension of length.

Retardation factor: Ratio of transport velocity of nonreacting solute to transport velocity of solute reacting with the solid phase.

Reynolds number: A dimensionless quantity that expresses the relative importance of inertia forces and viscous forces in a flow system. A small Reynolds number is associated with laminar flow; a large Reynolds number is associated with turbulent flow.

Salt-water interface: A surface forming a common boundary between adjacent salt water and fresh water.

Sorption: Includes adsorption and desorption.

Specific discharge: The flow rate through a cross-section of an aquifer divided by the area of that cross-section.

Streamline: A line that is everywhere tangent to the flow velocity vector, also called the flow line. It is used in the construction of flow nets.

Suction: A negative pressure head found in the unsaturated zone.

Total head: The sum of the elevation head, pressure head, and velocity head. In flow through porous media, the velocity head is small and often neglected in the calculation of the total head.

Tracer: A conservative solute that is used to track the path of groundwater movement.

Transverse dispersion coefficient: Dispersion coefficient in the direction perpendicular to the flow direction.

Velocity head: The kinetic energy of the flow per unit weight of fluid; it has the dimension of length.

Velocity potential: The product of the hydraulic conductivity and the total head. By virtue of Darcy's law, the negative of the gradient of the velocity potential is a flow velocity (specific discharge or Darcy velocity).

Water balance: An accounting of the inflows and outflows in a fluid control volume.

References

- Bear, J. 1979. *Hydraulics of Groundwater*, McGraw-Hill, New York.
- Bear, J. and Dagan, G. 1964. Some exact solutions of interface problems by means of the hodograph method, *Journal of Geophysical Research*, 69, 2, 1563–1572.
- Bear, J. and Verruijt, A. 1987. *Modeling Groundwater Flow and Pollution*, Reidel, Dordrecht, The Netherlands.
- Bedient, P.B., Rifai, H.S., and Newell, C.J. 1994. *Ground Water Contamination: Transport and Remediation*, Prentice Hall, Englewood Cliffs, NJ.
- Bouwer, H. 1969. Theory of seepage from open channels, in *Advances in Hydroscience*, Vol. 5, pp. 121–172, Ven Te Chow, Ed., Academic Press, New York.
- Bouwer, H. 1970. Groundwater Recharge Design for Renovating Waste Water, *Journal of the Sanitary Engineering Division, Proceedings of the ASCE*, 96, SA1, 59–74.
- Bouwer, H. 1978. *Groundwater Hydrology*, McGraw-Hill, New York.
- Bouwer, H. Artificial recharge of groundwater: hydrogeology and engineering, 2002. *Hydrogeology Journal*, 10, 1, 121–142.
- Bouwer, H., Back, J.T., and Oliver, J.M. 1999. Predicting infiltration and ground-water mounds for artificial recharge, *Journal of Hydrologic Engineering*, 4, 4, 350–357.
- Butler, J.J. Jr., 1997. *The Design, Performance and Analysis of Slug Tests*, Lewis Publishers, Boca Raton, FL.
- Charbeneau, R.J. 2000. *Groundwater Hydraulics and Pollution Transport*, Prentice Hall, Upper Saddle River, NJ.
- Chow, Ven Te. 1959. *Open-Channel Hydraulics*, McGraw-Hill, New York.
- Crank, J. 1956. *The Mathematics of Diffusion*, Oxford University Press, Oxford.
- Darcy, H. 1856. *Les fontaines publiques de la ville de Dijon*, V. Dalmont, Paris.
- Dagan, G. and Bear, J. 1968. Solving the problem of interface upconing in coastal aquifers by the method of small perturbations. *Journal of Hydraulic Research*, 6, 1, 15–44.
- Daugherty, L., Franzini, J.B., and Finnemore, E.J. 1985. *Fluid Mechanics with Engineering Applications*, 8th ed., McGraw-Hill, New York.

- de Marsily, G. 1986. *Quantitative Hydrogeology — Groundwater Hydrology for Engineers*. Academic Press, Orlando, FL.
- Deming, D. 2002. *Introduction to Hydrogeology*, McGraw-Hill, New York
- Domenico, P.A. and Schwartz, F.W. 1990. *Physical and Chemical Hydrogeology*, John Wiley & Sons, New York.
- Dupuit, J. 1863. *Études théoriques et pratiques sur le mouvement des eaux dans les canaux découverts et à travers les terrains perméables*, Dunod, Paris.
- Fetter, C.W. 1999. *Contaminant Hydrogeology*, 2nd ed. Prentice Hall, Upper Saddle River, NJ.
- Fetter, C.W. 2004. *Applied Hydrogeology*, 4th ed. Prentice Hall, Upper Saddle River, NJ.
- Fitts, C.R. 2002. *Groundwater Science*, Academic Press, New York.
- Forchheimer, P. 1930. *Hydraulik*, Teubner Verlagsgesellschaft, Stuttgart.
- Finnemore, E.J. and Franzini, J.B. 2002. *Fluid Mechanics with Engineering Applications*, McGraw-Hill, New York.
- Freeze, R.A. and Cherry, J.A. 1979. *Groundwater*, Prentice-Hall, Englewood Cliffs, NJ.
- Galloway, D., Jones, D.R., and Ingebritsen, S.E. 1999. *Land Subsidence in the United States*, U.S. Geological Survey, Circular 1182, Reston, VA.
- Gelhar, L.W. 1986. Stochastic subsurface hydrology from theory to application, *Water Resources Research*, 22, 9, 135S–145S.
- Gelhar, L.W., Welty, C., and Rehfeldt, K.R. 1992. A critical review of data on field-scale dispersion in aquifers. *Water Resources Research*, 28, 7, 1955–1974.
- Harr, M.E. 1962. *Ground Water and Seepage*, McGraw-Hill, New York.
- Heath, R.C. 1995. *Basic Ground-Water Hydrology*, U.S. Geological Survey, Water-Supply Paper 2200, 7th printing, Denver, CO.
- Hibbs, B.J. 1997. Ground-water modeling and land subsidence prediction in the Harris-Galveston coastal subsidence district, Texas, in *Proceedings of the 24th Annual Water Resources Planning and Management Conference*, D.H. Merritt, Ed. Huston, TX. April 6–9, 388–393, ASCE, Reston, VA.
- Hubbert, M.K. 1953. Entrapment of petroleum under hydrodynamic conditions, *Bulletin of American Association of Petroleum Geologists*, 37, 1954–2026.
- Hubbert, M.K. 1987. Darcy's law: its physical theory and application to entrapment of oil and gas, in *History of Geophysics*, Vol. 3, pp. 1–26. C.S. Gillmor, Series Editor, AGU, Washington, DC.
- Hubbert, M.K., 1956. Darcy's law and the field equations of flow of underground fluids, *Transactions, American Institute of Mining, Metallurgical, and Petroleum Engineers*, 297, 222–239. Dallas, TX.
- Hudak, P.F. 2005. *Principles of Hydrogeology*, 3rd ed. CRC Press, Boca Raton, FL.
- Johnston, R.H. 1986. Water quality issues — Factors affecting ground-water quality. U.S. Geological Survey, *Water Supply Paper* 2325, Washington, DC.
- Jumikis, A.R. 1984. *Soil Mechanics*, Robert E. Krieger Co., Malabar, FL.
- Karplus, W.J. 1958. *Analog Simulation*, McGraw-Hill, New York.
- LeBlanc, D.R., S.P. Garabedian, K.M. Hess, L.W. Gelhar, R.D. Quadri, K.G. Stollenwerk, W.W. Wood, 1991. Large-scale natural gradient tracer test in sand and gravel, Cape Cod, Massachusetts, 1, Experimental design and observed tracer movement, *Water Resources. Research*, 27(5), 895–910.
- Lowman, S.W. 1972. *Ground-Water Hydraulics*, Geological Survey Professional Paper 708, U.S. Government Printing Office, Washington, DC.
- Mariño, M.A. 1975. Artificial groundwater recharge, II rectangular recharging area, *Journal of Hydrology*, 26, 29–37.
- National Research Council 1984. *Groundwater Contamination*, National Academy Press, Washington, DC.
- Neuman, S.P. 1990. Universal scaling of hydraulic conductivities and dispersivities in geologic media. *Water Resources Research*, 26, 8, 1749–1758.
- Ogata, A. and Banks, R.B. 1961. *A Solution of the Differential Equation of Longitudinal Dispersion in Porous Media*, U.S. Geological Survey, Professional Paper 411A, U.S. Government Printing Office, Washington, DC.

- Perkins, T.K. and Johnson, O.C. 1963. A review of diffusion and dispersion in porous media, Society of Petroleum Engineers Journal, 3, 70–84.
- Poland, J.F. (Ed.) 1984. *Guidebook to Studies of Land Subsidence due to Ground-Water Withdrawal*, UNESCO, International Hydrological Programme, Working Group 8.4, Paris.
- Polubarinova-Kochina, P.Ya. 1962. *Theory of Ground Water Movement*, Translated from Russian by J.M. Roger De Wiest, Princeton University Press, Princeton, NJ.
- Palmer, C.M. 1996. *Contaminant Hydrogeology*, Lewis Publishers, Boca Raton, FL.
- Price, M. 1996. *Introducing Groundwater*, Chapman & Hall, London.
- Sauty, J.-P. 1978. Identification des paramètres du transport hydrodispersif dans les aquifères par interprétation de traçage en écoulement cylindrique convergent ou divergent, *Journal of Hydrology*, 39, 1/2, 69–103.
- Saquib, M.N., Davis, L.C., Khan, A.Q., and Taghavi, S.A. 1995. Imperial county groundwater evaluation: model applications, in *Proceedings of the 22nd Annual Conference on Integrated Water Resources for the 21st Century*, Cambridge, MA, May 7–11, M.F. Domenica, ed., ASCE, Reston, VA, pp. 929–932.
- Schneebeli, G. 1955. Expériences sur la limite de validité de la loi de Darcy et l'apparition de la turbulence dans un écoulement de filtration, *La Houille Blanche*, 10, 141–149.
- Schwartz, F.W. and Zhang, H. 2003. *Fundamentals of Groundwater*, John Wiley & Sons, New York.
- Spitz, K. and Moreno, J. 1996. *A Practical Guide to Groundwater and Solute Transport Modeling*, John Wiley & Sons, New York.
- Subramanya, K., Madhav, M.R., and Mishra, G.C. 1973. Studies on seepage from canals with partial lining, *Journal of Hydraulics Division, Proceedings of the ASCE*, 99, HY12, 2333–2351.
- Taghavi, S.A., Saquib, M.N., Khan, A.Q., and Davis, L.C. 1995. Imperial county groundwater evaluation: model development and calibration, in *Proceedings of the 22nd Annual Conference on Integrated Water Resources for the 21st Century*, Cambridge, MA, May 7–11, M.F. Domenica, ed., ASCE, Reston, VA, pp. 933–936.
- Todd, D.K. 1980. *Groundwater Hydrology*, 2nd ed., Wiley, New York.
- Todd, D.K. and Mays, L.W. 2005. *Groundwater Hydrology*, 3rd ed., John Wiley & Sons, Hoboken, NJ.
- Van der Leeden, F. 1991. *Geraghty & Miller's Groundwater Bibliography*, 5th ed., Water Information Center, Plainview, NY.
- van Genuchten, M.T. 1981. Analytical solutions for chemical transport with simultaneous adsorption, zero-order production and first order decay, *Journal of Hydrology*, 49, 213–233.
- Xu, M. and Eckstein, Y. 1995. Use of weighted least squares method in evaluation of the relationship between dispersivity and field scale, *Ground Water*, 33, 6, 905–908.
- Yussuff, S.M., Chauhan, H.S., Kumar, M. and Srivastava, V.K. 1994. Transient canal seepage to sloping aquifer, *Journal of the Irrigation and Drainage Engineering, Proceedings of the ASCE*, 120, 1, 97–109.
- Zheng, C., and Bennett, G.D. 2002. *Applied Contaminant Modeling*, 2nd ed., Wiley Interscience, New York.

Further Information

This chapter deals with elementary groundwater flow and transport problems. More advanced aspects are deferred to subsequent chapters. The following books give an introductory treatment of groundwater problems:

Heath R.C. [1995] provides an excellent, practical, and well-illustrated introduction to groundwater flow, well hydraulics, well tests, and groundwater pollution.

Hudak P.F. [2005] gives introductory principles and applications of hydrogeology for students with little or no background in hydrogeology. The principal topics covered are aquifers, monitoring wells, groundwater flow, wells and aquifer tests, groundwater quality and solute transport, and management of groundwater pollution.

Lowman S.W. [1972] gives a good introductory treatment of hydrologic properties of water-bearing materials, flowing wells, aquifer tests by well methods and by areal methods.

National Research Council [1984] provides a well-documented, nonmathematical introduction to groundwater contamination, including case studies.

Palmer C.M. [1996] gives a nonmathematical introduction to contaminant hydrogeology. A practical approach to completing investigations and the basics of collecting data are presented.

Price M. [1996] gives an excellent physical, nonmathematical, and exceptionally well-written description of the phenomena associated with groundwater occurrence, movement, pumping, and pollution.

There are a number of excellent textbooks on the subject of this chapter. Here we select a few:

Bedient et al. [1994] give a good introduction on groundwater hydraulics and then concentrate on groundwater quality, pollution of groundwater, contaminant transport, and groundwater management and remediation.

Deming D. [2002] discusses introductory hydrogeological theory and covers other topics such as environmental hydrogeology, heat transport, earthquakes, and petroleum migration. It is easy to read and is well illustrated.

Fetter C.W. [1999] focuses on the transport of contaminants in the saturated and unsaturated zones and gives practical information on monitoring and site remediation.

Fetter C.W. [2004] provides a general treatment of groundwater occurrence, movement, contamination, development and management, and models. It is a very popular textbook, now in its fourth edition.

Fitts C.R. [2002] has written a good undergraduate textbook. The mathematical modeling is treated in more detail than in the other chapters.

Freeze R.A. and Cherry J.A. [1979] have written a classical text that covers the fundamentals of groundwater geology, the flow and chemical evolution of groundwater, and groundwater contamination. It is still one of the favorite textbooks and a standard reference.

Schwartz F.W. and Zhang H. [2003] provide an excellent up-to-date textbook that integrates the concepts of hydrology with geology. Its organization is similar to that of Freeze R.A. and Cherry J.A. [1979] but includes many results developed in the interim time. It includes a section on groundwater modeling centered on MODFLOW and an excellent chapter on groundwater management. Transport processes are very well presented. There are some similarities to Domenico P.A. and Schwartz F.W. [1990] discussed below.

Todd D.K. [1980] wrote a good practical text on groundwater hydrology. This has recently been updated and extended as Todd D.K. and Mays L.W. [2005]. It contains a good listing of journals, web sites, and organizations with interests in subsurface flows and contaminants issues.

At a more advanced level there are also excellent textbooks among which we have selected the following:

Bear J. [1979] gives an excellent mathematical treatment of the laws and equations that describe the flow of groundwater and the management of groundwater.

Bear J. and Verruijt A. [1987] are concerned with the movement and accumulation of groundwater and pollutants in aquifers, and the construction of conceptual and mathematical models. A number of computer programs written in BASIC are included.

de Marsily G. [1986] presents a very lucid and classical treatment of the mathematics of groundwater flow and contaminant transport.

Domenico P.A. and Schwartz F.W. [1990] give an advanced treatment of physical and chemical hydrogeology. This book has more breadth and depth than Schwartz F.W. and Zhang H. [2003] discussed above.

Zheng C. and Bennett G.D. [2002] provide an excellent graduate level text on modeling of contaminant transport that is useful for students and practitioners. The first part includes advective transport, diffusion, dispersion, chemical reactions, and numerical formulation. Part 2 is concerned with field applications and part 3 is on advanced topics that include variable density, unsaturated flow and transport, and groundwater management.

If one needs a handy bibliographical reference of publications in the groundwater field there is Van der Leeden F. [1991]. He lists approximately 5600 references on general bibliographies, journals, texts, handbooks, and dictionaries, and references by subjects such as history, environment, geophysical exploration, well logging, hydraulics of soils and aquifers, theory of groundwater flow, pumping, well maintenance, pumping equipment, tracers, water quality, contamination, salt water intrusion, models, laws and regulations, management, etc.

Appendix 3.1 Values of the Error Function and Complementary Error Function

x	$\text{erf } x$	$\text{erfc } x$	x	$\text{erf } x$	$\text{erfc } x$
0	0	1.0	1.0	0.842701	0.157299
0.05	0.056372	0.943628	1.1	0.880205	0.119795
0.1	0.112643	0.887537	1.2	0.910314	0.089686
0.15	0.167996	0.823004	1.3	0.934008	0.065992
0.2	0.222703	0.777297	1.4	0.952285	0.047715
0.25	0.276326	0.723674	1.5	0.996105	0.033895
0.3	0.328627	0.671373	1.6	0.976348	0.023652
0.35	0.379382	0.620618	1.7	0.983790	0.016210
0.4	0.428392	0.571608	1.8	0.989091	0.010909
0.45	0.475482	0.524518	1.9	0.992790	0.007210
0.5	0.520500	0.479500	2.0	0.995322	0.004678
0.55	0.563323	0.436677	2.1	0.997021	0.002979
0.6	0.603856	0.396144	2.2	0.998137	0.001863
0.65	0.642029	0.357971	2.3	0.998857	0.001143
0.7	0.677801	0.322199	2.4	0.999311	0.000689
0.75	0.711156	0.288844	2.5	0.999593	0.000407
0.8	0.742101	0.257899	2.6	0.999764	0.000236
0.85	0.770668	0.229332	2.7	0.999866	0.000134
0.9	0.796908	0.203092	2.8	0.999925	0.000075
0.95	0.820891	0.179109	2.9	0.999959	0.000041
			3.0	0.999978	0.000022

$$\text{erf } x = (4/\pi)^{1/2} \int_0^x \exp(-z)^2 dz$$

$$\text{erf } (0) = 0; \text{erf } (\infty) = 1; \text{erf } (-x) = -\text{erf } (x); 1 - \text{erf } (x) = \text{erfc } s(x)$$

4

Two- and Three-Dimensional Flow of Groundwater

4.1	Fundamentals	4-1
	Introduction • Continuity Equation • Macroscopic Approach • Equation of Motion • Extensions of Darcy's Law	
4.2	Groundwater Flow Equations.....	4-10
	General • Groundwater Flow Equation • Boundary Conditions	
4.3	Hydraulic Approach to Groundwater Flow	4-15
	Concept • Motion Equations • Flow Equation for a Confined Aquifer • Flow Equation for a Phreatic Aquifer • Flow Equation for an Aquitard	
4.4	Examples of Solutions	4-24
	Dupuit Discharge Formula • Hvorslev Slug Test Equation • Theis Well Flow Equation • Hantush Leaky Well Flow Equation	
	Glossary.....	4-33
	References	4-34
	Further Information	4-36

F. De Smedt
Vrije Universiteit Brussel, Belgium

4.1 Fundamentals

4.1.1 Introduction

The objective of this chapter is to show the logical development of equations that explain and predict the movement of groundwater in two- and three-dimensions (2D and 3D) as observed in nature. Flow of groundwater is a special case of fluid flow in porous media, and is governed by the laws of physics, in particular the laws of fluid mechanics. Fluid mechanics deals with the motion of fluids and with the forces exerted on solid bodies in contact with fluids. The fundamental principles are *conservation of mass* and *conservation of momentum*. It will be shown that these fundamental principles are sufficient to explain groundwater flow.

The purpose of this chapter is to obtain flow equations that enable to calculate the state variables of groundwater in each point of the flow domain and if necessary also in time. Therefore, groundwater flow

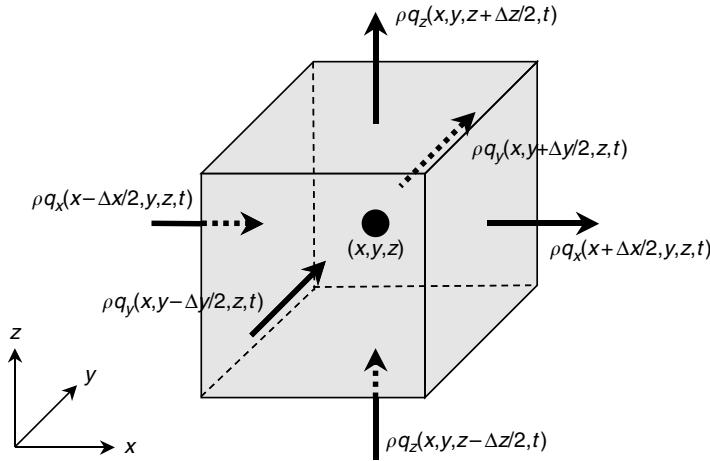


FIGURE 4.1 Mass conservation of the water phase in a reference elementary volume of saturated porous material.

equations are expressed in the form of partial differential equations, with the spatial coordinates, and time as independent variables. Hence, the variables that describe the state of groundwater are explicit functions of position and time, as for instance the groundwater potential

$$h = h(x, y, z, t) \quad (4.1)$$

where h represents the groundwater potential, the space position is denoted by Cartesian coordinates x, y, z , and time is represented by the symbol t ; x, y, z , and t are independent variables, while other variables, as h , are considered dependent, that is, with each set of values (x, y, z, t) variables can be associated, as $h(x, y, z, t)$.

Variables have dimensions. Whenever a new variable is introduced the dimensions will be indicated between brackets, []. The symbols used are: [L] for length, [T] for time, [M] for mass, and [F = ML/T²] for force. For practical applications the dimensions can be substituted by any consistent set of units. A logical choice would be the International Systems of Units (SI units), that is, meter (m), second (sec), kilogram (kg), and Newton (N), respectively; however, because groundwater flow is usually extremely slow often days are used instead of seconds for the time dimension.

4.1.2 Continuity Equation

The first fundamental law governing groundwater flow is the continuity equation, that expresses the principle of mass conservation. Consider in a fully saturated soil an *elementary control volume* centered around a point with Cartesian coordinates (x, y, z) as shown in Figure 4.1. It is customary and convenient to choose the z -axis vertical and pointing upward in the positive direction. The size of the elementary volume is Δx in the x -direction, Δy in the y -direction, and Δz in the z -direction. At a certain time instant, t , the mass of groundwater, M , present in the elementary control volume is given by

$$M = \rho n \Delta x \Delta y \Delta z \quad (4.2)$$

where ρ is the density of water with dimensions [M/L³], and n the porosity of the porous medium with dimensions [L³/L³]. The principle of mass conservation implies that the net result of inflow minus outflow is balanced by the change in storage versus time, or

$$\frac{\partial M}{\partial t} = \text{inflow} - \text{outflow} \quad (4.3)$$

Hence, it is necessary to calculate the groundwater flows through the sides of the elementary control volume in order to evaluate the net result of inflow minus outflow. The amount of groundwater flow is denoted by means of the flux, \mathbf{q} , which is the volumetric discharge or flow rate per cross-sectional area. The flux is a vector quantity with components along the x -, y - and z -directions: $\mathbf{q} = (q_x, q_y, q_z)$; the dimensions are [$L^3/T/L^2 = L/T$]. For instance, the mass inflow of groundwater along the left side, situated at $x - \Delta x/2$ of the control volume, is given by

$$\rho q_x(x - \Delta x/2, y, z, t) \Delta y \Delta z \quad (4.4)$$

As Δx is small, q_x at position $x - \Delta x/2$ can be approximated by a *Taylor series expansion*, where only zero and first order terms are maintained, such that this inflow can be calculated as

$$\left(\rho q_x - \frac{\Delta x}{2} \frac{\partial \rho q_x}{\partial x} \right) \Delta y \Delta z \quad (4.5)$$

In this expression ρq_x and its derivative with respect to x are evaluated at the center of the control volume. Similar expressions can be established for the other sides; for instance at the right side, at position $x + \Delta x/2$ of the control volume, the groundwater outflow is given by

$$\rho q_x(x + \Delta x/2, y, z, t) \Delta y \Delta z \approx \left(\rho q_x + \frac{\Delta x}{2} \frac{\partial \rho q_x}{\partial x} \right) \Delta y \Delta z \quad (4.6)$$

Hence, the total inflow minus outflow can be calculated as

$$\begin{aligned} & \left(\rho q_x - \frac{\Delta x}{2} \frac{\partial \rho q_x}{\partial x} \right) \Delta y \Delta z + \left(\rho q_y - \frac{\Delta y}{2} \frac{\partial \rho q_y}{\partial y} \right) \Delta x \Delta z + \left(\rho q_z - \frac{\Delta z}{2} \frac{\partial \rho q_z}{\partial z} \right) \Delta x \Delta y \\ & - \left(\rho q_x + \frac{\Delta x}{2} \frac{\partial \rho q_x}{\partial x} \right) \Delta y \Delta z - \left(\rho q_y + \frac{\Delta y}{2} \frac{\partial \rho q_y}{\partial y} \right) \Delta x \Delta z - \left(\rho q_z + \frac{\Delta z}{2} \frac{\partial \rho q_z}{\partial z} \right) \Delta x \Delta y \end{aligned} \quad (4.7)$$

Working out term by term and combining, yields following expression

$$-\left(\frac{\partial \rho q_x}{\partial x} + \frac{\partial \rho q_y}{\partial y} + \frac{\partial \rho q_z}{\partial z} \right) \Delta x \Delta y \Delta z \quad (4.8)$$

Using the *del operator* $\nabla = (\partial/\partial x, \partial/\partial y, \partial/\partial z)$ [L^{-1}], this can be written as

$$-\nabla \cdot (\rho \mathbf{q}) \Delta x \Delta y \Delta z \quad (4.9)$$

where the dot represents the scalar product of two vectors.

The principle of mass balance states that the inflow minus outflow is equal to the change in storage; using Equation 4.2 the change in storage is given by

$$\frac{\partial M}{\partial t} = \frac{\partial}{\partial t} (\rho n \Delta x \Delta y \Delta z) \quad (4.10)$$

In this expression, the variables that can change with time are the density of the water, ρ , because water is compressible, and the porosity, n , and size of the control volume, $\Delta x \Delta y \Delta z$, because the porous medium can be compressible. However, for the latter it is assumed that under natural conditions only vertical

deformation needs to be considered, such that only n , and Δz depend upon time while $\Delta x \Delta y$ remains constant. Hence, the storage term can be worked out by using the rules of differentiation as

$$\begin{aligned}\frac{\partial M}{\partial t} &= \rho n \Delta x \Delta y \frac{\partial \Delta z}{\partial t} + \rho \frac{\partial n}{\partial t} \Delta x \Delta y \Delta z + \frac{\partial \rho}{\partial t} n \Delta x \Delta y \Delta z \\ &= \rho \left(\frac{n}{\Delta z} \frac{\partial \Delta z}{\partial t} + \frac{\partial n}{\partial t} + \frac{n}{\rho} \frac{\partial \rho}{\partial t} \right) \Delta x \Delta y \Delta z\end{aligned}\quad (4.11)$$

In addition, it is generally assumed that Δz changes only because of changes in porosity, while the solid phase may be considered as incompressible. Hence, the solid volume per unit surface area, $(1 - n)\Delta z$, remains constant in time, or

$$\frac{\partial}{\partial t}[(1 - n)\Delta z] = (1 - n) \frac{\partial \Delta z}{\partial t} - \Delta z \frac{\partial n}{\partial t} = 0 \quad (4.12)$$

from which we obtain the relationship between changes in n , and Δz

$$\frac{\partial n}{\partial t} = \frac{1 - n}{\Delta z} \frac{\partial \Delta z}{\partial t} \quad (4.13)$$

so that Equation 4.11 becomes

$$\frac{\partial M}{\partial t} = \rho \left(\frac{1}{\Delta z} \frac{\partial \Delta z}{\partial t} + \frac{n}{\rho} \frac{\partial \rho}{\partial t} \right) \Delta x \Delta y \Delta z \quad (4.14)$$

The compression of a porous formation can be expressed in function of the water pressure (Bear, 1972)

$$\frac{1}{\Delta z} \frac{\partial \Delta z}{\partial t} = \alpha \frac{\partial p}{\partial t} \quad (4.15)$$

where α is the elastic compressibility coefficient of the porous formation, with dimensions $[L^2/F]$, and p represents the groundwater pressure, $[F/L^2]$. The compressibility of the water can be expressed by a similar law

$$\frac{1}{\rho} \frac{\partial \rho}{\partial t} = \beta \frac{\partial p}{\partial y} \quad (4.16)$$

with β the compressibility coefficient of water, $[L^2/F]$. It is assumed that other effects on the density, such as solutes or temperature, are of minor importance and can be ignored. Substitution of these relations into Equation 4.14 gives

$$\frac{\partial M}{\partial t} = \rho \left[(\alpha + n\beta) \frac{\partial p}{\partial t} \right] \Delta x \Delta y \Delta z \quad (4.17)$$

In the derivation of Equation 4.17 we have followed the classical approach initiated by Jacob (1940) and finalized by Jacob (1950), where the elemental volume is allowed to deform. This was later criticized by, for example, De Wiest (1966) and Cooper (1966), arguing that a control volume should not deform. A more rigorous development using a fixed control volume and fixed coordinates is possible. In this case the movement of the solid phase has to be taken into account and also its conservation of mass. This has been achieved by Gambolati and Freeze (1973), Freeze and Cherry (1979), and Dagan (1989), resulting in an expression that is independent of the coordinate system and that corresponds to Equation 4.17 for the Cartesian coordinate system used here. Another more general approach considers 3D deformation as proposed by Biot (1955); a very good overview of this approach is given by Verruijt (1969).

The continuity equation is obtained by putting the change in storage, Equation 4.17, equal to the net inflow given by Equation 4.9, and dividing by $\Delta x \Delta y \Delta z$ to express the mass balance per unit volume of porous medium. This gives following result

$$\rho \left[(\alpha + n\beta) \frac{\partial p}{\partial t} \right] = -\nabla \cdot (\rho \mathbf{q}) \quad (4.18)$$

This relationship states the principle of mass conservation of groundwater in its most general form in a notation that is independent from the choice of a particular coordinate system. However, often density changes of water are of limited importance, and a simplified continuity equation can be obtained by dividing Equation 4.18 by ρ , and neglecting spatial density differences as

$$\left[(\alpha + n\beta) \frac{\partial p}{\partial t} \right] = -\frac{1}{\rho} \nabla \cdot (\rho \mathbf{q}) = -\nabla \cdot \mathbf{q} - \frac{1}{\rho} \mathbf{q} \cdot \nabla \rho \approx -\nabla \cdot \mathbf{q} \quad (4.19)$$

In this simplified form, the continuity equation expresses the groundwater balance on a volumetric basis. The left-hand side of the equation gives the change in volume of groundwater present in the porous medium; this change in storage can be due to compression of the medium and the water. The right-hand side of the equation gives the convergence or divergence of the volumetric flow rate of groundwater. When the flow rate is converging, the storage increases, and vice versa; when it diverges the storage decreases.

4.1.3 Macroscopic Approach

The obtained continuity equation is elementary, yet fundamental, but not without ambiguity. The problem arises from the fact whether Δx , Δy , and Δz can be made small enough to justify the truncation of the Taylor series to terms of order zero and one in the derivation of the inflows and outflows, as in Equation 4.6. Indeed, when the limit is taken of Δx , Δy , and Δz going to zero, the elementary control volume reduces to a point, such that the concept of a porous medium becomes illusive, because a point is either situated in the solid phase, the air phase, or the water phase, and variables as water content and groundwater flux lose their meaning.

Hence, it seems that groundwater flow problems cannot be formulated and solved correctly at the microscopic level. Of course, such a microscopic approach is not really of any interest in practice. But nevertheless, this ambiguity needs to be cleared up. This is achieved by abandoning the microscopic pore-scale level and by moving to a coarser macroscopic level. At this level microscopic features as solid grains and in between pore spaces are ignored, and the medium is conceived as a continuous space with average properties, such that porosity exists in any point of the medium, regardless of whether this point is situated in the solid phase or in the pore space. Also the variables describing the state and movement of water through the medium become macroscopic, therefore the average behavior is described, and not the fate of individual fluid particles moving in the pores of the porous medium.

The question remains how macroscopic variables need to be defined and interpreted. Different approaches are in use. The first and still most popular technique is the method of the representative elementary volume (REV), developed by Bear (1972), where macroscopic variables are defined as mean values over an REV, and attributed to the center of the REV. The exact size of the REV is not determined, but it is assumed that the size is much larger than the pore scale and much smaller than the scale of the porous medium. The results obtained with this approach yield macroscopic values for the groundwater flow, representing the average behavior of the fluid over the REV. The main advantage of the method is that macroscopic quantities have a clear physical meaning as they are measurable in the field by taking and analyzing soil samples of an adequate size. The main disadvantages are the loss of detailed information on the microscopic pore scale level, and the uncertainty due to the assumption that the values of the macroscopic variables are independent of the size of the REV. A detailed discussion can be found in Pinder (1983).

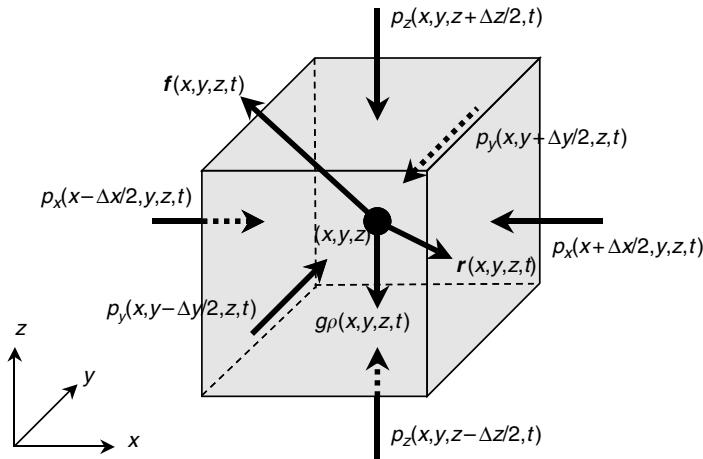


FIGURE 4.2 Forces acting in the water phase of a reference elementary volume of saturated porous material.

A second, more recent, technique is a statistical approach taking into account the uncertainty of the spatial distribution on a microscopic scale (Dagan, 1989). The microscopic arrangement of the porous medium is considered to be random, such that a set of media can be imagined with similar characteristics. The macroscopic variables are considered to be the averages of the variables of the media in the set. The main conceptual difficulty stems from the fact that statistical averaging should be carried out over an ensemble of realizations, whereas in practice usually only one particular porous formation is available from which the statistical information needs to be determined. This is feasible under the assumption of ergodicity, that is, the characteristics of one sample are assumed to be representative for the whole set. Hence, with the theory of random functions, macroscopic laws can be derived. As long as only relationships between averaged quantities are derived, with no special concern about their fluctuations, the results obtained are essentially the same as with the REV approach. However, the statistical technique is more powerful when dealing with fluctuations and deviations, as in case of particle or pollutant transport in groundwater.

4.1.4 Equation of Motion

The second fundamental law is the momentum equation, based on Newton's second law of motion, that is, forces induce motion or a change in motion. Consider the elementary control volume as used previously when deriving the continuity equation. An inventory of all forces acting on the water present in the control volume can be made in order to obtain a momentum equation. Under natural conditions, the forces to consider are pressure forces, gravity forces, and reaction forces of the solid matrix exerted on the fluid. As forces are vectors, it is necessary to consider different components along the different directions, as shown in Figure 4.2. For instance, along the left side of the control volume, pressure is acting on the water phase, which yields following contribution to the force balance in the x -direction

$$np(x - \Delta x/2, y, z, t)\Delta y\Delta z \quad (4.20)$$

where the porosity, n , appears in the expression because the water occupies only that fraction of the boundary. A similar force is acting on the right side, but in opposite direction as

$$-np(x + \Delta x/2, y, z, t)\Delta y\Delta z \quad (4.21)$$

Using truncated Taylor series expansions to relate the np -terms to the center of the control volume, the resulting pressure force component in the x -direction becomes

$$\left[\left(np - \frac{\Delta x}{2} \frac{\partial np}{\partial x} \right) - \left(np + \frac{\Delta x}{2} \frac{\partial np}{\partial x} \right) \right] \Delta y \Delta z = -\frac{\partial np}{\partial x} \Delta x \Delta y \Delta z \quad (4.22)$$

Similar expressions can be obtained for the pressure force components in the y -, and z -directions, acting on the other sides of the elementary control volume.

The gravity force only acts in the z -direction downwards, and is equal to the total weight of the water in the control volume

$$-\rho g n \Delta x \Delta y \Delta z \quad (4.23)$$

where g is the gravity constant, with dimension [L/T²].

The evaluation of the reaction force of the solid material on the water is more complicated. It consists of forces acting against the water pressure and friction forces due to the groundwater movement. These forces are extremely difficult to evaluate on a pore scale level, because the shape of the contact surface between the solid phase and the water phase is very complex from geometrical point of view. As, the exact contact surface is generally not known and would be very difficult to express in mathematical terms anyway, it is impossible to describe these forces on a microscopic scale, and one is forced to adapt a less precise macroscopic approach. Therefore, the reaction forces are defined as average body forces per water volume; the reaction force against the water pressure is denoted as $\mathbf{r} = (r_x, r_y, r_z)$, [F/L³], and the friction force against water movement, as $\mathbf{f} = (f_x, f_y, f_z)$, [F/L³]. Hence, the effect of these forces in the x -direction can be written as

$$(r_x + f_x) n \Delta x \Delta y \Delta z \quad (4.24)$$

with similar expressions in the y - and z -directions.

Now, the force balance in the three directions can be calculated as,

$$\left[-\frac{\partial np}{\partial x} + (r_x + f_x) n \right] \Delta x \Delta y \Delta z \quad (4.25)$$

$$\left[-\frac{\partial np}{\partial y} + (r_y + f_y) n \right] \Delta x \Delta y \Delta z \quad (4.26)$$

$$\left[-\frac{\partial np}{\partial z} + (r_z + f_z) n \right] \Delta x \Delta y \Delta z \quad (4.27)$$

Using the del operator, ∇ , these can be combined in one vector equation

$$[-\nabla(np) - \rho g n \nabla z + (\mathbf{r} + \mathbf{f}) n] \Delta x \Delta y \Delta z \quad (4.28)$$

This equation can be worked out further and simplified by dividing by the total volume of water present in the control volume, $n \Delta x \Delta y \Delta z$, yielding

$$-\nabla p - \frac{p}{n} \nabla n - \rho g \nabla z + \mathbf{r} + \mathbf{f} \quad (4.29)$$

Next, note that in case the fluid is at rest, the sum of all forces should be zero. Also, there would be no friction, $\mathbf{f} = 0$, and the pressure should be hydrostatic, $\nabla p = -\rho g \nabla z$, from which the overall reaction force, \mathbf{r} , of the grains opposing the water pressure can be evaluated (Dagan, 1989, p. 73) as

$$\mathbf{r} = \frac{p}{n} \nabla n \quad (4.30)$$

Substituting this result in the force balance simplifies Equation 4.29 to

$$-\nabla p - \rho g \nabla z + \mathbf{f} \quad (4.31)$$

In case of motion, the sum of forces is not zero, but equals the change in momentum of the fluid. Furthermore, the friction along the solid–water interface is non-zero and should be specified in function of the motion and friction properties. Several additional assumptions and considerations are necessary to arrive at a useful result.

From field observations it is known that groundwater flow under natural conditions is generally very slow, leading to a series of important simplifications. First, changes in momentum are also very small and can be neglected compared to other forces acting on the fluid. Hence, although the fluid is in motion, the forces acting on the fluid are approximately in equilibrium as seen in the force balance equation given below.

$$-\nabla p - \rho g \nabla z + \mathbf{f} \approx 0 \quad (4.32)$$

This type of flow is known in fluid dynamics as creeping motion. The active forces yielding motion, as pressure and gravity, are immediately balanced by resisting friction forces of equal strength. Of course, in porous media this is due to the large contact area between fluid and solid material which causes extensive friction, such that significant movement of the fluid is prevented.

Second, because water is a viscous fluid the friction force results from viscous momentum transfer between the stagnant pore walls and the moving fluid. In principle, the *Navier–Stokes equation* should enable to calculate the overall resistance force, but due to the complex geometry of the pore walls and pore sizes this is unfeasible in practice. However, in case of viscous creeping flow it is known that for an isotropic medium the overall resistance force has a component that is opposite in direction to the fluid flow, proportional to the viscosity of the fluid and the magnitude of the flow, and depending on the size of the obstacles in the flow field. Hence, it is justified to express the friction force as

$$\mathbf{f} = -\mu \underline{\underline{k}}^{-1} \cdot \mathbf{q} \quad (4.33)$$

where μ is the dynamic viscosity of the fluid, [FT/L²], \mathbf{q} is the flux as defined in the previous paragraph, here representing the amount and direction of the water movement, and $\underline{\underline{k}}$ is matrix with proportionality factors representing the geometry of the pore space; this matrix of coefficients has dimensions [L²] and is denoted as intrinsic permeability or sometimes permeability in short, for reasons that will become evident hereafter.

Substitution of Equation 4.33 into the force balance Equation 4.32, and rearranging, results in the following equation of motion

$$\mathbf{q} = -\mu^{-1} \underline{\underline{k}} \cdot (\nabla p + \rho g \nabla z) \quad (4.34)$$

The obtained expression, called Darcy's law, describes the flow of a fluid in a porous medium, also if the fluid has a variable density (and viscosity). Such conditions can be present in, for instance, coastal aquifers where salt and fresh water intermix, or in geothermal reservoirs where the density (and viscosity) of the fluid changes with temperature.

However, if density gradients are not significant, or when the density only changes due to compressibility of the fluid, the equation of motion can be simplified as

$$\mathbf{q} = -\rho g \mu^{-1} \underline{\underline{k}} \cdot \nabla(\phi + z) = -\underline{\underline{K}} \cdot \nabla h \quad (4.35)$$

where $\underline{\underline{K}} = \rho g \mu^{-1} \underline{\underline{k}}$ is the hydraulic conductivity matrix [L/T], $\phi = p/\rho g$ the pressure potential [L], and $h = \phi + z$ the groundwater potential [L]. Hence, Darcy's law, which originally was based on experimental evidence, is nothing else but an expression of equilibrium of forces reduced to a form suitable for describing flow of fluids in porous media. The derivation presented above sheds more light on the underlying principles and assumptions that result in Darcy's law, and enables to appreciate more its applicability in field conditions.

The basic assumption leading to Darcy's law is that movement of a fluid through a porous medium is very restricted, due to large friction forces that balance the driving forces for motion. Hence, Darcy's law is applicable in cases as water flow in soils or other types of granular porous media, or flow in fractured rocks, but not in cases as caves or other large sized openings, like cracks, fissures, and the like. These latter problems should be analyzed with viscous flow theories. Next, all ambiguity as discussed in case of the mass balance equation is also present here. Darcy's law is a macroscopic approach; on a microscopic level there are no such things as hydraulic conductivity or permeability. Hence, one might wonder about the errors involved in using a macroscopic approach as Darcy's law. However, this is not really an important issue, because the uncertainty in obtaining representative and accurate values for the hydraulic conductivity or permeability has a much larger effect on the overall accuracy. Indeed, field investigations show that natural porous media as aquifers and aquitards exhibit a large variability in conductive properties, which is difficult to quantify accurately by experimental or deterministic means.

4.1.5 Extensions of Darcy's Law

Darcy's law, like the momentum equation, is a vector relationship. When the flow is 3D, a Darcy's law can be written for each of the directions. For instance, in case of an isotropic porous medium, the conductivity matrix becomes a scalar, K , resulting in three Darcy equations

$$q_x = -K \frac{\partial h}{\partial x} \quad (4.36)$$

$$q_y = -K \frac{\partial h}{\partial y} \quad (4.37)$$

$$q_z = -K \frac{\partial h}{\partial z} \quad (4.38)$$

In case of other coordinate systems, the rules of vector-calculus should be applied in order to find the different expressions along the coordinate axes. For instance, in a cylindrical coordinate system (r, φ, z) , this becomes

$$q_r = -K \frac{\partial h}{\partial r} \quad (4.39)$$

$$q_\varphi = -\frac{K}{r} \frac{\partial h}{\partial \varphi} \quad (4.40)$$

$$q_z = -K \frac{\partial h}{\partial z} \quad (4.41)$$

In addition to heterogeneity, that is, porosity and conductivity variations from point to point, a dependence on direction is possible. This is the case for so-called anisotropic porous media, where due to some

direction-related properties, as preferential lining of fractures, stratifications, or layering, the conductivity changes depending upon direction. Such situations can be described by an extension of previous equations, where the conductivity becomes a symmetrical matrix, $\underline{\underline{K}}$, with following components

$$\underline{\underline{K}} = \begin{bmatrix} K_{xx} & K_{xy} & K_{xz} \\ K_{xy} & K_{yy} & K_{yz} \\ K_{xz} & K_{yz} & K_{zz} \end{bmatrix} \quad (4.42)$$

In Cartesian coordinates, Darcy's law becomes

$$q_x = -K_{xx} \frac{\partial h}{\partial x} - K_{xy} \frac{\partial h}{\partial y} - K_{xz} \frac{\partial h}{\partial z} \quad (4.43)$$

$$q_y = -K_{xy} \frac{\partial h}{\partial x} - K_{yy} \frac{\partial h}{\partial y} - K_{yz} \frac{\partial h}{\partial z} \quad (4.44)$$

$$q_z = -K_{xz} \frac{\partial h}{\partial x} - K_{yz} \frac{\partial h}{\partial y} - K_{zz} \frac{\partial h}{\partial z} \quad (4.45)$$

Notice that a potential gradient in one direction can yield flows in other directions. Especially in fractured porous media, anisotropy can be important. Anisotropy also plays a role in case of layered formations, usually of sedimentary origin. For horizontal layering only two conductivity components exist: a horizontal conductivity, K_h , and a vertical conductivity, K_v . In such case, Darcy's law becomes

$$q_x = -K_h \frac{\partial h}{\partial x} \quad (4.46)$$

$$q_y = -K_h \frac{\partial h}{\partial y} \quad (4.47)$$

$$q_z = -K_v \frac{\partial h}{\partial z} \quad (4.48)$$

These equations are useful in practice, also in view of the fact that effects of horizontal layering on the conductivity are rather the rule than the exception in ground layers. In case of folded or dipping layered formations also off-diagonal terms are important; all components of the conductivity tensor can be calculated from the dip angle and the conductivity values normal and parallel to the layering, following the rules of tensor calculus (Bear, 1972).

4.2 Groundwater Flow Equations

4.2.1 General

The groundwater flow equation is obtained by combining the continuity equation with Darcy's law. The most general form is obtained when the groundwater balance, Equation 4.19, is combined with the general Darcy Equation 4.34, yielding

$$(\alpha + n\beta) \frac{\partial p}{\partial t} = \nabla \cdot \left[\mu^{-1} \underline{\underline{k}} \cdot (\nabla p + \rho g \nabla z) \right] \quad (4.49)$$

Note that the porosity, n , on the left-hand side of the equation, depends upon the water pressure due to elastic deformation of the porous medium; actually the same holds for the permeability matrix, $\underline{\underline{k}}$.

However, these effects are generally ignored and the groundwater flow equation is expressed as a linear partial differential function of the water pressure as follows

$$S_p \frac{\partial p}{\partial t} = \nabla \cdot \left[\mu^{-1} \underline{\underline{K}} \cdot (\nabla p + \rho g \nabla z) \right] \quad (4.50)$$

where S_p is the storage coefficient of the porous medium related to water pressure changes, with dimensions $[L^2/F]$, and given by

$$S_p = \alpha + n\beta \quad (4.51)$$

This storage coefficient depends upon different soil and water properties. Hence, it is a rather complicated parameter, but fortunately it has a simple physical interpretation, that is, the storage coefficient gives the volume of water released by a unit volume of porous medium due to a unit decline of the water pressure.

It is instructive to write the obtained groundwater flow equation in Cartesian coordinates, taking into consideration different horizontal and vertical permeabilities

$$S_p \frac{\partial p}{\partial t} = \frac{\partial}{\partial x} \left(\frac{k_h}{\mu} \frac{\partial p}{\partial x} \right) + \frac{\partial}{\partial y} \left(\frac{k_h}{\mu} \frac{\partial p}{\partial y} \right) + \frac{\partial}{\partial z} \left[\frac{k_v}{\mu} \left(\frac{\partial p}{\partial z} + \rho g \right) \right] \quad (4.52)$$

This equation describes 3D movement of groundwater in its most general form. Notice that as a consequence the flow equation is expressed as a function of the water pressure, and not as a function of the groundwater potential. A state of the art of density-driven flow in porous media is for instance given by Holzbecher (1998).

4.2.2 Groundwater Flow Equation

In practice the general groundwater flow equation is not much used, because simplifications are usually introduced. When density-driven flow can be ignored, a groundwater flow equation can be obtained by combining the continuity equation with the simple form of Darcy's law, given by Equation 4.35, which yields

$$(\alpha + n\beta) \frac{\partial p}{\partial t} = \nabla \cdot (\underline{\underline{K}} \cdot \nabla h) \quad (4.53)$$

Making use of the fact that when the fluid density is constant, the water pressure differences in time can be related to the temporal variation of the groundwater potential

$$\frac{\partial p}{\partial t} = \rho g \frac{\partial h}{\partial t} \quad (4.54)$$

the resulting flow equation can be written as

$$\rho g (\alpha + n\beta) \frac{\partial h}{\partial t} = \nabla \cdot (\underline{\underline{K}} \cdot \nabla h) \quad (4.55)$$

This expression can be considered as a basic groundwater flow equation, because it enables to calculate the groundwater potential, h , in function of characteristics of the porous medium and the fluid. The equation is usually written in following form

$$S_s \frac{\partial h}{\partial t} = \nabla \cdot (\underline{\underline{K}} \cdot \nabla h) \quad (4.56)$$

where S_s is the specific storage coefficient, which depends upon compressibility of the porous medium and the fluid, then

$$S_s = \rho g(\alpha + n\beta) \quad (4.57)$$

The specific storage can be interpreted physically as the volume of water released by a unit volume of saturated porous medium due to a unit decline of the groundwater potential. The units are [L^{-1}]. Common values are usually very small because elastic deformation of ground layers or water is limited and certainly not noticeable in the field. Of course, plastic deformation of ground layers, consisting of materials like clay or peat, is excluded. Such deformations mainly occur under the influence of external loads or extensive groundwater pumping, and are noticeable in the field as *land subsidence*. Elastic deformations and storage in aquifers are described by Verruijt (1969). Theories on porous medium deformability under various soil conditions, with special attention to the subject of land subsidence, are discussed in part 2 of Bear and Corapcioglu (1984).

Written in Cartesian coordinates and using vertical and horizontal conductivities, the saturated groundwater flow equation becomes

$$S_s \frac{\partial h}{\partial t} = \frac{\partial}{\partial x} \left(K_h \frac{\partial h}{\partial x} \right) + \frac{\partial}{\partial y} \left(K_h \frac{\partial h}{\partial y} \right) + \frac{\partial}{\partial z} \left(K_v \frac{\partial h}{\partial z} \right) \quad (4.58)$$

This is the most used groundwater flow equation in practice. The equation is of the diffusive type, indicating that potential differences will be dissipated through the medium with a speed depending upon the value of K/S_s . Hence, because S_s is usually small, this process can be rather fast, unless K is also small as for instance in case of clay layers.

It is also interesting to write the flow equation in cylindrical coordinates (r, φ, z) as

$$S_s \frac{\partial h}{\partial t} = \frac{1}{r} \frac{\partial}{\partial r} \left(r K_h \frac{\partial h}{\partial r} \right) + \frac{1}{r^2} \frac{\partial}{\partial \varphi} \left(K_h \frac{\partial h}{\partial \varphi} \right) + \frac{\partial}{\partial z} \left(K_v \frac{\partial h}{\partial z} \right) \quad (4.59)$$

This equation is used for the prediction of groundwater flow toward wells, as for instance described by Hantush (1964), Walton (1970), Huisman (1972), and Kruseman and de Ridder (1991).

When groundwater flow is stationary, the variables become independent of time, and the flow equation is reduced to

$$\nabla \cdot (K \nabla h) = 0 \quad (4.60)$$

Written in Cartesian coordinates, and generalizing by using vertical and horizontal conductivities, we obtain

$$\frac{\partial}{\partial x} \left(K_h \frac{\partial h}{\partial x} \right) + \frac{\partial}{\partial y} \left(K_h \frac{\partial h}{\partial y} \right) + \frac{\partial}{\partial z} \left(K_v \frac{\partial h}{\partial z} \right) = 0 \quad (4.61)$$

This equation clearly shows that movement of groundwater is a *potential flow problem*. The driving forces for the movement are differences in groundwater potential and the resulting fluxes depend upon the conductive properties of the medium. In this way groundwater flow resembles other types of potential flow, like heat flow in heat conducting media under influence of temperature gradients, or electric flow in electric conductive materials under influence of electric potential differences. Many solutions have been obtained for groundwater flow problems based on the theory of potential functions, especially in 2D (Polubarnova-Kochina, 1962; Harr, 1962; Verruijt, 1970; Halek and Svec, 1979; Strack, 1989).

Another possibility is to assume that the groundwater flow occurs under unsaturated conditions, but we will not go into detail as these derivations can be found in Chapter 6.

4.2.3 Boundary Conditions

Groundwater flow equations are partial differential equations that relate the dependent variable, the groundwater potential, to the independent variables, the coordinates of the flow domain and the time. However, flow equations are only valid inside the flow domain, and not on the boundaries, because there the porous medium ends and other phenomena occur, which can influence the groundwater flow inside of the domain. These interactions at the boundaries are governed by other physical laws and have to be described separately by mathematical expressions, the so-called *boundary conditions*.

First, consider the time dimension. As transient groundwater flow equations contain first order partial derivatives vs. time, it can be shown mathematically that only conditions at the start, that is, at time zero, are needed. Hence, knowing the distribution of the dependent variable in the flow domain at the start is sufficient to describe the further evolution in time. For instance, in case the groundwater potential is the dependent variable, the so-called initial condition is

$$h(x, y, z, 0) = h_0(x, y, z) \quad (4.62)$$

where h_0 represents a known function of x , y , and z .

Boundary conditions for the physical domain are somewhat more complicated. Mathematically, it can be shown that because groundwater flow equations contain second order partial derivatives with respect to the space coordinates, conditions are required at every point of the boundary of the physical flow domain, even when the boundaries tend to infinity. Because the interaction between the inside groundwater flow and the outside world can be complex, and our ability to conceive such phenomena accurately is rather limited, boundary conditions are usually described in a simplified way.

Generally, three types of boundary conditions are considered. First type boundary conditions apply when the value of the dependent variable at the boundary is known. For instance, in case of groundwater potentials, such a boundary condition would be

$$h(x_b, y_b, z_b, t) = h_b(t) \quad (4.63)$$

where (x_b, y_b, z_b) represents a point on the boundary, and h_b is a known function of time. Such an expression is called a potential boundary condition, and is used when the groundwater is in contact with a water body, having a known potential h_b , as for instance a river, lake or reservoir, and so on. The key assumption is that whatever groundwater flow occurs at the inside of the flow domain and at the boundary, it will have no influence on the potential of the outside water body, such that this potential remains fixed as stated by the boundary condition.

Second type boundary conditions are so-called flux boundary conditions, where it is assumed that the amount of groundwater exchange through the boundary is known. The amount of water exchange is given by the groundwater flux component perpendicular to the boundary, such that the boundary condition can be expressed as

$$q_n(x_b, y_b, z_b, t) = q_b(t) \quad (4.64)$$

where q_n represents the flux component normal to the boundary, and $q_b(t)$ is a known function of time. A flux type condition can be expressed in function of the dependent variable, the groundwater potential, by using Darcy's law

$$q_n = -K \frac{\partial h}{\partial n} = q_b(t) \quad (4.65)$$

where $\partial h / \partial n$ represents the derivative of the potential perpendicular to the boundary. From mathematical point of view, the choice of the positive direction of the normal to the boundary determines whether q_n -values are positive or negative. However, in practice, often a simpler convention is used, by considering

fluxes entering the flow domain as positive, and outgoing fluxes as negative, regardless of the sense of the normal on the boundary.

A flux boundary condition implies that whatever the state and flow of the groundwater inside the flow domain and at the boundary, the normal flux is fixed by external conditions, and remains as stated by the boundary condition. Typical examples of such conditions are pumping wells, groundwater recharge, infiltration, and the like; in general any situation where the flux is supposedly fixed and known. An obvious example is an impervious boundary, as the flux component perpendicular to the boundary is strictly zero.

A third type boundary condition is a mixture of the two previous types. It applies when potential and normal flux components at the boundary are related to each other. Such condition applies in case of a flow domain in contact with a water body with a fixed potential, h_b , but where the exchange between the groundwater reservoir and the water body is restricted, due to the presence of some resistance. Mathematically, this condition is expressed as

$$q_n(x_b, y_b, z_b, t) = C_b[h_b(t) - h(x_b, y_b, z_b, t)] \quad (4.66)$$

where we have used the above mentioned sign convention for the normal boundary flux, and C_b represents the boundary conductance, with dimensions $[T^{-1}]$, between the groundwater reservoir and the outside water source or sink. This parameter can be explained physically as the conductive capacity of a permeable boundary layer, present between the groundwater flow domain and the water body, such that

$$C_b = K_b/D_b \quad (4.67)$$

where K_b is the hydraulic conductivity of the boundary layer, and D_b its thickness. For instance, this situation would apply when a river is in contact with the groundwater, but the interaction is restricted due to the presence of a mud layer in the river bed. Another example is a well with a clogged filter.

One of the most complicated boundary conditions is the water table. Actually, the water table is not a true boundary, because groundwater flow can cross the water table when transfer of water occurs between the unsaturated and saturated zones. However, because unsaturated flow is usually difficult to solve, it is often neglected and the water table is considered as an upper boundary of the groundwater flow domain. If such a simplification is used, it becomes necessary to specify the water table as a boundary condition. When the position of the water table is known, this is not so complicated, because the pressure at the water table is atmospheric such that the relative pressure is zero, and the following potential boundary condition applies

$$h(x_w, y_w, z_w, t) = z_w \quad (4.68)$$

where (x_w, y_w, z_w) represents a point on the water table. However, in most cases the position of the water table is not known exactly, and in addition to Equation 4.68 another boundary condition is needed to solve the groundwater flow problem, which now also includes the determination of the shape and position of the water table. Obviously, this additional condition should state something about the amount of flow crossing the water table. Several possibilities exist, the most simple one being the case where a steady water table acts as an impervious upper boundary, that is, no flow crosses this surface. This can mathematically be expressed as follows. An implicit equation describing the position of the water table is $h - z = 0$. The gradient of this expression is a vector perpendicular to the water table, and hence, the scalar product with the groundwater flux is zero if no water is crossing the water table, hence

$$\mathbf{q} \cdot \nabla(h - z) = 0 \quad (4.69)$$

When this is worked out further and Darcy's law is used to express the flux, the following boundary condition is obtained.

$$K_h \left(\frac{\partial h}{\partial x} \right)^2 + K_h \left(\frac{\partial h}{\partial y} \right)^2 + K_v \left(\frac{\partial h}{\partial z} \right)^2 - K_v \frac{\partial h}{\partial z} = 0 \quad (4.70)$$

This equation applies to every point (x_w, y_w, z_w) of the water table. Notice, that it is nonlinear, which is the price to pay for neglecting the unsaturated flow (there is no such thing as a free lunch). More complicated equations are needed when the water table is moving in time, or when certain flows are crossing the water table as for instance a groundwater recharge flux. It is not possible to go into more details here. For the interested reader these cases are discussed by Bear (1972). However, in practice, these types of boundary conditions are not used very much, because they are too complicated; therefore it becomes impossible to solve the groundwater flow problem accurately. Hence, approximate techniques are used. For instance, in case of numerical simulation models, the position of the water table is determined in an iterative way. For steady state problems, first a water table position is chosen above the expected position, then the groundwater flow equation is solved with a flux type boundary condition as Equation 4.65, and afterwards a new water table position is determined by identifying the surface where $h = z$, after which the procedure can be restarted. For transient conditions, the position of the water table is tracked in time, by equating the inputs and outputs of groundwater at the water table and calculating the resulting changes in water table position.

4.3 Hydraulic Approach to Groundwater Flow

4.3.1 Concept

In general, groundwater flow is 3D, but due to the geometry of ground layers, and differences in hydraulic conductivities, actual groundwater flows tend to be concentrated in certain directions, as illustrated in Figure 4.3. The reason for this is that ground layers usually extend horizontally over large distances, while the vertical dimensions are rather restricted. Also, due to the large variety in conductive properties, ground layers can be grouped into three classes: pervious formations or aquifers, semi-pervious formations or aquitards, and impervious formations. As water flows through a porous medium along the path of least resistance, groundwater moves predominantly longitudinally in the pervious formations, and transversely in the semi-pervious layers. Hence, due to the horizontal layering, groundwater flow is essentially horizontal in aquifers, and vertical in aquitards, as shown schematically in Figure 4.3b. Therefore, instead of considering the flow as 3D a simplified description of groundwater flow is possible, when horizontal flow components are considered to be dominant in aquifers and vertical components in aquitards, as shown in Figure 4.3c. This concept is termed the hydraulic approach to groundwater flow, because the movement is considered to be directed as water flow in pipes or channels (Polubarnova-Kochina, 1962; Bear, 1972, 1979). In case of a phreatic aquifer, this is also called the Dupuit–Forchheimer theory, after the two scientists who pioneered this approach in the second half of the 19th century.

4.3.2 Motion Equations

To have a clear understanding of the basic concepts of the hydraulic approach, it is convenient to start with the equation of motion, before discussing the continuity equation. Consider a confined aquifer, bounded by a bottom surface, $b(x, y)$, and an upper surface, $s(x, y)$, such that its thickness is given by $D = s - b$, as shown in Figure 4.4a. When the groundwater potential is measured by a piezometer, the actual depth of the piezometer is unimportant because vertical potential differences are very small. Hence, h is almost independent of z , and the total horizontal volumetric flow rate of the groundwater over the

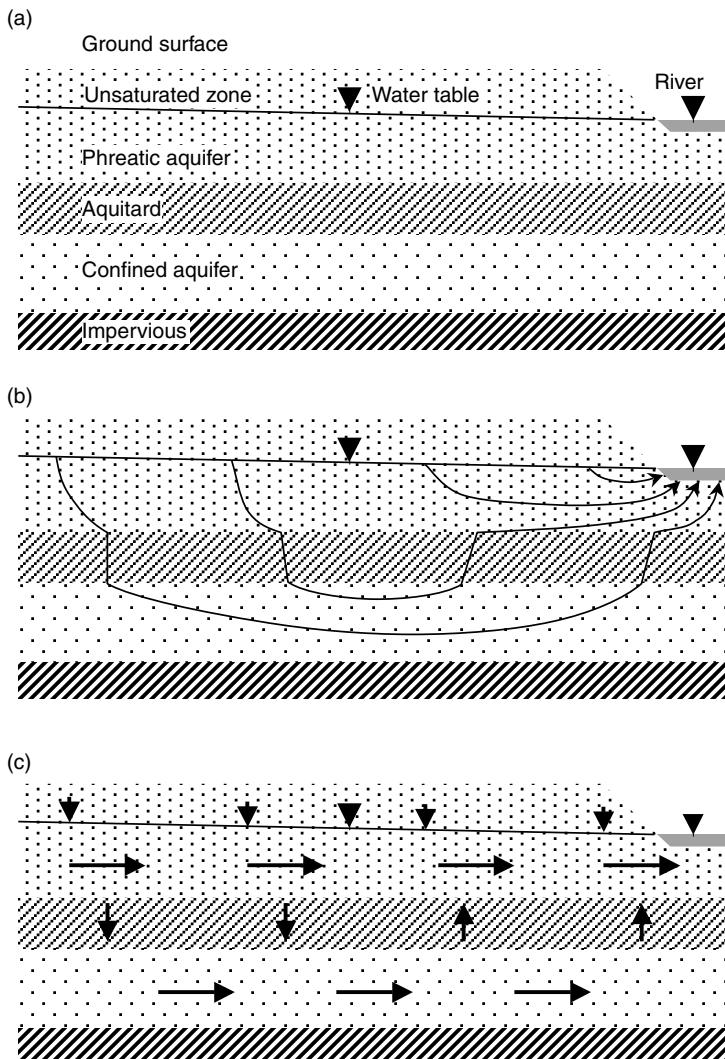


FIGURE 4.3 Occurrence and movement of groundwater: (a) types of hydrogeological layers, (b) actual groundwater flow paths, and (c) predominant groundwater flow directions.

entire thickness of the aquifer, in the x -direction, can be calculated as

$$Q_x = \int_b^s q_x dz = \int_b^s -K_h \frac{\partial h}{\partial x} dz \approx - \int_b^s K_h dz \frac{\partial h}{\partial x} = -T \frac{\partial h}{\partial x} \quad (4.71)$$

where Q_x is the x -component of the hydraulic groundwater flux, $\mathbf{Q} = (Q_x, Q_y)$, with dimensions volume per time per length [$L^3/TL = L^2/T$], and T is the transmissivity [L^2/T], defined as

$$T = \int_b^s K_h dz \quad (4.72)$$

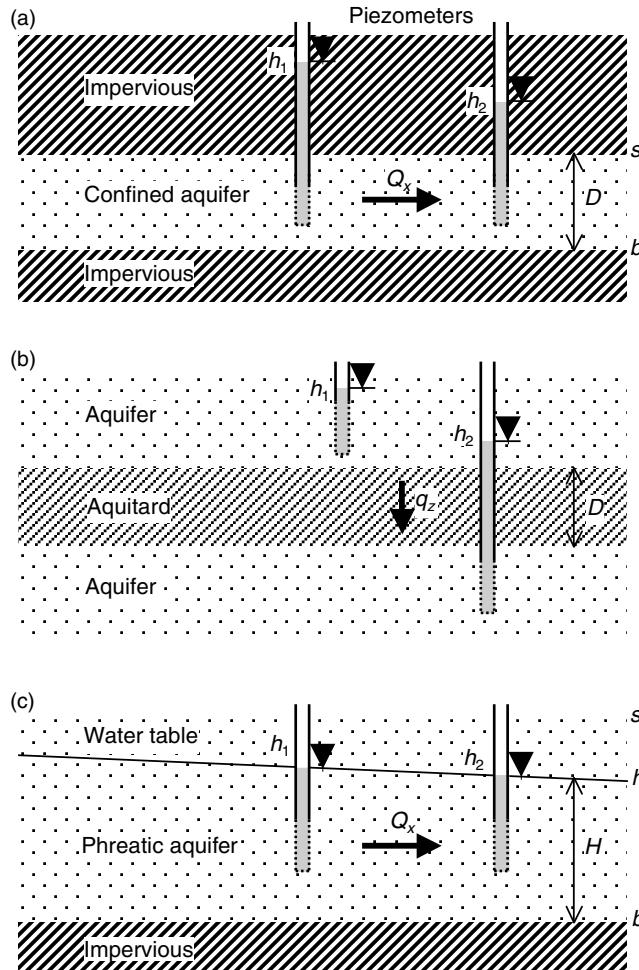


FIGURE 4.4 Hydraulic approach to groundwater flow: (a) horizontal flow in a confined aquifer, (b) vertical flow in an aquitard, and (c) horizontal flow in a phreatic aquifer.

In a similar way, the y -component of the hydraulic flux is given as

$$Q_y = -T \frac{\partial h}{\partial y} \quad (4.73)$$

The transmissivity is a parameter expressing the overall horizontal conductivity of a confined aquifer. Equation 4.72 and Equation 4.73 are extensions of Darcy's law, describing overall horizontal movement of groundwater in a confined aquifer.

In the case of an aquitard bounded by two aquifers, vertical flow can pass through the aquitard from one aquifer to the other, as depicted in Figure 4.4b. This vertical groundwater flux can be calculated by Darcy's law as

$$q_v = -K_v \frac{h_2 - h_1}{D} = C(h_1 - h_2) \quad (4.74)$$

where K_v is the vertical conductivity of the aquitard, D , its thickness, and the potential gradient is obtained from the difference in potential between the two aquifers bounding the aquitard. Parameter C ,

is the conductance, commonly called the leakage coefficient, and defined as

$$C = K_v/D \quad (4.75)$$

with dimensions $[T^{-1}]$; this coefficient expresses the leaking capacity of an aquitard, allowing vertical groundwater transfer between two aquifers bounding the aquitard. In case of a layered aquitard, the leakage coefficient is given by $C = [\int_b^s K_v^{-1} dz]^{-1}$. Hence, the leakage coefficient depends on the harmonic average of the conductivity over the aquitard thickness, while the transmissivity on the arithmetic average over the aquifer thickness.

The case of a phreatic aquifer is shown schematically in Figure 4.4c. The aquifer is bounded from below by its base, $b(x, y)$, and from above by the soil surface, $s(x, y)$. Actually, the true upper boundary of the phreatic aquifer is the water table. If a piezometer would be installed at a certain depth below the water table, the water level would reveal a potential, closer to the water table elevation, because vertical potential differences are insignificant as vertical groundwater flow is very small. Hence, h is nearly independent of z and almost coincides with the position of the water table, which is the upper surface of the saturated zone. Consequently, the total flow of groundwater passing horizontally through the aquifer in the x -direction, is given by

$$Q_x = \int_b^s q_x dz = \int_b^h q_x dz = \int_b^h -K_h \frac{\partial h}{\partial x} dz \approx - \int_b^h K_h dz \frac{\partial h}{\partial x} = -K_e(h - b) \frac{\partial h}{\partial x} \quad (4.76)$$

where horizontal flows in the unsaturated zone have been neglected, and K_e is an effective hydraulic conductivity given by

$$K_e = \frac{1}{h - b} \int_b^h K_h dz \quad (4.77)$$

In case of a homogeneous medium, K_e is equal to K_h . However, ground layers are generally heterogeneous, such that K_e needs to be determined by averaging; this is called upscaling or homogenization and requires careful consideration of the hydrogeological and stratigraphical characteristics of the subsurface and of the averaging procedure. Theoretical and mathematical aspects are discussed by, for example, Zijl and Nawalany (1993) and Rubin (2003).

An expression similar as Equation 4.76 can be obtained for the flow component in the y -direction

$$Q_y = -K_e(h - b) \frac{\partial h}{\partial y} \quad (4.78)$$

This is basically the approach of Dupuit, who postulated that groundwater flow in a phreatic aquifer is proportional to the slope of the groundwater table. Also, a transmissivity can be defined similar as for a confined aquifer, but this makes little sense, because the thickness of the phreatic aquifer changes with water table position, such that the transmissivity is not constant.

A simplification is possible when the base of the phreatic aquifer is horizontal. If the water table position is measured from this base, $H = h - b$, the hydraulic flow approximation can be written as

$$Q_x = -K_e H \frac{\partial H}{\partial x} = -K_e \frac{\partial}{\partial x} \left(\frac{H^2}{2} \right) \quad (4.79)$$

$$Q_y = -K_e H \frac{\partial H}{\partial y} = -K_e \frac{\partial}{\partial y} \left(\frac{H^2}{2} \right) \quad (4.80)$$

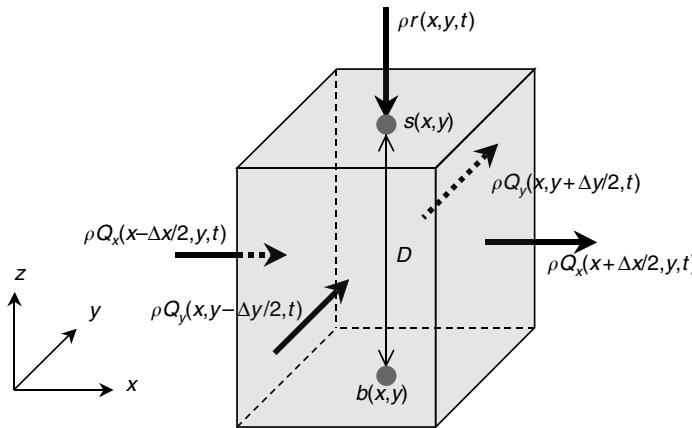


FIGURE 4.5 Mass conservation in a reference elementary volume in a confined aquifer.

Dupuit's equation is mostly known in this form. It replaces Darcy's law for describing overall horizontal groundwater flow in a phreatic aquifer.

4.3.3 Flow Equation for a Confined Aquifer

To obtain a flow equation, the equation of motion has to be combined with the continuity equation. The latter results from the principle of mass conservation applied to a control volume. Because, in the hydraulic approach the flow is integrated over the total depth of an aquifer, the control volume should be of a similar concept, that is, encompass the total vertical extent of the aquifer, as shown in Figure 4.5. Hence, the control box centered around a point (x, y) has sizes Δx and Δy in the horizontal plane, and extends from the bottom plane, $b(x, y)$, to the upper surface, $s(x, y)$, along the vertical, such that its vertical size is the aquifer thickness, D . The dimensions of the control volume reflect the hydraulic approach concept, that is, elementary sizes in the horizontal direction are of the same order as the thickness of the aquifer; hence, the aquifer is considered to be essentially a 2D object.

The groundwater mass inflow minus outflow in the control volume can be calculated as

$$[\rho Q_x(x - \Delta x/2, y, t) - \rho Q_x(x + \Delta x/2, y, t)]\Delta y + [\rho Q_y(x, y - \Delta y/2, t) - \rho Q_y(x, y + \Delta y/2, t)]\Delta x + \rho R(x, y, t)\Delta x\Delta y \quad (4.81)$$

where $R(x, y, t)$ [L/T] is a recharge or net exchange of groundwater passing through the base or through the upper surface of the aquifer. Actually, we consider here a semi-confined aquifer; in case of a truly confined aquifer R would be zero. The sign convention for R is that inputs are considered positive, irrespective of the direction along the z -axis.

Using Taylor series expansions, the total net inflow in the control box, becomes

$$\left(-\frac{\partial \rho Q_x}{\partial x} - \frac{\partial \rho Q_y}{\partial y} + \rho R \right) \Delta x \Delta y \quad (4.82)$$

This should be balanced by a change in storage. The amount of water mass, M , present in the control volume, is given by

$$M = \int_b^s \rho n dz \Delta x \Delta y = \rho n D \Delta x \Delta y \quad (4.83)$$

where in the last term we have considered average porosity and density values over the thickness of the aquifer. The change in storage can be calculated as

$$\frac{\partial M}{\partial t} = \left(\frac{\partial \rho}{\partial t} nD + n \frac{\partial n}{\partial t} D + \rho n \frac{\partial D}{\partial t} \right) \Delta x \Delta y \quad (4.84)$$

where compression of the water and aquifer is considered, and vertical deformation of the aquifer. Using the relationships for expressing the compressibility of water and porous medium and assuming that the solid grains are incompressible, this can be worked out as

$$\frac{\partial M}{\partial t} = \rho D(\alpha + n\beta) \frac{\partial p}{\partial t} \Delta x \Delta y = \rho^2 g D(\alpha + n\beta) \frac{\partial h}{\partial t} \Delta x \Delta y \quad (4.85)$$

Making use of the definition of specific storage coefficient, Equation 4.57, this can also be written as

$$\frac{\partial M}{\partial t} = \rho S_s D \frac{\partial h}{\partial t} \Delta x \Delta y = \rho S \frac{\partial h}{\partial t} \Delta x \Delta y \quad (4.86)$$

where S is the aquifer storage coefficient or storativity, defined as

$$S = S_s D = \rho g (\alpha + n\beta) D \quad (4.87)$$

which is a measure of the overall aquifer storage properties. The storativity is dimensionless. It can be defined as the volume of groundwater released by a confined aquifer per unit horizontal aquifer surface and per unit decline of groundwater potential, all in accordance with the hydraulic approach, where variables are considered averaged or integrated over the vertical dimensions of the aquifer.

Combining Equation 4.80 and Equation 4.84, dividing by $\rho \Delta x \Delta y$, and neglecting density gradients, results in

$$S \frac{\partial h}{\partial t} = - \frac{\partial Q_x}{\partial x} - \frac{\partial Q_y}{\partial y} + R \quad (4.88)$$

Using the expression for horizontal hydraulic flow given by Equation 4.71 and Equation 4.73, results in the following hydraulic groundwater flow equation for a (semi-)confined aquifer

$$S \frac{\partial h}{\partial t} = \frac{\partial}{\partial x} \left(T \frac{\partial h}{\partial x} \right) + \frac{\partial}{\partial y} \left(T \frac{\partial h}{\partial y} \right) + R \quad (4.89)$$

Compared to saturated 3D groundwater flow, Equation 4.58, it can be noted that this equation contains no z -dimension, which is the main simplification resulting from the hydraulic approach. Also, as a consequence, the needed porous medium properties are transmissivity and storativity, which are aquifer properties integrated over the vertical dimension. Therefore, Equation 4.89 is very useful in practice, because it is simpler and needs less detailed knowledge about the medium properties. Care should be taken when applying this equation to practical situations, because boundary conditions are required in accordance with the hydraulic approach. This implies that boundary conditions should be independent of the elevation, that is, they should apply to the total aquifer thickness. For instance, pumping by a fully penetrating well can be considered with the hydraulic approach, but not a partially penetrating well because this induces vertical potential gradients in the aquifer.

4.3.4 Flow Equation for a Phreatic Aquifer

A flow equation for a phreatic aquifer based on the hydraulic approach can be obtained as for a confined aquifer, but some complications occur. The control volume is similar; it is centered around the point (x, y) ,

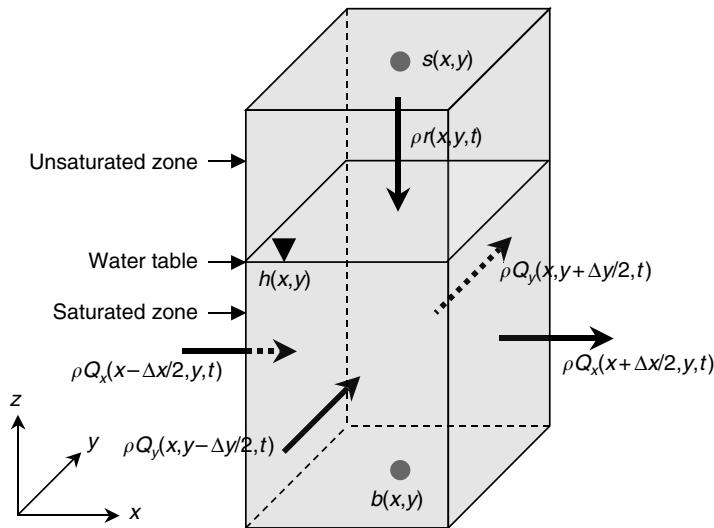


FIGURE 4.6 Mass conservation in a reference elementary volume in a phreatic aquifer.

with sizes Δx and Δy horizontally, and extends from the base to the soil surface, as shown in Figure 4.6. Notice that the control volume includes the unsaturated zone, because the groundwater flow and in particular the resulting changes in the water table elevation will also have effects on the water balance in the unsaturated zone, which have to be taken into account.

The inflow minus the outflow depends upon the hydraulic horizontal flow components in the saturated zone and possible vertical inputs or outputs through the upper or lower boundary. For the latter, we can think of water exchanges at the soil surface as infiltration and evapotranspiration. Actually, the true input from above in the phreatic aquifer occurs at the water table, and is also the result of processes occurring in the unsaturated zone. As groundwater flow is a slow process, it is reasonable to consider an average flux over long periods of time for the recharge, and if we consider no horizontal flow in the unsaturated zone (often a good approximation), the average flux at the upper boundary equals the average flux at passing through the water table. Hence, the water balance can be equated in a control box with sizes Δx and Δy in the horizontal plane, and extending from the bottom plane, $b(x, y)$, to the top surface, $s(x, y)$, similar as in Equation 4.81 and when developed further by Taylor series expansions the same expression as Equation 4.82 is obtained. Of course, in addition to recharge at the water table, transfer could also take place through the base of the phreatic aquifer if an aquitard is present connected to an underlying aquifer, in which case another R -term should be added to the equation.

Somewhat more complicated than in case of a confined aquifer is the calculation of the change in storage in a phreatic aquifer, because the position of the water table water table is not fixed in time, and part of the saturated zone can become unsaturated, or vice versa. The total mass of water present can be calculated by integration along the vertical

$$M = \Delta x \Delta y \int_b^s \rho \theta dz \quad (4.90)$$

where θ is the volumetric water content [L^3/L^3]; the water content is equal to the porosity in the saturated part below the water table, and is either smaller than or equal to the porosity in the unsaturated zone above the water table.

The change in storage is given by

$$\frac{\partial M}{\partial t} = \Delta x \Delta y \frac{\partial}{\partial t} \int_b^s \rho \theta dz \quad (4.91)$$

Now, it becomes clear why the control volume extends to the soil surface, as changes in storage can be significant when there are differences in water content especially in the unsaturated zone. In presence of such changes, storage effects due to compressibility of the water or porous medium become insignificant and can be neglected. It follows that the change in storage can be evaluated as

$$\frac{\partial M}{\partial t} \approx \rho \Delta x \Delta y \int_b^s \frac{\partial \theta}{\partial t} dz \quad (4.92)$$

The exact evaluation of the integral is not simple. The situation is shown schematically in Figure 4.7. If the water table drops by an amount Δh , the change in storage depends upon the moisture distribution above the water table, before, and after the drop in water table position. Actually, time will also have an influence because the water in the unsaturated zone needs time to percolate downwards to the water table. Hence, the exact change in storage is difficult to calculate, unless unsaturated flow is taken into consideration, which is exactly what one would like to avoid by applying the hydraulic approach. Hence, the storage needs to be approximated in a more simple way. This can be achieved as follows. Consider the situation depicted in Figure 4.7a, where the soil surface is high above the water table. Under normal average conditions, the distribution of the water content will vary with elevation above the water table. It will equal the porosity at the water table, and generally will decrease with height above the water table, gradually approaching a residual water content, which is characteristic for the unsaturated zone far above the water table. This residual water content is denoted by θ_r . Now, if the position of water table is changed by an amount Δh , and given sufficient time, the moisture distribution will eventually be similar as before, by

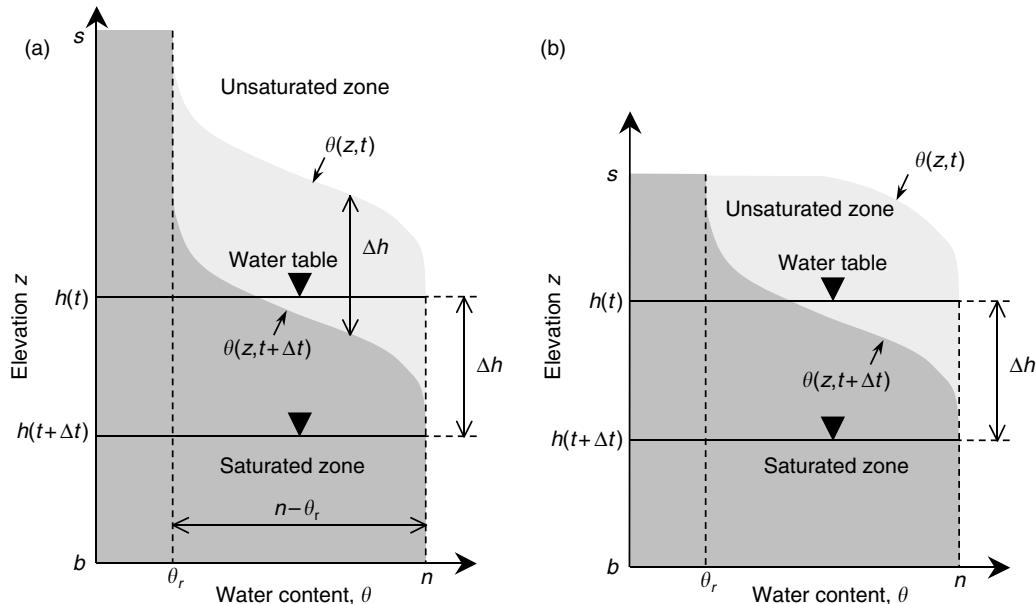


FIGURE 4.7 Storage and release of groundwater in a phreatic aquifer due to changes in water table position: (a) deep water table, and (b) shallow water table.

shifting over a distance Δh . Hence, the volumetric difference in storage can be calculated as

$$\int_b^a \frac{\partial \theta}{\partial t} dz = \int_b^a \lim_{\Delta t \rightarrow 0} \left[\frac{\theta(z, t + \Delta t) - \theta(z, t)}{\Delta t} \right] dz = \lim_{\Delta t \rightarrow 0} \left[(n - \theta_r) \frac{\Delta h}{\Delta t} \right] = (n - \theta_r) \frac{\partial h}{\partial t} = S_y \frac{\partial h}{\partial t} \quad (4.93)$$

where $S_y = n - \theta_r$ [−] is defined as the specific yield, that is, the amount of water released per volume of porous medium when changing from saturated state to unsaturated state high above the water table. In reality, things are not that simple. First, it could be possible that the soil surface is not situated far above the water table, as shown in Figure 4.7b, such that the mathematical derivation given above is not valid. Second, retardation effects will occur, because conductivities in the unsaturated zone are generally much smaller than in the saturated zone, and, consequently, water movement in the unsaturated zone lags behind saturated groundwater flow. Third, soil heterogeneity complicates the concept of residual water content and equilibrium moisture distributions in the unsaturated zone. Hence, in practice the change in storage will usually be smaller. Therefore, Equation 4.93 remains only useful when the concept of specific yield is extended to match any particular situation. This gross simplification is justified in view of the fact that with the hydraulic approach groundwater flow is described in a simplified way. However, the concept of specific yield has a clear physical interpretation, similar as storativity, that is, the volume of water released by a phreatic aquifer per unit horizontal aquifer surface and per unit decline of the water table. Of course, the position of the water table, the speed of water table decline, and the soil heterogeneity play an important role in this.

When we combine the obtained storage equation with the net inflow, Equation 4.82, divide by $\rho \Delta x \Delta y$, and use the hydraulic flow approximation, given by Equation 4.76 and Equation 4.78, the following hydraulic flow equation for a phreatic aquifer is obtained

$$S_y \frac{\partial h}{\partial t} = \frac{\partial}{\partial x} \left[K_e (h - b) \frac{\partial h}{\partial x} \right] + \frac{\partial}{\partial y} \left[K_e (h - b) \frac{\partial h}{\partial y} \right] + R \quad (4.94)$$

Also, in this equation, there are no terms with respect to the z -dimension, which is again the main advantage of the hydraulic approach. However, the equation is nonlinear, which complicates practical applications and finding exact analytical solutions. Nevertheless, the equation is extensively used, especially in numerical groundwater flow simulation models.

A simplification is possible for phreatic aquifers with a horizontal base, using the elevation of the water table above the base, $H = h - b$, as dependent variable; the flow equation becomes

$$S_y \frac{\partial H}{\partial t} = \frac{\partial}{\partial x} \left[K_e H \frac{\partial H}{\partial x} \right] + \frac{\partial}{\partial y} \left[K_e H \frac{\partial H}{\partial y} \right] + R \quad (4.95)$$

This is called the Boussinesq equation. The equation is still non-linear, but in case of steady state conditions, becomes linear in $H^2/2$

$$\frac{\partial}{\partial x} \left[K_e \frac{\partial}{\partial x} \left(\frac{H^2}{2} \right) \right] + \frac{\partial}{\partial y} \left[K_e \frac{\partial}{\partial y} \left(\frac{H^2}{2} \right) \right] + R = 0 \quad (4.96)$$

Several useful solutions have been derived with this equation (Polubarnova-Kochina, 1962; Bear, 1972, 1979). However, care should be taken with respect to boundary conditions, because these have to respect the hydraulic approach and consequently be independent of elevation. Consider for instance the very common boundary condition: a river in contact with a phreatic aquifer. The water level in the river can only be used as a potential boundary condition when the river is wide and deep such that the groundwater underneath the river is effectively having the same potential. If the surface water body is narrow and shallow, as in drains or ditches, there can be significant vertical flow components resulting in a potential

difference between the surface water and the groundwater underneath. In practice, such a situation can be treated with a third type boundary condition, where a boundary resistance is introduced that takes into account the effect of converging groundwater flow towards the drain or ditch.

4.3.5 Flow Equation for an Aquitard

To obtain a flow equation for an aquitard, it is sufficient to note that the flow is considered to be strictly vertical, such that the flow equation can be derived from the 3D saturated groundwater flow, Equation 4.56, by only considering the z -component of the flow

$$S_s \frac{\partial h}{\partial t} = \frac{\partial}{\partial z} \left(K_v \frac{\partial h}{\partial z} \right) \quad (4.97)$$

This equation describes flow through an aquitard, while the potentials of the aquifers bounding the aquitard can be used as boundary conditions. However, in practice this equation is not used much. Usually, an additional simplifying assumption is made, namely that the specific storage coefficient is negligibly small, such that

$$\frac{\partial}{\partial z} \left(K_v \frac{\partial h}{\partial z} \right) = 0 \quad (4.98)$$

This means that the flow is considered to be in quasi steady state, and the solution is given by Equation 4.74, that is, the flow adjusts immediately to changes in the aquifer potentials above and below the aquitard. This simplification makes vertical flows through aquitards easy to evaluate, such that these can be incorporated as inputs or outputs in above or underlying aquifers. However, the approach is not very accurate, because conductivities of aquitards are much smaller than for aquifers, as a result of which, in general, the flow through an aquitard lags behind the flow in the bounding aquifers, and time and storage effects can be important (Frind, 1983). However, lack of knowledge of aquitard properties often forces us to use a simplified approach.

4.4 Examples of Solutions

4.4.1 Dupuit Discharge Formula

Consider steady flow through a pervious homogeneous and isotropic dam with vertical walls and impervious horizontal base, as shown in Figure 4.8. The dam separates two reservoirs, so that the flow will be directed from the reservoir with the highest water level on the left to the reservoir with the lower level on the right, and bounded vertically by the impervious base and the water table. In addition, there will be a seepage face at the right-end side above the reservoir where groundwater seeps out freely in the atmosphere under zero pressure.

The flow equation, given by Equation 4.61, reduces to

$$\frac{\partial^2 h}{\partial x^2} + \frac{\partial^2 h}{\partial z^2} = 0 \quad (4.99)$$

If the groundwater potential is measured from the impervious base, the boundary conditions can be written as follows: for the left boundary

$$h(0, z) = h_1 \quad (4.100)$$

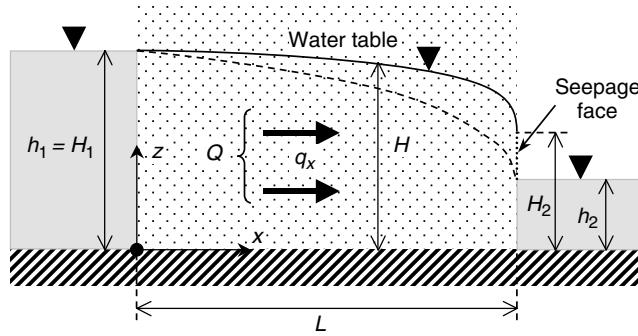


FIGURE 4.8 Steady state flow through a porous dam with vertical walls, situated between two reservoirs with constant water levels; the flow is vertically contained between the impervious base and water table and partially exists in the atmosphere through a seepage face; the dashed line represents the approximate position of the water table as obtained from the Dupuit equation.

where h_1 is the water depth in the left reservoir, for the right boundary at the contact between dam and reservoir

$$h(L, z \leq h_2) = h_2 \quad (4.101)$$

where h_2 is the water depth in the right reservoir, and L the width of the dam, while for the seepage face above the reservoir

$$h(L, z > h_2) = z \quad (4.102)$$

and for the impervious base

$$\left. \frac{\partial h}{\partial z} \right|_{z=0} = 0 \quad (4.103)$$

and finally for the water table Equation 4.70, which here becomes

$$\left(\frac{\partial h}{\partial x} \right)^2 + \left(\frac{\partial h}{\partial z} \right)^2 - \frac{\partial h}{\partial z} = 0 \quad (4.104)$$

An exact solution is very difficult to obtain because the position of the water table is unknown and the condition given by Equation 4.104 is nonlinear. A solution is given by Polubarinova-Kochina (1962) but is very complicated and hard to apply in practice. Hence, a simplified approach using the hydraulic approximation seems more appropriate. In this case Equation 4.96 applies, which reduces to

$$\frac{\partial^2 H^2}{\partial x^2} = 0 \quad (4.105)$$

where H is the average potential and, at the same time, the water table position measured from the base. The solution is

$$H^2 = C_1 x + C_2 \quad (4.106)$$

with C_1 and C_2 , integration constants, that have to be determined from the boundary conditions. Because the water table starts at h_1 , the water level of the left reservoir at $x = 0$, it follows that $C_2 = h_1^2$. At the right

end side, there is a problem, because we do not know the position of the water table due to the presence of the seepage face. Hence, we are forced to ignore the seepage face, and impose $H = h_2$ for $x = L$, which enables to determine C_1 , resulting in the following solution

$$H^2 = \frac{1}{L}(h_2^2 - h_1^2)x + h_1^2 \quad (4.107)$$

This equation implies that the water table has a parabolic shape, as depicted by the dashed line in Figure 4.8, which obviously deviates from reality (although differences will be small if L is large compared to h_1). The amount of flow Q , from the left reservoir to the right reservoir can be calculated with Equation 4.79, or

$$Q = -K \frac{\partial}{\partial x} \left(\frac{H^2}{2} \right) \quad (4.108)$$

Calculating the derivative using Equation 4.107 results in

$$Q = K \frac{h_1^2 - h_2^2}{2L} \quad (4.109)$$

This is the famous Dupuit equation, which is extensively used in practice to calculate ground water flow in phreatic aquifers. One would be tempted to question its validity or accuracy in view of the horizontal flow assumption. Hence, we return to the 2D flow problem, given by Equation 4.99 to Equation 4.103, and calculate the flow from by integrating the horizontal flux component q_x along a vertical from the base of the dam to the water table

$$Q = \int_0^H q_x dz = -K \int_0^H \frac{\partial h}{\partial x} dz \quad (4.110)$$

When the integration and derivative is interchanged, this becomes

$$Q = -K \frac{\partial}{\partial x} \left(\int_0^H h dz - \frac{1}{2} H^2 \right) \quad (4.111)$$

Integration with respect to x gives

$$xQ = -K \left(\int_0^H h dz - \frac{1}{2} H^2 \right) + C_3 \quad (4.112)$$

The integration constant, C_3 , is determined by equating terms at $x = 0$, for which we make use of Equation 4.100, so that

$$C_3 = K \left(\int_0^{h_1} h dz - \frac{1}{2} h_1^2 \right) = K \left(h_1^2 - \frac{1}{2} h_1^2 \right) = \frac{1}{2} K h_1^2 \quad (4.113)$$

Equation 4.112 now becomes

$$xQ = \frac{1}{2} K h_1^2 - K \left(\int_0^H h dz - \frac{1}{2} H^2 \right) \quad (4.114)$$

Next, all terms are equated at $x = L$, for which we use Equation 4.101 and Equation 4.102, yielding

$$\begin{aligned} LQ &= \frac{1}{2}Kh_1^2 - K \left(\int_0^{h_2} h_2 dz + \int_{h_2}^{H_2} zdz - \frac{1}{2}H_2^2 \right) \\ &= \frac{1}{2}Kh_1^2 - K \left(h_2^2 + \frac{1}{2}H_2^2 - \frac{1}{2}h_2^2 - \frac{1}{2}H_2^2 \right) = \frac{1}{2}K(h_1^2 - h_2^2) \end{aligned} \quad (4.115)$$

Surprisingly, the final result does not depend upon H_2 , although the seepage face contributes to the total flow, and we obtain back Equation 4.109. This remarkable finding is due to Charny (1951). Hence, it appears that although the shape of the water table is not correctly predicted by the Dupuit approach, the equation for the flux is nevertheless exact. Obviously the errors introduced by the parabolic shape of the water table and the disregard of the seepage face cancel out.

4.4.2 Hvorslev Slug Test Equation

Consider a piezometer, with a screen of length L and radius R , installed in an aquifer, as shown in Figure 4.9a. We assume that the screen is small compared to the aquifer dimensions, so that seen from the piezometer screen the aquifer can be considered infinite in all directions. The groundwater potential h_0 is revealed by the water level in the piezometer. We will assume that this potential is present everywhere in the aquifer. The question arises as to what will be the groundwater inflow through the screen into the piezometer if the water level inside the piezometer is artificially lowered by an amount, Δh . No exact solution has been found for this problem. Hence, some simplifications are needed. First, consider steady flow to a point sink as shown in Figure 4.9b. Using spherical coordinates and because of spherical symmetry, the flux directed to the sink q_r at a distance r can be found by dividing the total flow Q by the surface of a sphere with radius, r , or

$$q_r = -K \frac{\partial h}{\partial r} = \frac{-Q}{4\pi r^2} \quad (4.116)$$

which enables to find h , as function of r

$$h = \frac{-Q}{4\pi Kr} + C_1 \quad (4.117)$$

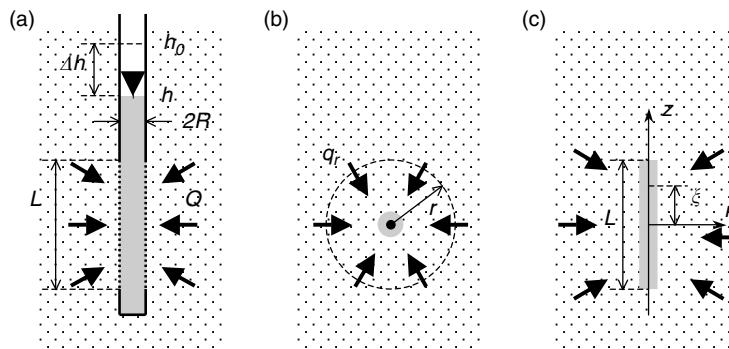


FIGURE 4.9 Groundwater flow through a piezometer screen when the water level in the piezometer is artificially lowered: (a) actual situation, (b) approximation by a point sink, and (c) approximation by a line sink.

The integration constant can be obtained by imposing that far from the sink, $r \rightarrow \infty$, the potential remains unchanged, that is, $h = h_0$, yielding

$$h = h_0 - \frac{Q}{4\pi Kr} \quad (4.118)$$

or

$$\Delta h = h_0 - h = \frac{Q}{4\pi Kr} \quad (4.119)$$

As a point sink is not a good representation of a piezometer screen, we now consider a line source sink, as shown in Figure 4.9c. The line sink has a length L and extracts a total flow Q , to mimic the piezometer screen as well as possible. To find the potentials, we use cylindrical coordinates and consider the line sink to be composed of an infinite series of point sinks distributed along z -axis. We use for this an additional coordinate ξ which runs from $-L/2$ to $L/2$ as shown in Figure 4.9c. Hence, each point sink extracts an elementary flow $Qd\xi/L$ and the distance from any point to a sink is given by $\sqrt{(z - \xi)^2 + r^2}$. The resulting potential distribution can be obtained by summing up all the contributions of each point sink, that is,

$$\Delta h = \int dh = \int_{-L/2}^{L/2} \frac{Q}{4\pi KL\sqrt{(z - \xi)^2 + r^2}} d\xi = \frac{Q}{4\pi KL} \left[\sinh^{-1} \left(\frac{L/2 + z}{r} \right) + \sinh^{-1} \left(\frac{L/2 - z}{r} \right) \right] \quad (4.120)$$

Finally, the potential drop in the piezometer is approximated by equating the above formula for $z = 0$ and $r = R$ yielding

$$\Delta h = \frac{Q}{2\pi KL} \sinh^{-1} \left(\frac{L}{2R} \right) = \frac{Q}{2\pi KL} \ln \left[\frac{L}{2R} + \sqrt{1 + \left(\frac{L}{2R} \right)^2} \right] \quad (4.121)$$

which for $L/R > 8$ can be approximated as

$$\Delta h = \frac{Q}{2\pi KL} \ln \left(\frac{L}{2R} \right) \quad (4.122)$$

This equation was presented by Hvorslev (1951) and forms the basis for the so-called piezometer or slug test, to determine the hydraulic conductivity in the field. As explained by Hvorslev (1951) and Freeze and Cherry (1979), the potential difference and the flow rate can be obtained by measuring the change in water level inside the piezometer, and used to calculate the hydraulic conductivity.

4.4.3 Theis Well Flow Equation

Consider a fully penetrating well in a homogeneous and perfect horizontal, and infinite in radial, extend confined aquifer, as shown in Figure 4.10. Initially the groundwater potential is constant and equal to h_0 . At time $t = 0$, pumping starts with a constant rate Q . The problem is described by the 3D groundwater flow equation, given by Equation 4.56. As the problem is axial-symmetric, it can best be described in cylindrical coordinates (r, φ, z) , as in Equation 4.59, but the term in φ disappears because of the axial symmetry and in z because the flow to the well is strictly horizontal; hence the flow equation reduces to

$$S_s \frac{\partial h}{\partial t} = \frac{1}{r} \frac{\partial}{\partial r} \left(r K_h \frac{\partial h}{\partial r} \right) \quad (4.123)$$

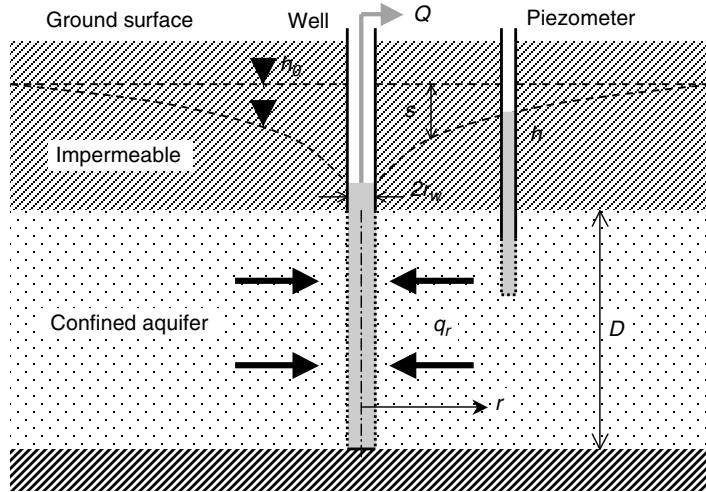


FIGURE 4.10 Groundwater flow in a confined aquifer to a fully penetrating well.

The **initial condition** is given by

$$h(r, 0) = h_0 \quad (4.124)$$

and the **boundary conditions** are: at infinity the given initial potential

$$h(\infty, t) = h_0 \quad (4.125)$$

and at the well screen the given radial flux

$$q_r(r_w, t) = -K_h \frac{\partial h}{\partial r} \Big|_{r=r_w} = -\frac{Q}{2\pi r_w D} \quad (4.126)$$

where r_w is the radius of the well screen. Notice that the term on the right-hand side is negative, because the flow is directed out of the aquifer. However, the boundary condition at the well is rather complicated; hence, this will be simplified by assuming that the radius of the well tends to zero, or

$$\lim_{r \rightarrow 0} \left(r \frac{\partial h}{\partial r} \right) = \frac{Q}{2\pi D K_h} \quad (4.127)$$

The solution is obtained by means of the Boltzmann transform $h(\lambda) = h(r, t)$, with $\lambda = r^2/t$. The transformation formulae are

$$\frac{\partial h}{\partial t} = \frac{dh}{d\lambda} \frac{\partial \lambda}{\partial t} = -\frac{r^2}{t^2} \frac{dh}{d\lambda} = -\frac{\lambda}{t} \frac{dh}{d\lambda} \quad (4.128)$$

$$r \frac{\partial h}{\partial r} = r \frac{dh}{d\lambda} \frac{\partial \lambda}{\partial r} = \frac{2r^2}{t} \frac{dh}{d\lambda} = 2\lambda \frac{dh}{d\lambda} \quad (4.129)$$

$$\frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial h}{\partial r} \right) = \frac{1}{r} \frac{\partial}{\partial r} \left(2\lambda \frac{dh}{d\lambda} \right) = \frac{1}{r} \frac{d}{d\lambda} \left(2\lambda \frac{dh}{d\lambda} \right) \frac{\partial \lambda}{\partial r} = \frac{4}{t} \frac{d}{d\lambda} \left(\lambda \frac{dh}{d\lambda} \right) \quad (4.130)$$

so that Equation 4.123 is transformed in

$$4K_h \frac{d}{d\lambda} \left(\lambda \frac{dh}{d\lambda} \right) + S_s \left(\lambda \frac{dh}{d\lambda} \right) = 0 \quad (4.131)$$

The boundary conditions for Equation 4.124 and Equation 4.125 become

$$\lim_{\lambda \rightarrow \infty} (h) = h_0 \quad (4.132)$$

and for Equation 4.127

$$\lim_{\lambda \rightarrow 0} \left(\lambda \frac{dh}{d\lambda} \right) = \frac{Q}{4\pi D K_h} \quad (4.133)$$

Equation 4.131 is a first order differential equation in $\lambda dh/d\lambda$, which can be solved as

$$\lambda \frac{dh}{d\lambda} = C_1 \exp \left(-\frac{S_s \lambda}{4K_h} \right) \quad (4.134)$$

where C_1 is an integration constant that can be determined with Equation 4.132, yielding

$$\lambda \frac{dh}{d\lambda} = \frac{Q}{4\pi D K_h} \exp \left(-\frac{S_s \lambda}{4K_h} \right) \quad (4.135)$$

This can be solved for h , as

$$h = \frac{Q}{4\pi D K_h} \int \lambda^{-1} \exp \left(-\frac{S_s \lambda}{4K_h} \right) d\lambda + C_2 \quad (4.136)$$

where C_2 is a second integration constant that can be determined with Equation 4.132, giving

$$h = h_0 - \frac{Q}{4\pi D K_h} \int_{r^2/t}^{\infty} \lambda^{-1} \exp \left(-\frac{S_s \lambda}{4K_h} \right) d\lambda \quad (4.137)$$

where we also have back-transformed $\lambda = r^2/t$. This equation was first derived by Theis (1935) through analogy with the flow and storage of heat in solids. It is usually written using the drawdown $s = h_0 - h$ [L], and overall aquifer properties transmissivity, T , and storativity, S , as

$$s = \frac{Q}{4\pi T} W \left(\frac{Sr^2}{4Tt} \right) \quad (4.138)$$

with the Theis well function, $W(u)$, given by

$$W(u) = \int_u^{\infty} \lambda^{-1} e^{-\lambda} d\lambda \quad (4.139)$$

The Theis well flow equation is probably the most commonly used and fundamental equation in groundwater flow, especially for the design of wells and the analysis of pumping tests.

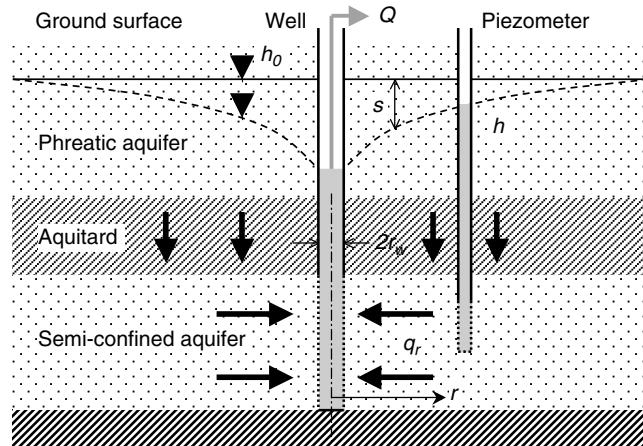


FIGURE 4.11 Groundwater flow in a semi-confined aquifer to a pumped, fully penetrating well and induced leakage through the overlying aquitard.

4.4.4 Hantush Leaky Well Flow Equation

Consider a fully penetrating well in a homogeneous and perfect horizontal semi-confined or leaky aquifer, as shown in Figure 4.11. The setting and conditions are very similar as for the confined aquifer considered in the previous case, but an essential difference is that the pumping of the well in the semi-confined aquifer will induce leakage through the aquitard. It is assumed that this water is coming from an aquifer on top, which can yield sufficient water without being depleted. Hence, it is assumed that initially the groundwater potential in both aquifers is constant and equal to h_0 , and that even when the well is pumping the potential in the top aquifer will remain unchanged.

The leakage through the aquitard will induce vertical flows in the semi-confined aquifer, resulting in a rather complex flow pattern. In order to simplify the problem, we will use the hydraulic approach for flow in the semi-confined aquifer, given by Equation 4.89, and for leakage through the aquitard, given by Equation 4.74. Because of axial symmetry, the flow equation becomes

$$S \frac{\partial h}{\partial t} = \frac{1}{r} \frac{\partial}{\partial r} \left(rT \frac{\partial h}{\partial r} \right) + C(h_0 - h) \quad (4.140)$$

The last term in this equation represents the leakage through the aquitard. The boundary conditions are similar as for the previous case, that is, Equations 4.124, 4.125, and 4.127. The flow equation can also be expressed in function of the drawdown, $s = h_0 - h$, as follows

$$S \frac{\partial s}{\partial t} = \frac{1}{r} \frac{\partial}{\partial r} \left(rT \frac{\partial s}{\partial r} \right) - Cs = T \frac{\partial^2 s}{\partial r^2} + \frac{T}{r} \frac{\partial s}{\partial r} - Cs \quad (4.141)$$

with initial condition

$$s(r, 0) = 0 \quad (4.142)$$

and boundary conditions

$$s(\infty, t) = 0 \quad (4.143)$$

$$\lim_{r \rightarrow 0} \left(r \frac{\partial s}{\partial r} \right) = -\frac{Q}{2\pi T} \quad (4.144)$$

The solution is obtained with the *Laplace transform*; let $\tilde{s}(r, p)$ be the Laplace transform of $s(r, t)$, where the tilde indicates the Laplace transform and p is the transform parameter. After transformation Equation 4.141 becomes

$$\frac{\partial^2 \tilde{s}}{\partial r^2} + \frac{1}{r} \frac{\partial \tilde{s}}{\partial r} - \frac{Sp + C}{T} \tilde{s} = 0 \quad (4.145)$$

where we have made use of the initial condition given by Equation 4.142. Equation 4.145 is the modified *Bessel equation* of order zero with general solution

$$\tilde{s} = C_1 I_0 \left(r \sqrt{\frac{Sp + C}{T}} \right) + C_2 K_0 \left(r \sqrt{\frac{Sp + C}{T}} \right) \quad (4.146)$$

where I_0 is the modified *Bessel function* of the first kind of order zero and K_0 the modified *Bessel function* of the second kind of order zero, and C_1 and C_2 are arbitrary constants, which have to be determined from the boundary conditions. The Laplace transform of Equation 4.143 is

$$\tilde{s}(\infty, p) = 0 \quad (4.147)$$

which implies that C_1 should be zero, since $I_0(\infty) = \infty$ and $K_0(\infty) = 0$. The Laplace transform of Equation 4.144 is

$$\lim_{r \rightarrow 0} \left(r \frac{\partial \tilde{s}}{\partial r} \right) = -\frac{Q}{2\pi T p} \quad (4.148)$$

and as the derivative of $K_0(x)$ is $K_1(x)$, the modified Bessel function of the second kind of order one, which tends to x^{-1} for $x \rightarrow 0$, the second constant can be determined and substituted in Equation 4.146, yielding

$$\tilde{s} = \frac{Q}{2\pi T p} K_0 \left(r \sqrt{\frac{Sp + C}{T}} \right) \quad (4.149)$$

Finally, making use of properties and tables of the Laplace transform (Abramowitz and Stegun, 1970; Zwillinger, 2002) to back-transform the solution, results in

$$s = \frac{Q}{4\pi T} \int_0^t \tau^{-1} \exp \left(-\frac{Sr^2}{4T\tau} - \frac{C\tau}{S} \right) d\tau \quad (4.150)$$

After substituting $\lambda = Sr^2/4T\tau$ as integration variable, this becomes

$$s = \frac{Q}{4\pi T} W \left(\frac{Sr^2}{4Tt}, \frac{r}{B} \right) \quad (4.151)$$

where $B = \sqrt{T/C}$ is the leakage factor [L], and $W(u, \beta)$ is the Hantush leaky well function, given by

$$W(u, \beta) = \int_u^\infty \lambda^{-1} \exp \left(-\lambda - \frac{\beta^2}{4\lambda} \right) d\lambda \quad (4.152)$$

This equation was first developed by Hantush and Jacob (1955), and is extensively used for design and analysis of wells in semi-confined aquifers. Although it resembles very much the Theis equation, it cannot be considered as a straightforward extension of the Theis equation, because the Theis equation is an exact solution of the 3D groundwater flow equation (albeit trivial because there are no vertical flow components), while the Hantush leaky well flow equation is based on the hydraulic flow approximation, where vertical flows are ignored, that is, the flow in the semi-confined aquifer is assumed to be horizontal and the leakage is uniformly distributed over the thickness of the aquifer. An exact solution of well flow in a semi-confined aquifer is much more difficult, but possible to obtain and has been given by Bruggeman (1999, p. 392), which in our notation becomes

$$s = \frac{Q}{2\pi T} \sum_{i=0}^{\infty} \frac{\sin \alpha_i}{\alpha_i + \sin \alpha_i \cos \alpha_i} W\left(\frac{Sr^2}{4Tt}, \frac{\alpha_i r}{D}\right) \cos\left(\frac{\alpha_i z}{D}\right) \quad (4.153)$$

where z is measured from the base of the aquifer, and α_i ($i = 0, 1, 2, \dots$) are the roots of $\alpha \tan \alpha = D^2/B^2$. Comparing Equation 4.153 with Equation 4.151 clearly illustrates the complexity which arises from considering full 3D groundwater flow vs. the hydraulic approximation.

Glossary

Bessel equation, Bessel function: Bessel's differential equation arises in many problems in physics, especially when using polar or cylindrical coordinates

$$x^2 \frac{d^2 y}{dx^2} + \frac{dy}{dx} \pm (x^2 \mp n^2)y = 0$$

The solutions of this equation $y(x, n)$ are called Bessel functions and have various and often interesting properties depending upon their type and order n . For more information consult standard mathematical books, as for instance Abramowitz and Stegun (1970) or Zwillinger (2002).

Boundary conditions: all partial differential equations have an infinite number of possible solutions, each of which corresponds to a particular case; to obtain from this multitude of possible solutions one particular solution corresponding to a certain specific case, it is necessary to provide supplementary information that is not contained in the partial differential equation; the supplementary information are termed boundary conditions and should include the initial state of the considered situation and how the considered domain interacts with its surroundings.

Conservation of mass: also termed continuity equation; a mathematical statement of the mass balance, that is, in the absence of mass-producing or mass-ingesting processes, the mass present in an arbitrary portion of a moving fluid remains constant.

Conservation of momentum: also termed Newton's second law of motion or momentum or impulse equation; a mathematical statement of the physical fact that forces are needed to induce deviation from uniform motion or to change motion; the law states that the time rate of change of an arbitrary portion of a moving fluid is equal to the resulting forces acting upon the considered portion.

Del operator: symbolic notation of a differential vector operator; several important operations as gradient, divergence, and curl can be described with this operator in a coordinate-independent way.

Elementary control volume: a volume of infinitesimal size, which is used to equate some physical property or process; when the size of the volume is taken in the limit going to zero, a mathematical equation is obtained that describes the property or process in every point of the continuum.

Navier–Stokes equation: equation describing the motion of a Newtonian fluid, that is, a viscous fluid for which the viscous stress components are linearly related to the deformation; when the Navier–Stokes equation is put in dimensionless form the Reynolds number appears, which expresses the ratio of inertia to viscous forces; for very small Reynolds numbers the inertia terms can be

neglected and the Navier–Stokes equation reduces to the Stokes equation describing a situation called Stokes flow, or creeping motion; the movement of groundwater under natural conditions is always considered to be creeping because of the very small groundwater flow velocities.

Land subsidence: a phenomenon that involves the lowering or settling of the land surface due to various factors, mostly under the impact of man's activities; often the extensive exploitation of groundwater resources have been accompanied in many places by significant land subsidence; in each case the ground layers consist of a sequence of sand and gravel intermixed with poorly consolidated and highly compressible clays or peat; the pumping of water from the conductive layers induces a drop in water potentials and gradual drainage of water from the clays or peat, which leads to compaction of the formations and sinking of the land surface.

Laplace transform: integral transformation that can reduce a differential equation to a simpler form or even to an algebraic expression; it is usually applied to time variables to eliminate time derivative, that is, if $f(t)$ is a function of t , the Laplace transform $\tilde{f}(p)$ is defined by

$$\tilde{f}(p) = L\{f(t); t \rightarrow p\} = \int_0^{\infty} f(t)e^{-pt} dt$$

where p is a complex parameter; the Laplace transform of the derivative df/dt becomes

$$L\left\{\frac{df}{dt}; t \rightarrow p\right\} = p\tilde{f}(p) - f(0)$$

Hence, the derivation becomes a multiplication. This reduces the complexity of a differential equation and can often enable to obtain the solution of the differential equation, whereby the true solution can be obtained by back transformation, which is usually achieved by look-up tables, as for instance given in Abramowitz and Stegun (1970) or Zwillinger (2002).

Piezometer: observation well tapping a saturated porous formation at a certain position and depth, indicating the groundwater potential present in the formation at that point.

Potential flow problem: potential flow is used to denote movement of fluids, without rotations in the driving force, which is a sufficient condition for the existence of a potential function, such that the gradient of that function gives the direction and magnitude of the driving force; in a homogeneous medium, potential functions are harmonic functions as they are solutions of the Laplace equation; there many mathematical techniques to solve these type of problems, especially for 2D flow problems that can be solved with the theory of complex functions.

Taylor series expansion: let $f(x)$ be a differentiable function of x , then the following series expansion holds

$$\begin{aligned} f(x) &= \sum_{n=0}^{\infty} \frac{x - x_0}{n!} \frac{d^n f(x_0)}{dx^n} \\ &= f_0 + \Delta x \frac{df_0}{dx} + \frac{(\Delta x)^2}{2} \frac{d^2 f_0}{dx^2} + \frac{(\Delta x)^3}{6} \frac{d^3 f_0}{dx^3} + \dots \end{aligned}$$

where $\Delta x = x - x_0$, and $f_0 = f(x_0)$.

References

- Abramowitz, M. and Stegun, I.A. 1970. *Handbook of Mathematical Functions*. Dover Publications Inc., New York.
- Bear, J. 1972. *Dynamics of Fluids in Porous Media*. American Elsevier Publ. Co., Inc., New York (reprinted in 1988 by Dover Publ. Inc., New York).

- Bear, J. 1979. *Hydraulics of Groundwater*. McGraw-Hill Inc., New York.
- Bear, J. 1988. *Dynamics of Fluids in Porous Media*. Dover Publications, Inc., New York.
- Bear, J. and Corapcioglu, M.Y. (eds). 1984. *Fundamentals of Transport Phenomena in Porous Media*. Martinus Nijhoff Publishers, Dordrecht, The Netherlands.
- Bear, J. and Verruijt, A. 1987. *Modeling Groundwater Flow and Pollution*. D. Reidel Publ. Co., Dordrecht, The Netherlands.
- Biot, M.A. 1955. Theory of elasticity and consolidation for a porous anisotropic solid. *J. Appl. Phys.* 26, 182–185.
- Bruggeman, G.A. 1999. *Analytical Solutions of Geohydrological Problems*. Elsevier, Amsterdam.
- Charny, I.A. 1951. A rigorous derivation of Dupuit's formula for unconfined seepage with seepage surface (in Russian). *Dokl. Akad. Nauk. SSSR* 79, 937–940.
- Cooper, H.H. 1966. The equation of groundwater flow in fixed and deforming coordinates. *J. Geophys. Res.* 71, 4785–4790.
- Dagan, G. 1989. *Flow and Transport in Porous Formations*. Springer-Verlag, Berlin, Heidelberg, New York.
- Dagan, G. and Neuman, S.P. (eds). 1997. *Subsurface Flow and Transport: A Stochastic Approach*. International Hydrology Series, Cambridge University Press, Cambridge, U.K.
- De Wiest, R.G.M. 1966. On the storage coefficient and the equations of groundwater flow. *J. Geophys. Res.* 71, 1117–1122.
- Freeze, R.A. and Cherry, J.A. 1979. *Groundwater*. Prentice Hall, Inc., Englewood Cliffs, NJ.
- Frind, E.O. 1983. Exact aquitard response functions for multiple aquifer mechanics. In *Flow through Porous Media*, ed. G.F. Pinder, pp. 86–92. CML U.K. Publications, Southampton, U.K.
- Gambolati, G. and Freeze, R.A. 1973. Mathematical simulation of the subsidence of Venice: 1. Theory. *Water Resources Res.* 9, 721–733.
- Halek, V. and Svec, J. 1979. *Groundwater Hydraulics*. Elsevier Scientific Publications Co., Amsterdam, Oxford, New York.
- Hantush, M.S. and Jacob, C.E. 1955. Non-steady radial flow in an infinite leaky aquifer. *Trans. Amer. Geoph. Union* 36, 95–100.
- Hantush, M.S. 1964. Hydraulics of wells. In *Advances in Hydrosciences*, vol. 1, ed. V.T. Chow, pp. 281–432. Academic Press, New York.
- Harr, M.E. 1962. *Groundwater and Seepage*. McGraw-Hill Inc., New York (reprinted in 1990 by Dover Publications Inc., New York).
- Holzbecher, E.K. 1998. *Modeling Density-Driven Flow in Porous Media*. Springer-Verlag, Berlin, Heidelberg, New York.
- Huisman, L. 1972. *Groundwater Recovery*. Macmillan Press Ltd., London and Basingstoke, U.K.
- Hvorslev, M.J. 1951. Time lag and soil permeability in ground-water observations. U.S. Army Corps of Engineers, Waterways Experimental Station, Bulletin 36, 50 pp.
- Jacob, C.E. 1940. On the flow of water in an elastic artesian aquifer. *Trans. Amer. Geoph. Union* 2, 574–586.
- Jacob, C.E. 1950. Flow of groundwater. In *Engineering Hydraulics*, ed. H. Rouse, pp. 321–386. John Wiley & Sons, New York.
- Kruseman, G.P. and de Ridder, N.A. 1991. *Analysis and Evaluation of Pumping Test Data*. International Institute for Land Reclamation and Improvement/ILRI. Wageningen, The Netherlands.
- de Marsily, G. 1986. *Quantitative Hydrogeology*. Academic Press, New York.
- McWhorter, P.B. and Sunada, D.K. 1977. *Groundwater Hydrology and Hydraulics*. Water Resources Publications, Colorado.
- Pinder, G.P. (ed.). 1983. *Flow through Porous Media*. CML Publications, Southampton, U.K.
- Polubarnova-Kochina, P.Ya. 1962. *Theory of Ground Water Movement*. Princeton University Press, Princeton, NJ.
- Rubin, Y. 2003. *Applied Stochastic Hydrogeology*. Oxford University Press, Oxford, New York.
- Strack, O.D.L. 1989. *Groundwater Mechanics*. Prentice Hall, Inc., Englewood Cliffs, NJ.
- Theis, C.V. 1935. The relation between the lowering of the piezometric surface and the rate and duration of discharge of a well using groundwater storage. *Trans. Amer. Geophys. Union* 16, 519–524.

- Todd, D.K. 1980. *Groundwater Hydrology*, 2nd ed. Wiley, New York.
- Verruijt, A. 1969. Elastic storage of aquifers. In *Flow through Porous Media*, ed. J.M. De Wiest, pp. 331–376, Academic Press, New York.
- Verruijt, A. 1970. *Theory of Groundwater Flow*. Macmillian, London and Basingstoke, U.K.
- Walton, W.A. 1970. *Groundwater Resource Evaluation*. McGraw-Hill, Inc., New York.
- Zijl, W. and Nawalany, M. 1993. *Natural Groundwater Flow*. Lewis Publishers, Boca Raton, FL, Ann Arbor, London, Tokyo.
- Zwillinger, D. (ed.) 2002. *CRC Standard Mathematical Tables and Formulae*. CRC Press, Boca Raton, FL.

Further Information

A standard reference in the field of mathematical aspects related to groundwater hydrology is Bear (1972). This work gives a comprehensive coverage of the dynamics of fluid flow through porous media. This book has been reprinted (Bear, 1988) and is highly recommended for scientists and engineers. Also parts that have been revised or updated and republished, as (Bear, 1979) dealing with principles of regional groundwater management and with special emphasis on the hydraulic approach to groundwater flow, and Bear and Verruijt (1987) in which especially mathematical and numerical modeling of groundwater flow and pollution transport are discussed.

Another classic work is Polubarnova-Kochina (1962), which gives an extensive overview of mathematical aspects of groundwater flow, with many analytical solutions to practical problems. Other notable works are McWhorter and Sunada (1977), Freeze and Cherry (1979), Todd (1980), and de Marsily (1986). A mathematical in-depth discussion of groundwater flow is Zijl and Nawalany (1993). Stochastic groundwater flow approach is discussed by Dagan (1989), Dagan and Neuman (1997), and Rubin (2003). A very comprehensive overview of analytical solutions of geohydrological problems and analytical solution methods is Bruggeman (1999).

New developments can be found in journals. Well known is *Water Resources Research*, published by the American Geophysical Union; this journal covers all scientific hydrology subjects. Also well known is the *Journal of Hydrology* of the European Geophysical Society, published by Elsevier. This journal also covers all aspects of hydrology. The most specialized journal in the field of mathematical groundwater hydrology as well as flow and transport through porous media for non-hydrological applications is *Transport in Porous Media*, edited by J. Bear and published by Kluwer Academic Publishers. Other interesting journals are *Journal of Groundwater* of the U.S. National Ground Water Association, and *Hydrogeology Journal* of the International Association of Hydrogeologists; these journals cover all aspects of hydrogeology.

5

Groundwater and Seepage: Accounting for Variability

5.1	Introduction.....	5-1
5.2	Some Groundwater Fundamentals	5-2
	Bernoulli's Equation • Darcy's Law • Reynolds Number • Homogeneity and Isotropy • Streamlines and Equipotential Lines	
5.3	Some Probabilistic Fundamentals	5-9
	Moments • Probability Distributions	
5.4	The Flow Net.....	5-21
5.5	The Point Estimate Method (PEM)	5-25
	One Random Variable • Regression and Correlation • Point Estimate Method — Several Random Variables	
5.6	Method of Fragments.....	5-31
5.7	Flow in Layered Systems.....	5-37
5.8	Piping: Reliability	5-43
	Glossary.....	5-47
	References	5-47
	Further Information	5-49

Milton E. Harr
Purdue University

5.1 Introduction

Figure 5.1 shows the pore space available for flow in two highly idealized soil models: *regular cubic* and *rhombohedral*. It is seen that even for these special cases, the pore space is not regular, but consists of cavernous cells interconnected by narrower channels. Pore spaces in real soils can range in size from molecular interstices to tunnel-like caverns. They can be spherical (as in concrete) or flat (as in clays), or display irregular patterns that defy description. Add to this the fact that pores may be *isolated* (inaccessible) or *interconnected* (accessible from both ends) or may be *dead-ended* (accessible through one end only).

In spite of the apparent irregularities and complexities of the available pores, there is hardly an industrial or scientific endeavor that does not concern itself with the passage of matter, solid, liquid, or gaseous, into, out of, or through porous media. Contributions to the literature can be found among such diverse fields (to name only a few) as soil mechanics, groundwater hydrology, petroleum, chemical and metallurgical engineering, water purification, materials of construction (ceramics, concrete, timber, paper), chemical

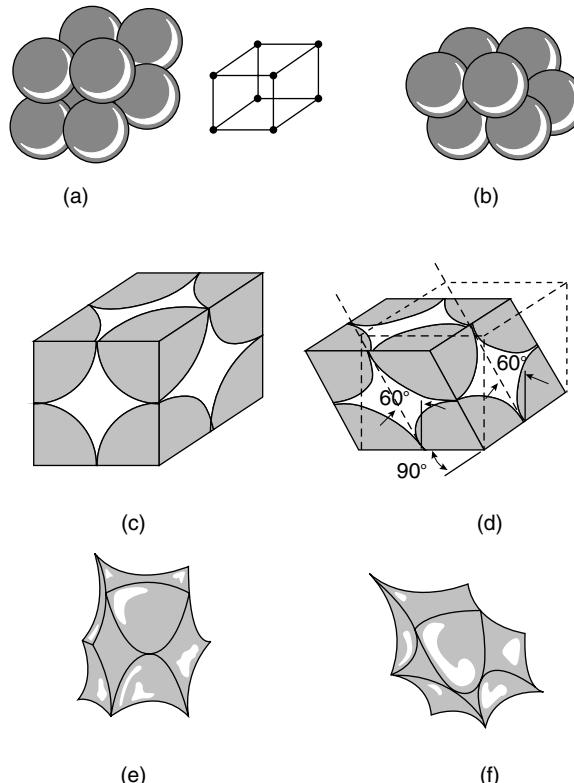


FIGURE 5.1 Idealized void space.

industry (absorbents, varieties of contact catalysts, and filters), pharmaceutical industry, traffic flow, and agriculture.

In no field of engineering are practitioners faced with more complex and uncertain sets of conditions than those concerned with groundwater systems. Unlike their colleagues in other engineering disciplines, who have the advantage of observing the performance of many prototypes under relatively known states of input, groundwater engineers generally deal with systems that are custom-built and tailored to special demands and specific locations. Add to these the very size of projects and the expense of underground sampling and data acquisition.

This chapter was written with the following main objectives:

1. To provide the groundwater engineer with an organized, logical, and systematic body of knowledge for the solution of groundwater and seepage problems.
2. To provide the groundwater engineer with probabilistic concepts and techniques to account for uncertainty in designs and to illustrate their relevance with application to practical problem situations.

5.2 Some Groundwater Fundamentals

The literature is replete with derivations and analytical excursions of the basic equations of steady-state groundwater flow (Polubarnova-Kochina, 1952; Harr, 1962; Cedergren, 1967; Bear, 1972; Domenico and Schwartz, 1990; to name only a few). A summary and brief discussion of these will be presented below for the sake of completeness.

5.2.1 Bernoulli's Equation

Underlying the analytical approach to groundwater flow is the representation of the actual physical system by a tractable mathematical model. In spite of their inherent shortcomings, many such analytical models have demonstrated considerable success in simulating the action of their prototypes.

As is well known from fluid mechanics, for steady flow of nonviscous incompressible fluids, Bernoulli's equation (Lamb, 1945)

$$\frac{p}{\gamma_w} + z + \frac{\bar{v}^2}{2g} = \text{constant} = h \quad (5.1)$$

where h = total head, ft, p = pressure, lb/ft², γ_w = unit weight of fluid, lb/ft³, g = gravitational constant, 32.2 ft/sec², \bar{v} = seepage velocity, ft/sec demonstrates that the sum of the *pressure head*, p/γ_w , *elevation head*, z , and *velocity head*, $\bar{v}^2/2g$, at any point within the region of flow is a constant.

To account for the loss of energy due to the viscous resistance within the individual pores, Bernoulli's equation is taken as

$$\frac{p_A}{\gamma_w} + z_A + \frac{\bar{v}_A^2}{2g} = \frac{p_B}{\gamma_w} + z_B + \frac{\bar{v}_B^2}{2g} + \Delta h \quad (5.2)$$

where Δh represents the total head loss (energy loss per unit weight of fluid) of the fluid over the distance Δs . The ratio

$$i = - \lim_{\Delta s \rightarrow 0} \frac{\Delta h}{\Delta s} = - \frac{dh}{ds} \quad (5.3)$$

is called the *hydraulic gradient* and represents the space rate of energy dissipation per unit weight of fluid (a pure number).

In most problems of interest the velocity heads (the kinetic energy) are so small they can be neglected. For example, a velocity of 1 ft/sec, which is large compared to typical seepage velocities through soils, produces a velocity head of only 0.015 ft. Hence, Equation 5.2 can be simplified to

$$\frac{P_A}{\gamma_w} + z_A = \frac{P_B}{\gamma_w} + z_B + \Delta h$$

and the total head at any point in the flow domain is simply

$$h = \frac{P}{\gamma_w} + z \quad (5.4)$$

5.2.2 Darcy's Law

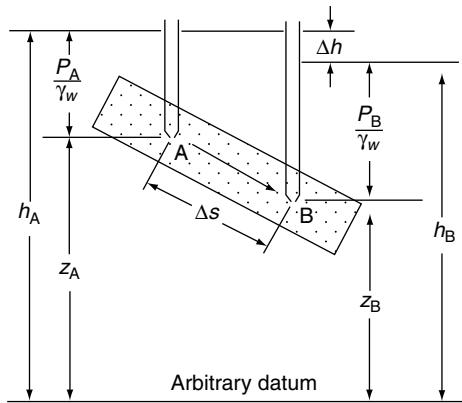
The flow of groundwater is taken to be governed by *Darcy's law*, which states that the velocity of the flow is directly proportional to the hydraulic gradient. A similar statement in an electrical system is *Ohm's law* and in a thermal system, *Fourier's law*. The grandfather of all such relations is *Newton's law of motion*. Table 5.1 presents some other points of similarity.

Prior to 1856, the formidable nature of the flow through porous media defied rational analysis. In that year, Henry Darcy published a simple relation based on his experiments on the flow of water in vertical sand filters in *Les Fontaines Publiques de la Ville de Dijon*, namely,

$$v = ki = -k \frac{dh}{ds} \quad (5.5)$$

TABLE 5.1 Some Similarities of Flow Models

Form of Energy	Name of Law	Quantity	Storage	Resistance
Electrical	Ohm's law	Current (voltage)	Capacitor	Resistor
Mechanical	Newton's law	Force (velocity)	Mass	Damper
Thermal	Fourier's law	Heat flow (temperature)	Heat capacity	Heat resistance
Fluid	Darcy's law	Flow rate (pressure)	Liquid storage	Permeability

**FIGURE 5.2** Heads in Bernoulli's equation.

Equation 5.5, commonly called Darcy's law, demonstrates a linear dependency between the hydraulic gradient and the *discharge velocity* v . The discharge velocity, $v = n\bar{v}$, is the product of the porosity n and the seepage velocity, \bar{v} . The coefficient of proportionality k in Equation 5.5 is called by many names depending on its use; among these are the *coefficient of permeability*, *hydraulic conductivity*, and *permeability constant*. As shown in Equation 5.5, k has the dimensions of a velocity. It should be carefully noted in this equation that flow is a consequence of differences in total head and not of pressure gradients. This is demonstrated in Figure 5.2 where the flow is directed from A to B, even though the pressure at point B is greater than that at point A.

Defining Q as the total volume of flow per unit time through a cross-sectional area A , Darcy's law takes the form

$$Q = Av = Aki = -Ak \frac{dh}{ds} \quad (5.6)$$

Darcy's law offers the single parameter k to account for both the characteristics of the medium and the fluid. It has been found that k is a function of γ_w , the unit weight of the fluid, μ , the coefficient of viscosity, and n , the porosity, as given by

$$k = C \frac{\gamma_w n}{\mu} \quad (5.7)$$

where C (dimensionally an area) typifies the structural characteristics of the media independent of the fluid properties. The principal advantage of Equation 5.7 lies in its use when dealing with more than one fluid or with temperature variations. When employing a single relatively incompressible fluid subjected to small changes in temperature, such as in groundwater and seepage-related problems, it is more convenient to use k as a single parameter. Some typical values for k are given in Table 5.2.

Although Darcy's law was obtained initially from considerations of one-dimensional macroscopic flow, its practical utility lies in its generalization into two or three spatial dimensions. Accounting for the

TABLE 5.2 Some Typical Values of Coefficient of Permeability

Soil Type	Coefficient of Permeability <i>k</i> , cm/sec
Clean gravel	1.0 and greater
Clean sand (coarse)	1.0–0.01
Sand (mixtures)	0.01–0.005
Fine sand	0.05–0.001
Silty sand	0.002–0.0001
Silt	0.0005–0.00001
Clay	0.000001 and smaller

directional dependence of the coefficient of permeability, Darcy's law can be generalized to

$$v_s = -k_s \frac{\partial h}{\partial s} \quad (5.8)$$

where k_s is the coefficient of permeability in the "s" direction, and v_s and $\partial h/\partial s$ are the components of the velocity and the hydraulic gradient, respectively, in that direction.

5.2.3 Reynolds Number

There remains now the question of the determination of the extent to which Darcy's law is valid in actual flow systems through soils. Such a criterion is furnished by the Reynolds number R (a pure number relating inertial to viscous force), defined as

$$R = \frac{vd\rho}{\mu} \quad (5.9)$$

where v = discharge velocity, cm/sec, d = average of diameter of particles, cm, ρ = density of fluid, g/(mass)/cm³, μ = coefficient of viscosity, g-sec/cm²

The critical value of the Reynolds number at which the flow in aggregations of particles changes from laminar to turbulent flow has been found by various investigators (see Muskat, 1937) to range between 1 and 12. However, it will generally suffice to accept the validity of Darcy's law when the Reynolds number is taken as equal to or less than unity, or

$$\frac{vd\rho}{\mu} \leq 1 \quad (5.10)$$

Substituting the known values of ρ and μ for water into Equation 5.10 and assuming a conservative velocity of 1/4 cm sec, we have d equal to 0.4 mm, which is representative of the average particle size of coarse sand.

5.2.4 Homogeneity and Isotropy

If the coefficient of permeability is independent of the direction of the velocity, the medium is said to be isotropic. Moreover, if the same value of the coefficient of permeability holds at all points within the region of flow, the medium is said to be *homogeneous* and *isotropic*. If the coefficient of permeability is dependent on the direction of the velocity and if this directional dependence is the same at all points of the flow region, the medium is said to be *homogeneous* and *anisotropic* (or *aleotropic*).

5.2.5 Streamlines and Equipotential Lines

Physically, all flow systems extend in three dimensions. However, in many problems the features of the motion are essentially planar, with the flow pattern being substantially the same in parallel planes. For these problems, for steady-state incompressible, isotropic flow in the xy plane, it can be shown (Harr, 1962) that the governing differential equation is

$$k_x \frac{\partial^2 h}{\partial x^2} + k_y \frac{\partial^2 h}{\partial y^2} = 0 \quad (5.11)$$

Here the function $h(x, y)$ is the distribution of the total head (of energy available to do work) within and on the boundaries of a flow region, and k_x and k_y are the coefficients of permeability in the x - and y -directions, respectively. If the flow system is isotropic, $k_x = k_y$, and Equation 5.11 reduces to

$$\frac{\partial^2 h}{\partial x^2} + \frac{\partial^2 h}{\partial y^2} = 0 \quad (5.12)$$

Equation 5.12, called *Laplace's equation*, is the governing relationship for steady-state, laminar-flow conditions (Darcy's law is valid). The general body of knowledge relating to Laplace's equation is called *potential theory*. Correspondingly, incompressible steady-state fluid flow is often called *potential flow*. The correspondence is more evident upon the introduction of the *velocity potential* ϕ , defined as

$$\phi(x, y) = -kh + C = -k \left(\frac{p}{\gamma_w} + z \right) + C \quad (5.13)$$

where h is the total head, p/γ_w is the pressure head, z is the elevation head, and C is an arbitrary constant. It should be apparent that, for isotropic conditions,

$$v_x = \frac{\partial \phi}{\partial x} \quad v_y = \frac{\partial \phi}{\partial y} \quad (5.14)$$

and, Equation 5.12 will produce

$$\frac{\partial^2 \phi}{\partial x^2} + \frac{\partial^2 \phi}{\partial y^2} = 0 \quad (5.15)$$

The particular solutions of Equation 5.12 or Equation 5.15 which yield the locus of points within a porous medium of equal potential, curves along with $h(x, y)$ or $\phi(x, y)$ are equal to a series of constants, are called *equipotential lines*.

In analyses of groundwater flow, the family of flow paths is given by the function $\psi(x, y)$, called the *stream function*, defined in two dimensions as (Harr, 1962)

$$v_x = \frac{\partial \psi}{\partial y} \quad v_y = -\frac{\partial \psi}{\partial x} \quad (5.16)$$

where v_x and v_y are the components of the velocity in the x - and y -directions, respectively.

Equating the respective potential and stream functions of v_x and v_y produces

$$\frac{\partial \phi}{\partial x} = \frac{\partial \psi}{\partial y} \quad \frac{\partial \phi}{\partial y} = -\frac{\partial \psi}{\partial x} \quad (5.17)$$

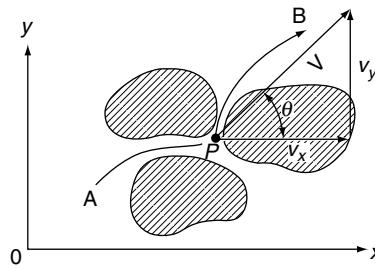


FIGURE 5.3 Path of flow.

Differentiating the first of these equations with respect to y and the second with respect to x and subtracting, we obtain Laplace's equation

$$\frac{\partial^2 \psi}{\partial x^2} + \frac{\partial^2 \psi}{\partial y^2} = 0 \quad (5.18)$$

We shall examine the significance of this relationship following a little more discussion of the physical meaning of the stream function.

Consider AB of Figure 5.3 as the path of a particle of water passing through point P with a tangential velocity v . We see from the figure that

$$\frac{v_y}{v_x} = \tan \theta = \frac{dy}{dx}$$

and hence

$$v_y dx - v_x dy = 0 \quad (5.19)$$

Substituting Equation 5.16, it follows that

$$\frac{\partial \psi}{\partial x} dx + \frac{\partial \psi}{\partial y} dy = 0$$

which states that the total differential $d\psi = 0$ and

$$\psi(x, y) = \text{constant}$$

Thus, we see that the family of curves generated by the function $\psi(x, y)$ equal to a series of constants are tangent to the resultant velocity at all points in the flow region and hence define the path of flow.

The potential, $\phi = -kh + \text{constant}$, is a measure of the energy available at a point in the flow region to move the particle of water from that point to the tailwater surface. Recall the locus of points of equal energy, say $\phi(x, y) = \text{constants}$, are called equipotential lines. The total differential along any curve, $\phi(x, y) = \text{constant}$, produces

$$d\phi = \frac{\partial \phi}{\partial x} dx + \frac{\partial \phi}{\partial y} dy = 0$$

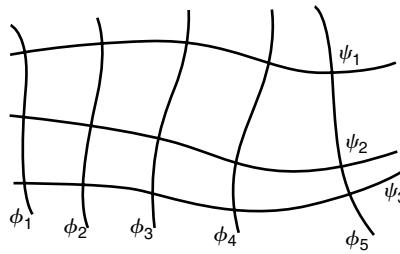


FIGURE 5.4 Streamlines and equipotential lines.

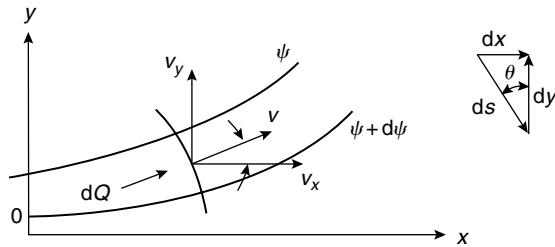


FIGURE 5.5 Flow between streamlines.

Substituting for $\partial\phi/\partial x$ and $\partial\phi/\partial y$ from Equation 5.16, we have

$$v_x dx + v_y dy = 0$$

and

$$\frac{dy}{dx} = -\frac{v_x}{v_y} \quad (5.20)$$

Noting the negative reciprocal relationship between their slopes, Equation 5.19 and Equation 5.20, we see that, within the flow domain, the families of streamlines $\psi(x, y) = \text{constants}$ and equipotential lines $\phi(x, y) = \text{constants}$ intersect each other at right angles. It is customary to signify the sequence of constants by employing a subscript notation, such as $\phi(x, y) = \phi_i$, $\psi(x, y) = \psi_j$ (Figure 5.4).

As only one streamline may exist at a given point within the flow medium, streamlines cannot intersect one another if flow is to occur between them. Consequently, if the medium is saturated, any pair of streamlines acts to form a flow channel between them.

Consider the flow between the two streamlines ψ and $\psi + d\psi$ in Figure 5.5; v represents the resultant velocity of flow. The quantity of flow through the flow channel per unit length normal to the plane of flow (say, cubic feet per second per foot) is

$$\begin{aligned} dQ &= v_x ds \cos \theta - v_y ds \sin \theta \\ &= v_x dy - v_y dx \\ &= \frac{\partial \psi}{\partial y} dy + \frac{\partial \psi}{\partial x} dx \end{aligned}$$

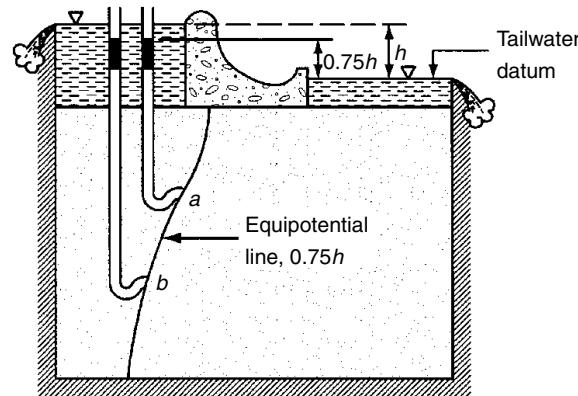


FIGURE 5.6 Pressure head along equipotential line.

and

$$dQ = d\psi \quad (5.21)$$

Hence, the quantity of flow (also called the *quantity of discharge* and *discharge quantity*) between any pair of streamlines is a constant whose value is numerically equal to the difference in their respective ψ values. Thus, once a sequence of streamlines of flow has been obtained, with neighboring ψ values differing by a constant amount, their plot will not only show the expected direction of flow but the relative magnitudes of the velocity along the flow channels; that is, the velocity at any point in the flow channel varies inversely with the streamline spacing in the vicinity of that point.

An equipotential line was defined previously as the locus of points where there is an expected level of available energy sufficient to move a particle of water from a point on that line to the tailwater surface. Thus, it is convenient to reduce all energy levels relative to a tailwater datum. For example, a piezometer located anywhere along an equipotential line, say at $0.75h$ in Figure 5.6, would display a column of water extending to a height of $0.75h$ above the tailwater surface. Of course, the pressure in the water along the equipotential line would vary with its elevation, Equation 5.4. This will be illustrated in Section 5.4.

5.3 Some Probabilistic Fundamentals

Within the context of engineering usage there are two primary definitions of the concept of probability: *relative frequency* and *subjective interpretation*. Historically, the measure first offered for the probability of an outcome was its *relative frequency*. If an outcome A can occur T times in N equally likely trials, the probability of the outcome A is

$$P[A] = \frac{T}{N} \quad (5.22a)$$

Implied in Equation 5.22a is that the probability of an outcome A equals the number of outcomes favorable to A (within the meaning of the experiment) divided by the total number of possible outcomes, or

$$P[A] = \frac{\text{Favorable outcomes}}{\text{Total possible outcomes}} \quad (5.22b)$$

Example 5.1

Find the probability of drawing a red card from an ordinary well-shuffled deck of 52 cards.

Solution. Of the 52 equally likely outcomes, there are 26 favorable (red card) outcomes. Hence,

$$P[\text{drawing red card}] = \frac{26}{52} = \frac{1}{2}$$

Understood from the example is that if one were to repeat the process a very large number of times, a red card would appear in one half of the trials. This is an example of the relative frequency interpretation. Now, what meaning could be associated with the statement, the probability of the failure of a proposed dam is 2% ($P[\text{failure}] = 0.02$)? The concept of repeated trials is meaningless: the structure will be built only *once* and it will either fail or be successful during its design lifetime. It cannot do both. Here, we have an example of the *subjective interpretation* of probability. It is a measure of information as to the likelihood of the occurrence of an outcome.

Subjective probability is generally more useful than the relative frequency concept in engineering applications. However, the basic rules governing both are identical. As an example, we note that both concepts specify the probability of an outcome to range between and to include numerical values from zero to one. The lower limit indicates there is *no* likelihood of occurrence: the upper limit corresponds to a *certain* outcome.

$$\langle \text{Axiom I} \rangle \quad 0 \leq P[A] \leq 1 \quad (5.23a)$$

The certainty of an outcome C is a probability of unity

$$\langle \text{Axiom II} \rangle \quad P[C] = 1 \quad (5.23b)$$

Equations 5.23 provide two of the three axioms of the theory of probability. The third axiom requires the concept of *mutually exclusive* outcomes. Outcomes are mutually exclusive if they cannot occur simultaneously. The third axiom states the probability of the occurrence of the sum of a number of mutually exclusive outcomes $A(1), A(2), \dots, A(N)$ is the sum of their individual probabilities (*addition rule*), or

$$\langle \text{Axiom III} \rangle \quad P[A(1) + A(2) + \dots + A(N)] = P[A(1)] + P[A(2)] + \dots + P[A(N)] \quad (5.23c)$$

As a very important application of these axioms, consider the proposed design of a spillway. After construction, only one of two outcomes can be obtained in the absolute structural sense: either it is successful or it fails. These are mutually exclusive outcomes. They are also exhaustive in that, within the sense of the example, no other outcomes are possible. Hence, the second axiom, Equation 5.23b, requires

$$P[\text{success+failure}] = 1$$

As they are mutually exclusive, the third axiom specifies that

$$P[\text{success}] + P[\text{failure}] = 1$$

The probability of the success of a structure is called its *reliability*, R . Symbolizing the probability of failure as $p(f)$, we have the important expression

$$R + p(f) = 1 \quad (5.24)$$

5.3.1 Moments

Consider a system of *discrete* parallel (vertical) forces, $P(1), P(2), \dots, P(N)$, acting on a rigid beam at the respective distances $x(1), x(2), \dots, x(N)$ (Figure 5.7a). From statics we have that the magnitude of the *equilibrant* M , is

$$M = \sum_{i=1}^N P(i) \quad (5.25a)$$

and its point of application, \bar{x} , is

$$\bar{x} = \frac{\sum_{i=1}^N x(i)P(i)}{\sum_{i=1}^N P(i)} \quad (5.25b)$$

For a *continuously* distributed parallel force system (Figure 5.7b) over a finite distance, say from $x(a)$ to $x(b)$, the corresponding expressions are

$$M = \int_{x(a)}^{x(b)} p(x)dx \quad (5.26a)$$

and

$$\bar{x} = \frac{\int_{x(a)}^{x(b)} xp(x)dx}{M} \quad (5.26b)$$

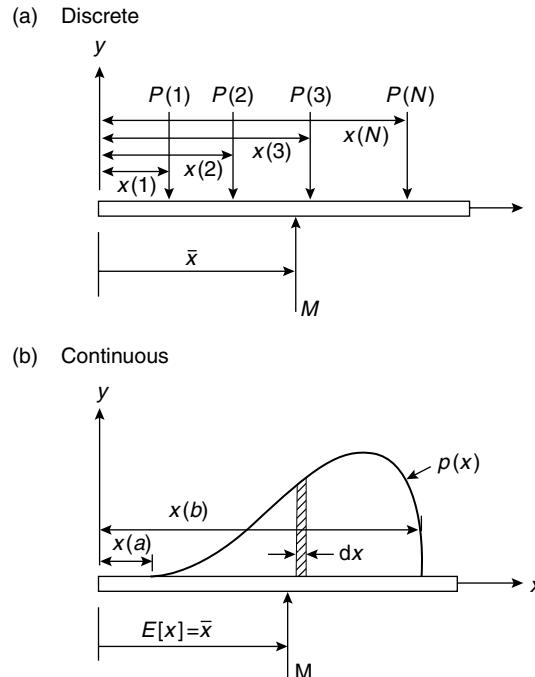


FIGURE 5.7 Equilibrant for discrete and continuous distributions.

Suppose that the discrete forces $P(i)$ in Figure 5.7a represent the frequencies of the occurrence of the N outcomes $x(1), x(2), \dots, x(N)$. As the distribution is exhaustive, from Axiom II, Equation 5.23b, the magnitude of the equilibrant must be unity, $M = 1$. Hence, Equation 5.25b becomes

$$\text{(Discrete)} \quad E[x] = \bar{x} = \sum_{i=1}^N x(i)P(i) \quad (5.27a)$$

Similarly, for the continuous distribution (Figure 5.7b), as all probabilities $p(x)dx$ must lie between $x(a)$ and $x(b)$, in Equation 5.26a $M = 1$. Hence, Equation 5.26b becomes

$$\text{(Continuous)} \quad E[x] = \bar{x} = \int_{x(a)}^{x(b)} xp(x)dx \quad (5.27b)$$

The symbol $E[x]$ in Equations 5.27 is called the *expected value* or the *expectation* or simply the *mean* of the variable x . As is true of the equilibrant, it is a measure of the central tendency, the *center of gravity* in statics.

Example 5.2

What is the expected value of the number of dots that will appear if a fair die is tossed?

Solution. Here, each of the possible outcomes 1, 2, 3, 4, 5, and 6 has the equal probability of $P(i) = 1/6$ of appearing. Hence, from Equation 5.27a

$$E[\text{toss of a fair die}] = \frac{1}{6}[1 + 2 + 3 + 4 + 5 + 6] = 3.5$$

We note in the above example that the expected value of 3.5 is an impossible outcome. There is no face on the die that will show 3.5 dots; however, it is still the best measure of the central tendency.

Example 5.3

Find the expected value of a continuous probability distribution wherein all values are equally likely to occur (called a *uniform distribution* [Figure 5.8]) between $y(a) = 0$ and $y(b) = 1/2$.

Solution. From Equation 5.26a as $M = 1$ and $p(y) = C$ is a constant, we have

$$1 = C \int_0^{1/2} dy, \quad C = 2$$

From Equation 5.27b, $E[y] = 2 \int_0^{1/2} y dy = 1/4$, as expected.

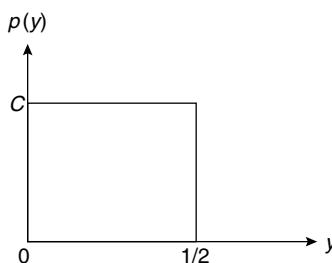


FIGURE 5.8 Uniform distribution.

As the variables x and y in the above examples are determined by the outcomes of random experiments they are said to be *random variables*. In classical probability theory, random variables are generally represented by capital letters, such as X and Y . The individual values are customarily denoted by their corresponding lowercase letters x and y ; however, no such distinction will be made here.

The expected value (mean) provides the locus of the central tendency of the distribution of a random variable. To characterize other attributes of the distribution, recourse is had to higher moments. Again, returning to statics, a measure of the dispersion of the distribution of the force system about the centroidal axis, at $x = E[x]$ in Figure 5.7b, is given by the *moment of inertia* (the *second central moment*),

$$I(y) = \int_{x(a)}^{x(b)} (x - \bar{x})^2 p(x) dx \quad (5.28)$$

The equivalent measure of the scatter (variability) of the distribution of a random variable is called its *variance*, denoted in this text as $\nu[x]$ (lowercase “ ν ”) and defined as

$$\langle \text{Discrete} \rangle \quad \nu[x] = \sum_{\text{all } x(i)} [x(i) - \bar{x}]^2 P(i) \quad (5.29a)$$

$$\langle \text{Continuous} \rangle \quad \nu[x] = \int_{x(a)}^{x(b)} (x - \bar{x})^2 p(x) dx \quad (5.29b)$$

In terms of the expectation these can be written as

$$\nu[x] = E[(x - \bar{x})^2] \quad (5.30)$$

which, after expansion, leads to a form more amenable to computations,

$$\nu[x] = E[x^2] - (E[x])^2 \quad (5.31)$$

This expression is the equivalent of the *parallel-axis theorem* for the moment of inertia.

Example 5.4

Find the expected value, and the variance of the *exponential distribution*, $p(x) = a \exp(-ax)$, $x > 0$, a is a constant.

Solution. We first show that $p(x)$ is a valid probability distribution

$$\int_0^\infty p(x) dx = a \int_0^\infty e^{-ax} dx = 1$$

The expected value is

$$E[x] = a \int_0^\infty x e^{-ax} dx = \frac{1}{a}$$

Continuing,

$$E[x^2] = a \int_0^\infty x^2 e^{-ax} dx = \frac{1}{a^2}$$

whence, using Equation 5.31,

$$\nu[x] = \frac{2}{a^2} - \left(\frac{1}{a}\right)^2 = \frac{1}{a^2}$$

The variance has the units of the square of those of the random variable. A more meaningful measure of dispersion of a random variable (x) is the positive square root of its variance (compare with *radius of gyration* of mechanics) called the *standard deviation*, $\sigma[x]$,

$$\sigma[x] = \sqrt{\nu[x]} \quad (5.32)$$

From the results of the previous example, it is seen that the standard deviation of the exponential distribution is $\sigma[x] = 1/a$.

An extremely useful relative measure of the scatter of a random variable $\sigma[x]$ is its *coefficient of variation* $V(x)$, capital "V" with parenthesis, usually expressed as a percentage,

$$V(x) = \frac{\sigma[x]}{E[x]} \times 100(%) \quad (5.33)$$

For the exponential distribution we found, $\sigma[x] = 1/a$ and $E[x] = 1/a$, hence V (exponential distribution) = 100%. In Table 5.3 are given representative values of the coefficients of variation. Original sources should be consulted for details.

The coefficient of variation expresses a measure of the reliability of the central tendency. For example, given a mean value of a parameter of 10 with a coefficient of variation of 10% would indicate a standard deviation of 1, whereas a similar mean with a coefficient of variation of 20% would demonstrate a standard deviation of 2. The coefficient of variation has been found to be a fairly stable measure of variability for homogeneous conditions. Additional insight into the standard deviation and the coefficient of variation as measures of uncertainty is provided by *Chebyshev's inequality* (for the derivation see Lipschutz, 1965).

The spread of a random variable is often spoken of as its *range*, the difference between the largest and smallest outcomes of interest. Another useful measure is the range between the mean plus-and-minus- h -standard deviations, $\bar{x} \pm h\sigma[x]$, called the *h-sigma bounds* (Figure 5.9). If x is a random variable with mean value \bar{x} and standard deviation σ , then Chebyshev's inequality states

$$P[(\bar{x} - h\sigma) \leq x \leq (\bar{x} + h\sigma)] \geq \frac{1}{h^2} \quad (5.34)$$

In words, it asserts that for any probability distribution (with finite mean and standard deviation) the probability that random values of the variate will lie within h -sigma bounds is at least $(1 - 1/h^2)$. Some numerical values are given in Table 5.4. It is seen that quantitative probabilistic statements can be made without complete knowledge of the probability distribution function: only its expected value and coefficient of variation (or standard deviation) are required. In this regard, the values for the coefficients of variation given in the table may be used in the absence of more definitive information.

Example 5.5

The expected value for the ϕ -strength parameter of a sand is 30° . What is the probability that a random sample of this sand will have a ϕ -value between 20° and 40° ?

TABLE 5.3 Representative Coefficient of Variation

Parameter	Coefficient of Variation (%)	Source
Soil		
Porosity	10	Schultze (1972)
Specific gravity	2	Padilla and Vanmarcke (1974)
Water content		
Silty clay	20	Padilla and Vanmarcke (1974)
Clay	13	Fredlund and Dahlman (1972)
Degree of saturation	10	Fredlund and Dahlman (1972)
Unit weight	3	Hammitt (1966)
Coefficient of permeability	(240 at 80% saturation to 90 at 100% saturation)	Nielsen et al. (1973)
Compressibility factor	16	Padilla and Vanmarcke (1974)
Preconsolidation pressure	19	Padilla and Vanmarcke (1974)
Compression index		
Sandy clay	26	Lumb (1966)
Clay	30	Fredlund and Dahlman (1972)
Standard penetration test	26	Schultze (1975)
Standard cone test	37	Schultze (1975)
Friction angle ϕ		
Gravel	7	Schultze (1972)
Sand	12	Schultze (1972)
c , strength parameter (cohesion)	40	Fredlund and Dahlman (1972)
Structural Loads, 50-Year Maximum		
Dead load	10	Ellingwood et al. (1980)
Live load	25	Ellingwood et al. (1980)
Snow load	26	Ellingwood et al. (1980)
Wind load	37	Ellingwood et al. (1980)
Earthquake load	> 100	Ellingwood et al. (1980)
Structural Resistance		
Structural steel		
Tension members, limit state, yielding	11	Ellingwood et al. (1980)
Tension members, limit state, tensile strength	11	Ellingwood et al. (1980)
Compact beam, uniform moment	13	Ellingwood et al. (1980)
Beam, column	15	Ellingwood et al. (1980)
Plate, girders, flexure	12	Ellingwood et al. (1980)
Concrete members		
Flexure, reinforced concrete, grade 60	11	Ellingwood et al. (1980)
Flexure, reinforced concrete, grade 40	14	Ellingwood et al. (1980)
Flexure, cast-in-place beams	8–9.5	Ellingwood et al. (1980)
Short columns	12–16	Ellingwood et al. (1980)
Ice		
Thickness	17	Bercha (1978)
Flexural strength	20	Bercha (1978)
Crushing strength	13	Bercha (1978)
Flow velocity	33	Bercha (1978)
Wood		
Moisture	3	Borri et al. (1983)
Density	4	Borri et al. (1983)
Compressive strength	19	Borri et al. (1983)
Flexural strength	19	
Glue-laminated beams		Borri et al. (1983)
Live load	18	Galambos et al. (1982)
Snow load	18	Galambos et al. (1982)

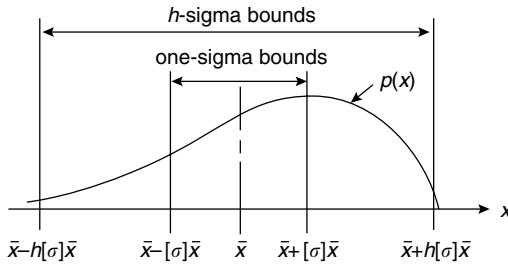


FIGURE 5.9 Range in h -sigma bounds.

TABLE 5.4 Probabilities for Range of Expected Values $\pm h$ -Sigma Bounds

h	Chebyshev's Inequality	Gauss's Inequality	Exact Exponential Distribution	Exact Normal Distribution	Exact Uniform Distribution
1/2	0	0	0.78	0.38	0.29
1	0	0.56	0.86	0.68	0.58
2	0.75	0.89	0.95	0.96	1.00
3	0.89	0.95	0.982	0.9973	1.00
4	0.94	0.97	0.993	0.999934	1.00

Solution. From Table 5.3, $V(\phi) = 12\%$ for sand; hence, $\sigma[\phi]$ is estimated to be $(0.12)(30) = 3.6^\circ$ and $h = (\bar{\phi} - \phi)/\sigma = 10^\circ/3.6^\circ = 2.8$. Hence, $P[20^\circ \leq \phi \leq 40^\circ] \geq 1 - (1/2.8)^2 \geq 0.87$. That is, the probability is at least 0.87 that the ϕ -strength parameter will be between 20° and 40° .

If the unknown probability distribution function is symmetrical with respect to its expected value and the expected value is also its maximum value (said to be *unimodal*), it can be shown that (Freeman, 1963)

$$P[(\bar{x} - h\sigma) \leq x \leq (\bar{x} + h\sigma)] \geq 1 - \frac{4}{9h^2} \quad (5.35)$$

This is sometimes called *Gauss' inequality*. Some numerical values are given in the third column of Table 5.4.

Example 5.6

Repeat the previous example if it is assumed that the distribution of the ϕ -value is symmetrical with its maximum at the mean value ($\bar{\phi} = 30^\circ$).

Solution. For this case, Gauss' inequality asserts

$$P[20^\circ \leq \phi \leq 40^\circ] \geq 1 - \frac{4}{9(2.8)^2} \geq 0.94$$

Recognizing symmetry we can also claim $P[\phi \leq 20^\circ] = P[\phi \geq 40^\circ] = 0.03$.

Example 5.7

Find the general expression for the probabilities associated with h -sigma bounds for the exponential distribution, $h \geq 1$.

Solution. From Example 5.4, we have (with $E[x] = 1/a$, $\sigma[x] = 1/a$)

$$P[(\bar{x} - h\sigma) \leq x \leq (\bar{x} + h\sigma)] = \int_0^{(h+1)/a} ae^{-ax} dx = 1 - e^{-(h+1)}$$

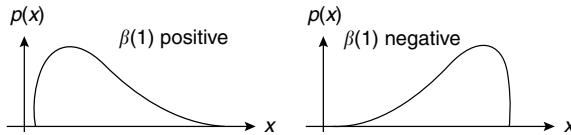


FIGURE 5.10 Coefficient of skewness.

Some numerical values are given in Table 5.4. The normal distribution noted in this table will be developed subsequently.

The results in Table 5.4 indicate that, lacking information concerning a probability distribution beyond its first two moments, from a practical engineering point of view, it may be taken to range within 3-sigma bounds. That is, in Figure 5.7b, $x(a) \approx \bar{x} - 3\sigma[x]$ and $x(b) \approx \bar{x} + 3\sigma[x]$.

For a symmetrical distribution all moments of odd order about the mean (central moments) must be zero. Consequently, any odd-ordered moment may be used as a measure of the degree of *skewness* or *asymmetry* of a probability distribution. The fourth central moment $E[(x - \bar{x})^4]$ provides a measure of the peakedness (called *kurtosis*) of a distribution.

As the units of the third central moment are the cube of the units of the variable, to provide an absolute measure of skewness, Pearson [1894, 1895] proposed that its value be divided by the standard deviation cubed to yield the dimensionless *coefficient of skewness*,

$$\beta(1) = \frac{E[(x - \bar{x})^3]}{(\sigma[x])^3} \quad (5.36)$$

If $\beta(1)$ is positive, the long tail of the distribution is on the right side of the mean: if it is negative the long tail is on the left side (Figure 5.10). Pearson also proposed the dimensionless *coefficient of kurtosis* $\beta(2)$ as a measure of peakedness

$$\beta(2) = \frac{E[(x - \bar{x})^4]}{(\sigma[x])^4} \quad (5.37)$$

5.3.2 Probability Distributions

In Figure 5.11 are shown the regions occupied by a number of probability distribution types identified by their coefficients of skewness and kurtosis. Examples of the various types are also shown schematically. We note in the figure that the type IV distribution and the symmetrical type VII are unbounded (infinite) below and above. From the point of view of practical groundwater engineering applications, this represents an extremely unlikely distribution. For example, *all* parameters or properties (see Table 5.3) are positive numbers (including zero).

The type V (the *lognormal distribution*), type III (the gamma), and type VI distributions are unbounded above. Hence, their use would be confined to those variables with an extremely large range of possible values. Some examples are the coefficient of permeability, the state of stress at various points in a body, distribution of annual rainfall, and so on.

The *normal (Gaussian) distribution* [$\beta(1) = 0, \beta(2) = 3$], even though it occupies only a single point in the universe of possible distributions, is the most frequently used probability model. Some associated properties were given in Table 5.4. The normal distribution is the well-known symmetrical bell-shaped curve (Figure 5.12). Some tabular values are given in Table 5.5. The table is entered by forming the standardized variable z , for the normal variate x as

$$z = \left| \frac{x - \bar{x}}{\sigma[x]} \right| \quad (5.38)$$

Tabular values yield the probabilities associated with the shaded areas shown in the figure: area = $\psi(z)$.

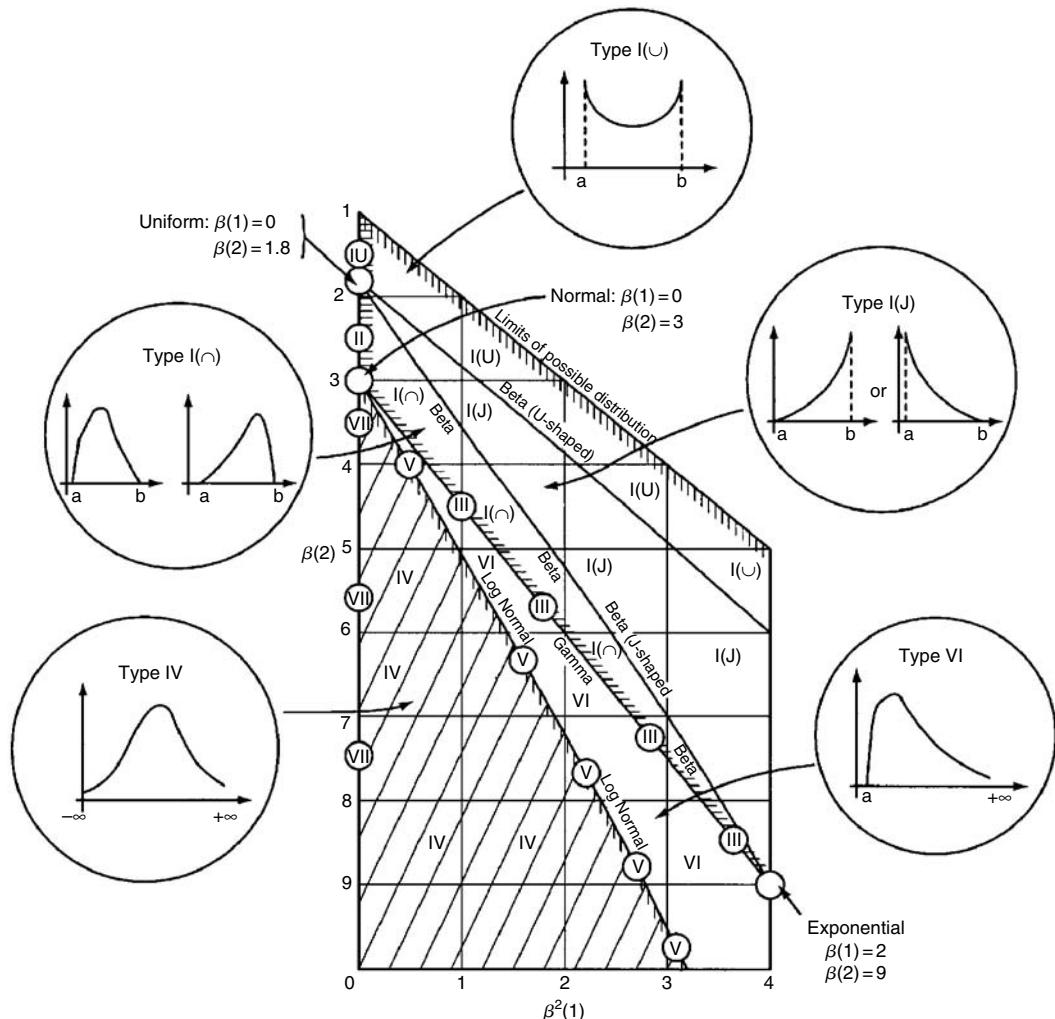


FIGURE 5.11 Space of probability distributions. (Adapted from Pearson, E.S. and Hartley, H.O. 1972. *Biometrika Tables for Statisticians*, vol. II. Cambridge University Press, London.)

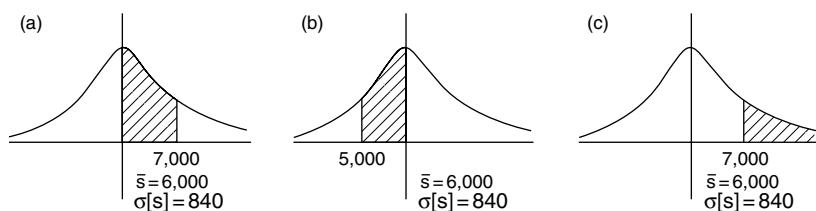


FIGURE 5.12 Example 5.8.

Example 5.8

Assuming the strength s of concrete to be used in a weir to be a normal variate with an expected value of $\bar{s} = 6000$ psi and a coefficient of variation of 14%, find (a) $P[6000 \leq s \leq 7000]$, (b) $P[5000 \leq s \leq 6000]$, and (c) $P[s \geq 7000]$.

TABLE 5.5 Standardized Normal Variates $N[0,1]$

for $z > 2.2$, $\psi(z) \approx \frac{1}{2} - \frac{1}{z} (2\pi)^{-1/2} \exp[-\frac{z^2}{2}]$

$$z = \frac{x-\bar{x}}{\sigma[x]}$$

z	0	1	2	3	4	5	6	7	8	9
0	0	.003969	.007978	.011966	.015953	.019939	.023922	.027903	.031881	.035856
.1	.039828	.043795	.047758	.051717	.055670	.059618	.063559	.067495	.071424	.075345
.2	.079260	.083166	.087064	.090954	.094835	.098706	.102568	.106420	.110251	.114092
.3	.117911	.121720	.125516	.129300	.133072	.136831	.140576	.144309	.148027	.151732
.4	.155422	.159097	.162757	.166402	.170031	.174645	.177242	.180822	.184386	.187933
.5	.191462	.194974	.198466	.201944	.205401	.208840	.212260	.215661	.219043	.222405
.6	.225747	.229069	.232371	.235653	.238914	.242154	.245373	.248571	.251748	.254903
.7	.258036	.261148	.264238	.267305	.270350	.273373	.276373	.279350	.282305	.285236
.8	.288145	.291030	.293892	.296731	.299546	.302337	.305105	.307850	.310570	.313267
.9	.315940	.318589	.321214	.323814	.326391	.328944	.331472	.333977	.336457	.338913
1.0	.341345	.343752	.346136	.348495	.350830	.353141	.355428	.357690	.359929	.362143
1.1	.364334	.366500	.368643	.370762	.372857	.374928	.376976	.379000	.381000	.382977
1.2	.384930	.386861	.388768	.390651	.392512	.394350	.396165	.397958	.399727	.401475
1.3	.403200	.404902	.406582	.408241	.409877	.411492	.413085	.414657	.416207	.417736
1.4	.419243	.420730	.422196	.423641	.425066	.426471	.427855	.429219	.430563	.431888
1.5	.433193	.434476	.435745	.436992	.438220	.439429	.440620	.441792	.442947	.444083
1.6	.445201	.446301	.447384	.448449	.449497	.450529	.451543	.452540	.453521	.454486
1.7	.455435	.456367	.457284	.458185	.459070	.459941	.460796	.461636	.462462	.463273
1.8	.464070	.464852	.465620	.466375	.467116	.467843	.468557	.469258	.469946	.470621
1.9	.471283	.471933	.472571	.473197	.473610	.474412	.475002	.475581	.476148	.476705
2.0	.477250	.477784	.478308	.478822	.479325	.479818	.480301	.480774	.481237	.481691
2.1	.482136	.482571	.482997	.483414	.483823	.484222	.484614	.484997	.485371	.485738
2.2	.486097	.486447	.486791	.487126	.487455	.487776	.488089	.488396	.488696	.488989
2.3	.489297	.489556	.489830	.490097	.490358	.490613	.490863	.491106	.491344	.491576
2.4	.491802	.492024	.492240	.492451	.492656	.492857	.493053	.493244	.493431	.493613
2.5	.493790	.493963	.494132	.494297	.494457	.494614	.494766	.494915	.495060	.495201
2.6	.495339	.495473	.495604	.495731	.495855	.495975	.496093	.496207	.496319	.496427
2.7	.496533	.496636	.496736	.496833	.496928	.497020	.497110	.497197	.497282	.497365
2.8	.497445	.497523	.497599	.497673	.497744	.497814	.497882	.497948	.498012	.498074
2.9	.498134	.498293	.498250	.498305	.498359	.498411	.498462	.498511	.498559	.498605
3.0	.498650	.498694	.498736	.498777	.498817	.498856	.498893	.498930	.498965	.498999
3.1	.499032	.499065	.499096	.499126	.499155	.499184	.499211	.499238	.499264	.499289
3.2	.499313	.499336	.499359	.499381	.499402	.499423	.499443	.499462	.499481	.499499
3.3	.499517	.499534	.499550	.499566	.499581	.499596	.499610	.499624	.499638	.499651
3.4	.499663	.499675	.499687	.499698	.499709	.499720	.499730	.499740	.499749	.499758
3.5	.499767	.499776	.499784	.499792	.499800	.499807	.499815	.499822	.499828	.499835
3.6	.499841	.499847	.499853	.499858	.499864	.499869	.499874	.499879	.499883	.499888
3.7	.499892	.499896	.499900	.499904	.499908	.499912	.499915	.499918	.499922	.499925
3.8	.499928	.499931	.499933	.499936	.499938	.499941	.499943	.499946	.499948	.499950
3.9	.499952	.499954	.499956	.499958	.499959	.499961	.499963	.499964	.499966	.499967

Solution. The standard deviation is $\sigma[s] = (0.14)(6000) = 840$ psi. Hence (see Figure 5.12),

- (a) $z = |(7000 - 6000)/840| = 1.19$; hence, from Table 5.5 $\psi(1.19) = 0.383$.
- (b) By symmetry, $P[5000 \leq s \leq 6000] = 0.383$.
- (c) $P[s \geq 7000] = 0.500 - 0.383 = 0.117$.

As might be expected from its name, the *lognormal distribution* (type V) is related to the normal distribution. If x is a normal variate and $x = \ln y$ or $y = \exp(x)$, then y is said to have a lognormal distribution. It is seen that the distribution has a minimum value of zero and is unbounded above. The probabilities associated with lognormal variates can be obtained very easily from those of mathematically corresponding normal variates (Table 5.5). If $E(y)$ and $V(y)$ are the expected value and coefficient of variation of a lognormal variate, the corresponding normal variate x will have the expected value and standard deviation (Benjamin and Cornell, 1970).

$$(\sigma[x])^2 = \ln\{1 + [V(y)]^2\} \quad (5.39a)$$

$$E[x] = \ln E(y) - (\sigma[x])^2/2 \quad (5.39b)$$

Example 5.9

A live load of 20 tons acts on a footing. If the loading is assumed to be lognormally distributed, estimate the probability that a loading of 40 tons will be exceeded.

Solution. From Table 5.3 we have that the coefficient of variation for a live load, say L , can be estimated as 25%; hence, from Equations 5.39 we have for the corresponding normal variate, say x ,

$$\sigma[x] = \sqrt{\ln[1 + (0.25)^2]} = 0.25$$

and

$$E[x] = \ln 20 - (0.25)^2/2 = 2.96$$

As $x = \ln L$, the value of the normal variate x equivalent to 40 tons is $\ln 40 = 3.69$. We seek the equivalent normal probability $P[3.69 \leq x]$. The standardized normal variate is $z = (3.69 - 2.96)/0.25 = 2.92$. Hence, using Table 5.5:

$$P[40 \leq L] = 0.50 - \psi(2.92) = 0.500 - 0.498 = 0.002$$

Particular attention should be paid to the regions of Figure 5.11, labeled the beta distributions. It is seen that many distributions (such as the normal, lognormal, uniform, and exponential) can be obtained as special cases of the very versatile beta distribution. Inherent in selecting a probability distribution to characterize a random variable is the imposition of the limits or range of its applicability. For example, for the normal it is required that the variable range from $-\infty$ to $+\infty$; the range of the lognormal and the exponential is from 0 to $+\infty$. Such assignments may not be critical if knowledge of distributions are desired in the vicinity of their expected values and their coefficients of variation are not excessive (say, less than 25%). On the other hand, many probabilistic assessments (such as reliability or the probability of failure) are vested in the tails of distributions. It is in such characterizations that the beta distribution is of great value.

The beta distribution is treated in great detail by Harr (1977, 1987). The latter reference also contains FORTRAN programs for beta probability distributions. This writer modified Pearson's beta distribution by fitting it to the four parameters: the expected value, the standard deviation, the minimum, and the

maximum values. Pearson's original work used the coefficients of skewness and kurtosis instead of the minimum and maximum values. Motivating the change was the sensitivity of the distribution to the coefficients. On the other hand, from an engineering point of view, one generally has a feel for the extremes of a design variable. Often, zero is the optional minimum value. In the event that limits are not defined, the specification of a range as the mean plus or minus, say, three standard deviations, generally places the generated beta distribution well within the accuracy required for most groundwater engineering applications (see Table 5.4).

We next return to groundwater and seepage.

5.4 The Flow Net

The graphical representation of special members of the families of streamlines and corresponding equipotential lines within a flow region form a *flow net*. The orthogonal network shown in Figure 5.13 represents such a system. Although the construction of a flow net often requires tedious trial-and-error adjustments, it is one of the more valuable methods employed to obtain the solutions for two-dimensional flow problems. Of additional importance, even a hastily drawn flow net will often provide a check on the reasonableness of solutions obtained by other means. Noting that, for steady-state conditions, Laplace's equation also models the action of (see Table 5.1) thermal, electrical, acoustical, odoriferous, torsional, and other systems, the flow net is seen to be a significant tool for analysis.

If, in Figure 5.13, Δw denotes the distance between a pair of adjacent streamlines and Δs is the distance between a pair of adjacent equipotential lines in the near vicinity of a point within the region of flow, the approximate velocity (in the mathematical sense) at the point, according to Darcy's law, will be

$$v \approx \frac{k\Delta h}{\Delta s} = \frac{\Delta \psi}{\Delta w} \quad (5.40)$$

As the quantity of flow between any two streamlines is a constant, ΔQ , and equal to $\Delta \psi$ (Equation 5.21) we have

$$\Delta Q \approx k \frac{\Delta w}{\Delta s} \Delta h \quad (5.41)$$

Equation 5.40 and Equation 5.41 are approximate. However, as the distances Δw and Δs become very small, Equation 5.40 approaches the velocity at a point and Equation 5.41 yields the quantity of discharge through the flow channel.

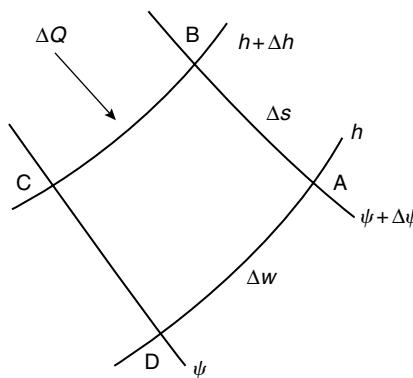


FIGURE 5.13 Flow at point.

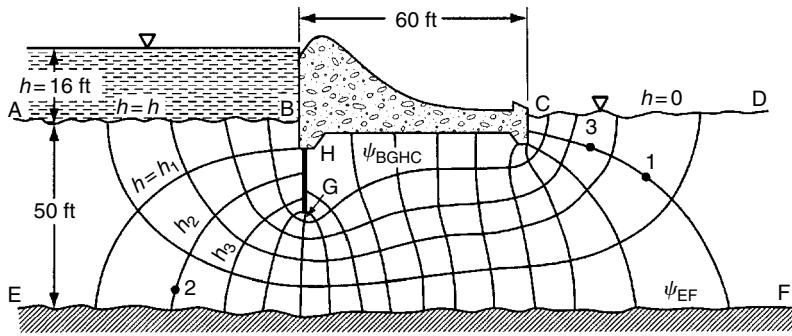


FIGURE 5.14 Example of flow net.

In Figure 5.14 is shown the completed flow net for a common type of structure. We first note that there are four boundaries: the bottom impervious contour of the structure $BGHC$, the surface of the impervious layer EF , the headwater boundary AB , and the tailwater boundary CD . The latter two boundaries designate the equipotential lines $h = h$ and $h = 0$, respectively. For steady-state conditions the quantity of discharge through the section Q , and the head loss ($h=16$ ft) must be constant. If the flow region is saturated, it follows that the two impervious boundaries are not only streamlines but their difference must be identically equal to the discharge quantity

$$Q = \psi_{BGHC} - \psi_{EF}$$

From among the infinite number of possible streamlines between the impervious boundaries, we sketch only a few specifying the same quantity of flow between neighboring streamlines. Designating N_f as the number of flow channels, we have, from above,^{*}

$$Q = N_f \Delta Q = k N_f \frac{\Delta w}{\Delta s} \Delta h$$

Similarly, from among the infinite number of possible equipotential lines between headwater and tailwater boundaries, we sketch only a few and specify the same drop in head, say Δh , between adjacent equipotential lines. If there are N_e equipotential drops along each of the channels,

$$h = N_e \Delta h$$

and

$$Q = k \frac{N_f}{N_e} \frac{\Delta w}{\Delta s} h \quad (5.42)$$

If, now, we also require that the ratio $\Delta w/\Delta s$ be the same at all points in the flow region, for convenience, and as a square is most sensitive to visual inspection, we take this ratio to be unity,

$$\frac{\Delta w}{\Delta s} = 1$$

and obtain

$$Q = k \frac{N_f}{N_e} h \quad (5.43)$$

^{*}There is little to be gained by retaining the approximately equal sign \approx .

Recalling that Q , k , and h are all constants, Equation 5.43 demonstrates that the resulting construction, with the obvious requirement that everywhere within the flow domain streamlines and equipotential lines meet at right angles, will yield a unique value for the ratio of the number of flow channels to the number of equipotential drops, N_f/N_e . In Figure 5.14, N_f equals about 5 and N_e equals 16; hence, $N_f/N_e = 5/16$.

The graphical technique of constructing flow nets by sketching was first suggested by Prasil (1913) although it was developed formally by Forchheimer (1930); however, the adoption of the method by engineers followed Casagrande's classical paper in 1940. In this paper and in the highly recommended flow nets of Cedergren (1967) are to be found some of the best examples of the art of drawing flow nets. Harr (1962) also warrants a peek!

Unfortunately, there is no "royal road" to drawing a good flow net. The speed with which a successful flow net can be drawn is highly contingent on the experience and judgment of the individual. In this regard, the beginner will do well to study the characteristics of well-drawn flow nets: *labor omnia vincit*.

In summary, a flow net is a sketch of distinct and special streamlines and equipotential lines that preserve right-angle intersections within the flow regime, satisfy the boundary conditions, and form curvilinear squares.* The following procedure is recommended:

1. Draw the boundaries of the flow region to a scale so that all sketched equipotential lines and streamlines terminate on the figure.
2. Sketch lightly three or four streamlines, keeping in mind that they are only a few of the infinite number of possible curves that must provide a smooth transition between the boundary streamlines.
3. Sketch the equipotential lines, bearing in mind that they must intersect all streamlines, including boundary streamlines, at right angles and that the enclosed figures must form curvilinear rectangles (except at singular points) with the same ratio of $\Delta s/\Delta s$ along a flow channel. Except for partial flow channels and partial head drops, these will form curvilinear squares with $\Delta w = \Delta s$.
4. Adjust the initial streamlines and equipotential lines to meet the required conditions. Remember that the drawing of a good flow net is a trial-and-error process, with the amount of correction being dependent upon the position selected for the initial streamlines.

Example 5.10

Obtain the quantity of discharge, Q/kh , for the section shown in Figure 5.15.

Solution. This represents a region of horizontal flow with parallel horizontal streamlines between the impervious boundaries and vertical equipotential lines between reservoir boundaries. Hence, the flow net will consist of perfect squares, and the ratio of the number of flow channels to the number of drops will be $N_f/N_e = a/L$ and $Q/kh = a/L$.

Example 5.11

Find the pressure in the water at points A and B in Figure 5.15.

Solution. For the scheme shown in Figure 5.15, the total head loss is linear with distance in the direction of flow. Equipotential lines are seen to be vertical. The total heads at points A and B are both equal to $2h/3$ (datum at the tailwater surface). This means that a piezometer placed at these points would show a column of water rising to an elevation of $2h/3$ above the tailwater elevation. Hence, the pressure at each point is simply the weight of water in the columns above the points in question: $p_A = (2h/3 + h_1)\gamma_w$, $p_B = (2h/3 + h_1 + a)\gamma_w$.

*We except singular squares such as the five-sided square at point H in Figure 5.14 and the three-sided square at point G. (It can be shown [Harr, 1962, p. 84] that a five-sided square designates a point of turbulence.) With continued subdividing into smaller squares, the deviations, in the limit, act only at singular points.

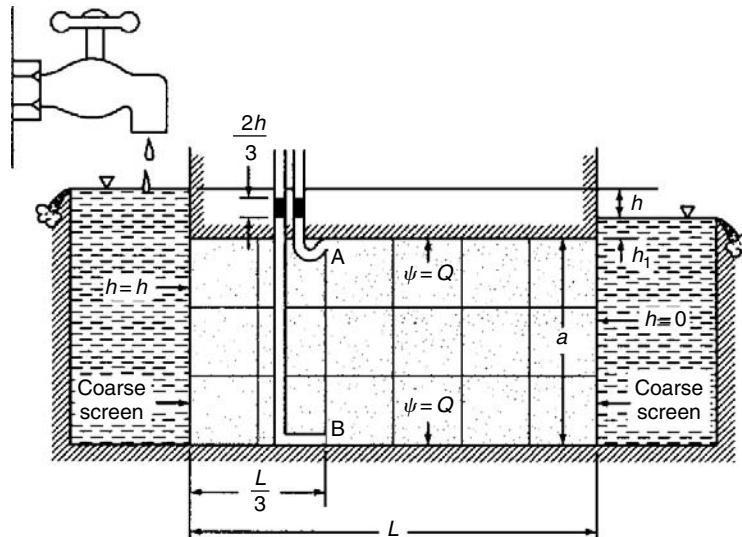


FIGURE 5.15 Example of flow regime.

Example 5.12

Using flow nets, obtain a plot of Q/kh as a function of the ratio s/T for the single impervious sheetpile shown in Figure 5.16a.

Solution. We first note that the section is symmetrical about the y axis, hence only one half of a flow net is required. Values of the ratio s/T range from 0 to 1; with 0 indicating no penetration and 1 complete cutoff. For $s/T = 1$, $Q/kh = 0$ (see point a in Figure 5.16b). As the ratio of s/T decreases, more flow channels must be added to maintain curvilinear squares and, in the limit as s/T approaches zero, Q/kh becomes unbounded (see arrow on Figure 5.16b). If $s/T = 1/2$ (Figure 5.16c), each streamline will evidence a corresponding equipotential line in the half-strip; consequently, for the whole flow region, $N_f/N_e = 1/2$ for $s/T = 1/2$ (point b in Figure 5.16b). Thus, without actually drawing a single flow net, we have learned quite a bit about the functional relationship between Q/kh and s/T . If Q/kh were known for $s/T=0.8$ we would have another point and could sketch, with some reliability, the portion of the plot in Figure 5.16b for $s/T \geq 0.5$. In Figure 5.16d is shown one half of the flow net for $s/T = 0.8$, which yields the ratio of $N_f/N_e = 0.3$ (point c in Figure 5.16b). As shown in Figure 5.16e, the flow net for $s/T=0.8$ can also serve, geometrically, for the case of $s/T=0.2$, which yields approximately $N_f/N_e=0.8$ (plotted as point d in Figure 5.16b). The portion of the curve for values of s/T close to 0 is still in doubt. Noting that for $s/T = 0$, $kh/Q = 0$, we introduce an ordinate scale of kh/Q to the right of Figure 5.16b and locate on this scale the corresponding values for $s/T=0, 0.2$, and 0.5 (shown primed). Connecting these points (shown dotted) and obtaining the inverse, Q/kh , at desired points, the required curve can be obtained.

A plot giving the quantity of discharge (Q/kh) for symmetrically placed pilings as a function of depth of embedment (s/T), as well as for an impervious structure of width $(2b/T)$, is shown in Figure 5.17. This plot was obtained by Polubarnova-Kochina (1952) using a mathematical approach. The curve labeled $b/T=0$ applies for the conditions in Example 5.12. It is interesting to note that this whole family of curves can be obtained, with reasonable accuracy,^{*} by sketching only two additional half flow nets (for special values of b/T).

* At least commensurate with the determination of the coefficient of permeability, k .

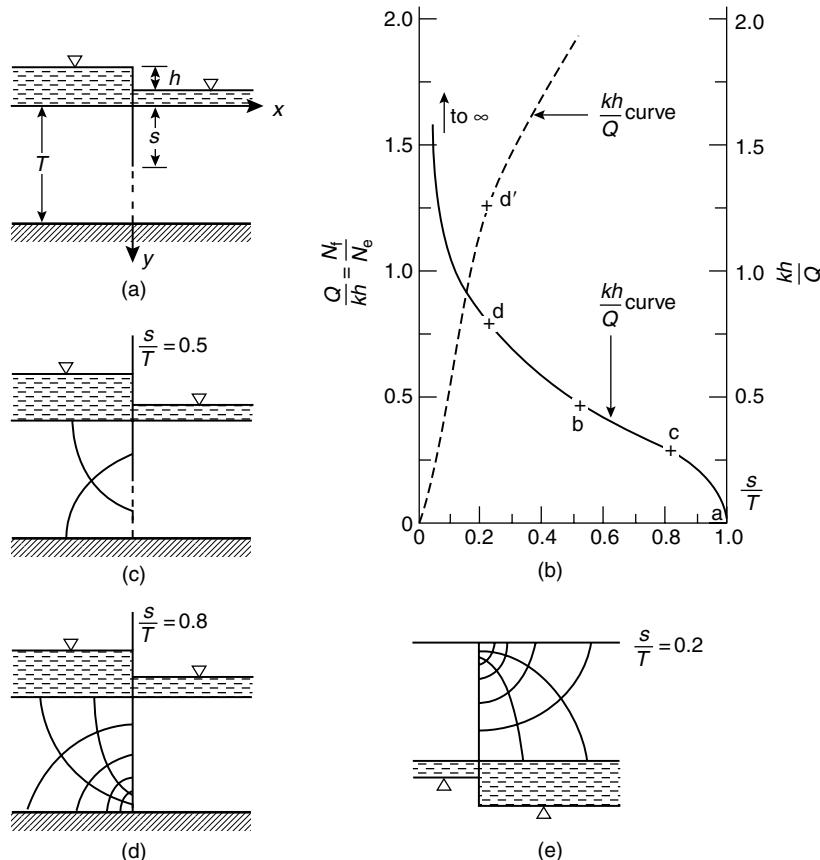


FIGURE 5.16 Example 5.12.

It was tacitly assumed in the foregoing that the medium was homogeneous and isotropic ($k_x = k_y$). Had isotropy not been realized, a transformation of scale* in the y direction of $Y = y(k_x/k_y)^{1/2}$ or in the x direction of $X = x(k_y/k_x)^{1/2}$ would render the system as an equivalent isotropic system (for details, see Harr, 1962, p. 29). After the flow net has been established, by applying the inverse of the scaling factor, the solution can be had for the anisotropic system. The quantity of discharge for a homogeneous and anisotropic section is

$$Q = \sqrt{k_x k_y} h \frac{N_f}{N_e} \quad (5.44)$$

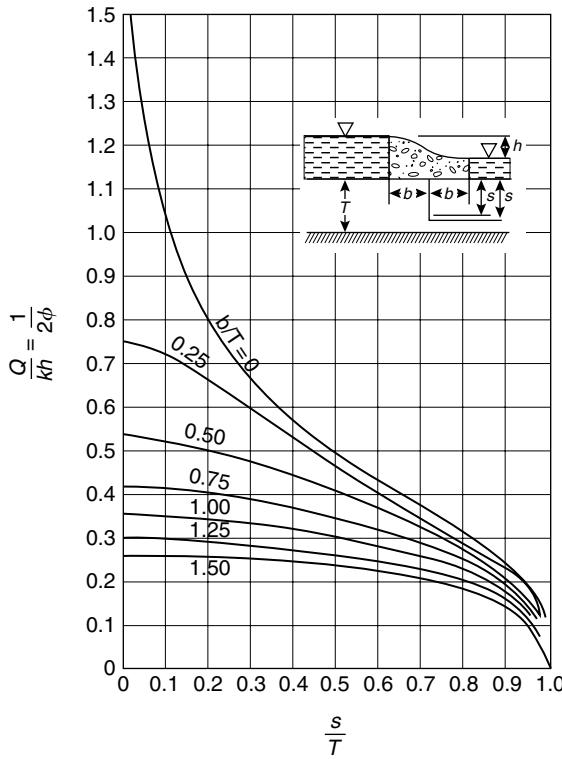
We next introduce some powerful probabilistic tools.

5.5 The Point Estimate Method (PEM)

5.5.1 One Random Variable

Many methods have been offered in the past to accommodate uncertainty. The most common, by far, is to assign single-valued point estimates that reflect central tendencies or implied levels of conservatism. Analyses are then reduced to deterministic treatments. More direct probabilistic methods employ Monte Carlo simulations or truncated Taylor series (Harr, 1987). However, analyses rapidly become exceedingly

*Here x and y are taken as the directions of maximum and minimum permeability.



$$\text{If } m = \cos \frac{\pi s}{2T} \sqrt{\tanh^2 \frac{\pi b}{2T} + \tan^2 \frac{\pi b}{2T}}$$

$$\text{for } m \leq 0.3, \frac{Q}{kh} = \frac{1}{\pi} \ln \frac{4}{m}$$

$$\text{for } m^2 \geq 0.3, \frac{Q}{kh} = \frac{-\pi}{2 \ln \left(\frac{1-m^2}{16} \right)}$$

FIGURE 5.17 Quantity of flow for given geometry.

difficult, if not impossible, for these methods for any but a very few uncorrelated random variables. A simple and very versatile procedure called the *point estimate method* is advocated by this writer and will be developed in some detail. This method, first presented by Rosenblueth (1975) and later extended by him in 1981, has since seen considerable use and expansion by this writer and his coauthors. The methodology is presented in considerable detail in Harr (1987).

Consider $y = p(x)$ to be the probability distribution of the random variable x , with expected value \bar{x} and standard deviation σ . With analogy to Figure 5.7, we replace the load on the beam by *two* reactions, $p(-)$ and $p(+)$ acting at $x(-)$ and $x(+)$, as shown in Figure 5.18.

Pleading symmetry, probabilistic arguments produce for a random variable x :

$$p(+) = p(-) = \frac{1}{2} \quad (5.45a)$$

$$x(+) = \bar{x} + \sigma[x] \quad (5.45b)$$

$$x(-) = \bar{x} - \sigma[x] \quad (5.45c)$$

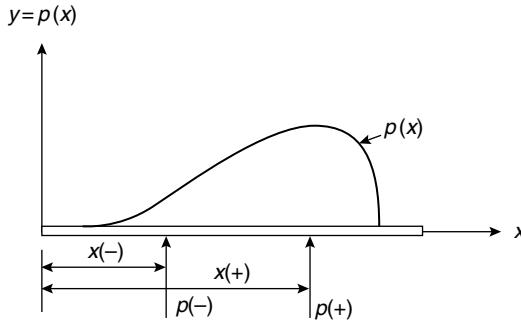


FIGURE 5.18 Point estimate approximations.

With the distribution $p(x)$ approximated by the point estimates $p(-)$ and $p(+)$, the moments of $y = p(x)$ are

$$E[y] = \bar{y} = p(-)y(-) + p(+)y(+) \quad (5.46a)$$

$$E[y^2] = p(-)y^2(-) + p(+)y^2(+) \quad (5.46b)$$

where $y(-)$ and $y(+)$ are the values of the function $p(x)$ at $x(-)$ and $x(+)$, respectively. These reduce to the simpler expressions

$$\bar{y} = \frac{y(+) + y(-)}{2} \quad (5.47a)$$

$$\sigma[y] = \left| \frac{y(+) - y(-)}{2} \right| \quad (5.47b)$$

Example 5.13

Estimate the expected value and the coefficient of variation for the well-known coefficient of active earth pressure $K_A = \tan^2(45 - \phi/2)$, if $E[\phi] = 30^\circ$.

Solution. With the standard deviation of the ϕ -parameter not given, we again return to Table 5.3, $V(\phi) = 12\%$, and $\sigma[\phi] = 3.6^\circ$. Hence, $\phi(+) = 33.6^\circ$, $\phi(-) = 26.4^\circ$. Therefore, $K_A(+) = 0.29$, $K_A(-) = 0.38$, and Equations 5.47 produce $\bar{K}_A = 0.34$, $\sigma[K_A] = 0.05$; hence, $V(K_A) = 13\%$.

Before proceeding with functions of many random variables, we shall first consider regression and correlation.

5.5.2 Regression and Correlation

We next study the functional relationship between random variables called *regression analysis*. It is regression analysis that provides the grist of being able to predict the value of one variable from that of another or of others. The measure of the degree of correspondence within the developed relationship belongs to *correlation analysis*.

Let us suppose we have N pairs of data $[x(1), y(1)], \dots, [x(N), y(N)]$ for which we postulate the *linear* relationship

$$y = Mx + B \quad (5.48)$$

where M and B are constants. Of the procedures available to estimate these constants (including *best fit by eye*), the most often used is the *method of least squares*. This method is predicated on minimizing the

sum of the squares of the distances between the data points and the corresponding points on a straight line. That is, M and B are chosen so that

$$\sum(y - Mx - B)^2 = \text{minimum}$$

This requirement is met by the expressions

$$M = \frac{N \sum xy - \sum x \sum y}{N \sum x^2 - (\sum x)^2} \quad (5.49a)$$

$$B = \frac{\sum x^2 \sum y - \sum x \sum xy}{N \sum x^2 - (\sum x)^2} \quad (5.49b)$$

It should be emphasized that a straight line fit was assumed. The reasonableness of this assumption is provided by the *correlation coefficient* ρ , defined as

$$\rho = \frac{\text{cov}[x, y]}{\sigma[x]\sigma[y]} \quad (5.50)$$

where $\sigma[x]$ and $\sigma[y]$ are the respective standard deviations and $\text{cov}[x, y]$ is their *covariance*. The covariance is defined as

$$\text{cov}[x, y] = \frac{1}{N} \sum_{i=1}^N [x(i) - \bar{x}][y(i) - \bar{y}] \quad (5.51)$$

With analogy to statics the covariance corresponds to the *product of inertia*.

In concept, the correlation coefficient is a measure of the tendency for two variables to vary together. This measure may be zero, negative, or positive; wherein the variables are said to be *uncorrelated*, *negatively correlated*, or *positively correlated*.

The variance is a special case of the covariance as

$$\text{cov}[x, x] = \nu[x] \quad (5.52)$$

Application of their definitions produces (Ditlevsen, 1981) the following identities (a, b, c are constants):

$$E[a + bx + cy] = a + bE[x] + cE[y] \quad (5.53a)$$

$$\nu[a + bx + cy] = b^2\nu[x] + c^2\nu[y] + 2bc \text{cov}[x, y] \quad (5.53b)$$

$$\text{cov}[x, y] \leq \sigma[x]\sigma[y] \quad (5.53c)$$

$$\nu[a + bx + cy] = b^2\nu[x] + c^2\nu[y] + 2bc\sigma[x]\sigma[y]\rho \quad (5.53d)$$

Equation 5.53c demonstrates that the correlation coefficient, Equation 5.50, must satisfy the condition

$$-1 \leq \rho \leq +1 \quad (5.54)$$

If there is perfect correlation between variables and in the same sense of direction, $\rho = +1$. If there is perfect correlation in opposite directions (one variable increases as the other decreases), $\rho = -1$. If some scatter exists, $-1 < \rho < +1$, with $\rho = 0$ if there is no correlation. Some examples are shown in Figure 5.19.

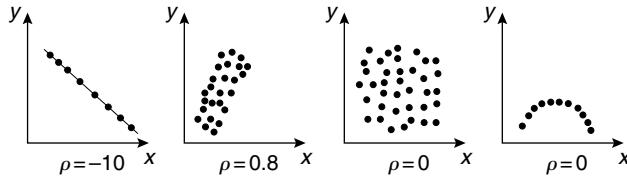


FIGURE 5.19 Example of scatter and correlation coefficients.

TABLE 5.6 Example of Extraction of Moments and Correlation Coefficients

Test	x_1	x_2	x_3	dx_1	dx_2	dx_3	ds_1	ds_2	ds_3
1	2	1	2	-5.25	-6.67	-4.92	-1.92	-1.85	-1.29
2	4	3	1	-3.25	-4.67	-5.92	-1.18	-1.30	-1.55
3	6	7	5	-1.25	-0.67	-1.92	-0.46	-0.19	-0.50
4	7	6	3	-0.25	-1.67	-3.92	-0.09	-0.46	-1.02
5	5	4	7	-2.25	-3.67	0.08	-0.82	-1.02	0.02
6	8	9	6	0.75	1.33	-0.92	0.27	0.37	-0.25
7	9	8	7	1.75	0.33	0.08	0.64	0.09	0.02
8	7	10	9	-0.25	2.33	2.08	-0.09	0.65	0.55
9	10	11	11	2.75	3.33	4.08	1.01	0.93	1.07
10	8	9	8	0.75	1.33	1.08	0.27	0.37	0.28
11	12	11	10	4.75	3.33	3.08	1.74	0.93	0.81
12	9	13	14	1.75	5.33	7.08	0.64	1.48	1.85
Expected value	7.25	7.67	6.92	0	0	0	0	0	0
Standard deviations	2.73	3.60	3.82	2.73	3.60	3.82	1	1	1

We shall next present an extension of the above correlation methodology that will accommodate very many random variables.

Suppose that 12 tests are performed and sample measurements of the three variables x_1 , x_2 , and x_3 yield the values in columns so labeled in Table 5.6. Also shown are the sample expected values (\bar{x}_i) obtained as

$$\bar{x}_i = E[x_i] = \frac{1}{12} \sum_{i=1}^{12} x_i, \quad i = 1, 2, 3 \quad (5.55)$$

Columns labeled dx_1 , dx_2 , and dx_3 are *mean-corrected*; that is, the respective expected values have been subtracted from the raw data,

$$dx_i = x_i - \bar{x}_i, \quad i = 1, 2, 3 \quad (5.56)$$

Designating the 12×3 matrix composed of the three columns (dx_1 , dx_2 , dx_3) as D , the *mean-corrected matrix*, and performing the matrix multiplication, D' is the transpose of the matrix D , produces the symmetrical *covariance matrix*, C :

$$C = \frac{1}{11} D'D = \begin{bmatrix} 7.48 & 8.73 & 7.84 \\ 8.73 & 12.97 & 12.24 \\ 7.84 & 12.24 & 14.63 \end{bmatrix} \quad (5.57a)$$

The elements on the principal diagonal, 7.48, 12.97, 14.63, are the respective variances of the variables x_1 , x_2 , and x_3 . Hence, the respective standard deviations are 2.73, 3.60, and 3.82.

$$\sigma_1 = 2.73, \quad \sigma_2 = 3.60, \quad \sigma_3 = 3.82 \quad (5.57b)$$

The off-diagonal elements are the respective covariances; for example, $8.73 = \sigma_1\sigma_2\rho_{1,2}$, where $\rho_{1,2}$ is the correlation coefficient between the variables x_1 and x_2 . Hence, $\rho_{1,2} = 8.73/(2.73 \times 3.60) = 0.89$. In general, dividing a row and column of the covariance matrix by the corresponding standard deviations produces the symmetrical *correlation matrix*, K , which for the example at hand becomes

$$K = \begin{bmatrix} 1.00 & 0.89 & 0.75 \\ 0.89 & 1.00 & 0.89 \\ 0.75 & 0.89 & 1.00 \end{bmatrix} \quad (5.58)$$

It should be noted that the correlation matrix could also have been obtained directly from the raw data, the columns of x_1 , x_2 , and x_3 in Table 5.6. The respective standard deviations (σ_i) are given by (Note: $12 - 1 = 11$):

$$\sigma_i^2 = \frac{1}{11} \sum_{i=1}^{12} (x_i - \bar{x})^2, \quad i = 1, 2, 3 \quad (5.59)$$

The last three columns of Table 5.6, labeled ds_1 , ds_2 , and ds_3 are *standardized variates*, obtained by subtracting the expected values and dividing by the standard deviations, respectively; that is

$$ds_i = \frac{x_i - \bar{x}_i}{\sigma_i} \quad (5.60)$$

Then, forming the 12×3 matrix D_s , with the columns the same as the last three columns of Table 5.6, the correlation matrix is obtained as

$$K = \frac{1}{11} D_s' D_s \quad (5.61)$$

Thus, the data in Table 5.6 have been reduced to provide the expected values, standard deviations, and correlation coefficients of the three uncertain variables. It should be apparent that the foregoing concepts can be readily extended to m test results for n variables to produce the $n \times n$ correlation matrix and the respective expected values and standard deviations of the n -variates. The coefficient in Equation 5.59 and Equation 5.61 would be $1/(m - 1)$. Computer software abounds that readily produces these results.

5.5.3 Point Estimate Method — Several Random Variables

Rosenblueth (1975) generalized his earlier methodology to any number of correlated variables. For example, for a function of three random variables; say, $y = y[x(1), x(2), x(3)]$; where $\rho(i, j)$ is the correlation coefficient between variables $x(i)$ and $x(j)$ he obtained:

$$E[y^N] = p(++)y^N(++) + p(+-)y^N(+-) + \dots + p(--y^N--) \quad (5.62a)$$

where

$$y(\pm \pm \pm) = y[\bar{x}(1) \pm \sigma[x1], \bar{x}(2) \pm \sigma[x2], \bar{x}(3) \pm \sigma[x3]] \quad (5.62b)$$

$$p(+++)=p(---)=\frac{1}{2^3}[1+\rho(1,2)+\rho(2,3)+\rho(3,1)]$$

$$p(++-)=p(-+)=\frac{1}{2^3}[1+\rho(1,2)-\rho(2,3)-\rho(3,1)]$$

$$p(+--)=p(-+-)=\frac{1}{2^3}[1-\rho(1,2)-\rho(2,3)+\rho(3,1)]$$

$$p(+--)=p(-++)=\frac{1}{2^3}[1-\rho(1,2)+\rho(2,3)-\rho(3,1)] \quad (5.62c)$$

where $\sigma[xi]$ is the standard deviation of $x(i)$. The sign of $\rho(i,j)$ is determined by the multiplication rule of i and j ; that is, if the sign of $i = (-)$, and of $j = (+)$; $(i)(j) = (-)(+) = (-)$.

Equation 5.62a has $2^3 = 8$ terms, all permutations of the three +’s and –’s. In general, for M variables there are 2^M terms and $M(M - 1)/2$ correlation coefficients, the number of combinations of M objects taken two at a time. The coefficient on the right-hand side of Equation 5.62c, in general, is $(1/2)^M$.

Example 5.14

The recommendation of the American Concrete Institute (Galambos et al., 1982) for the design of reinforced concrete structures is (in simplified form)

$$R \geq 1.6D + 1.9L$$

where R is the strength of the element, D is the dead load, and L is the lifetime live load. If $E[D] = 10$, $E[L] = 8$, $V(D) = 10\%$, $V(L) = 25\%$, $\rho(D, L) = 0.75$ (a) Find the expected value and standard deviation of R for the case $R = 1.6I + 1.9L$. (b) If the results in part (a) generate a normal variate and the maximum strength of the element R is estimated to be 40, estimate the implied probability of failure.

Solution. The solution is developed in Figure 5.20.

Generalizations of the point estimate method to more than three random variables is given by Harr (1987). We next return to groundwater and seepage.

5.6 Method of Fragments

In spite of its many uses, a graphical flow net provides the solution for a singular problem configuration only. Should one wish to investigate the influence of a range of characteristic dimensions (such as is generally the case in a design problem) many flow nets would be required. Consider, for example, the section shown in Figure 5.21, and suppose we wish to investigate the influence of the dimensions A, B, and C on the characteristics of flow, all other dimensions being fixed. Taking only three values for each of these dimensions would require 27 individual flow nets. As noted previously, a rough flow net should always be drawn as a check. In this respect it may be thought of as being analogous to a free-body diagram in mechanics, wherein the physics of a solution can be examined with respect to satisfying conditions of necessity.

An approximate analytical method of solution, directly applicable to design, was developed by Pavlovsky in 1935 (Pavlovsky, 1956) and was expanded and advanced by Harr (1962, 1977). The underlying assumption of the method, called the *method of fragments*, is that equipotential lines at various critical parts of the flow region can be approximated by straight vertical lines (as, for example, the dotted lines in Figure 5.22) that divide the region into sections or fragments. The groundwater flow region in the figure is shown divided into four fragments.

$$R = 1.6D + 1.9L$$

Variable,	\bar{x}	$\sigma[x]$	$x(+)$	$x(-)$
x				
D	10	1	11	9
L	8	2	10	6

$$\rho(D, L) = +0.75$$

	$R(ij)$	$R(ij)^2$
$R(++) :$	36.6	1340
$R(+-) :$	29.0	841
$R(-+) :$	33.4	1116
$R(--)$:	25.8	666

$$p(++) = 1/4(1 + \rho) = 0.44$$

$$E[R] = \bar{R} = \sum R(ij)p(ij)$$

$$p(+-) = 1/4(1 - \rho) = 0.06$$

$$= 0.44(36.6 + 25.8) + 0.06(29.0 + 33.4)$$

$$p(-+) = 1/4(1 - \rho) = 0.06$$

$$= 31.20$$

$$p(--)= 1/4(1 + \rho) = 0.44$$

$$E[R^2] = \sum R(ij)^2 p(ij)$$

$$v[R] = E[R^2] - (E[R])^2 = 1000.06 - (31.20)^2 = 26.62$$

$$= 0.44(1340 + 666) + 0.06(841 + 1116)$$

$$\sigma[R] = 5.16; \quad V(R) = 16.5\%$$

$$= 1000.06$$

$$P_f = P[R \geq 40] = \frac{1}{2} - \Psi\left[\frac{40 - 31.20}{5.16}\right] = \frac{1}{2} - \Psi[1.71] = 0.044 \text{ (Table 5.3)}$$

The exact solution is 0.043

FIGURE 5.20 Solution to Example 5.14.

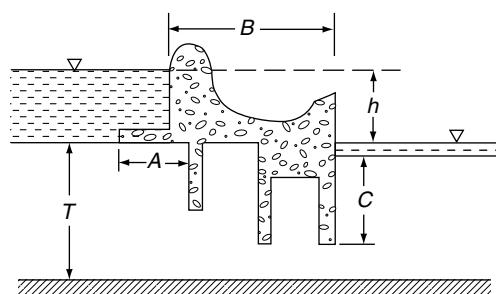


FIGURE 5.21 Example of a complicated structure.

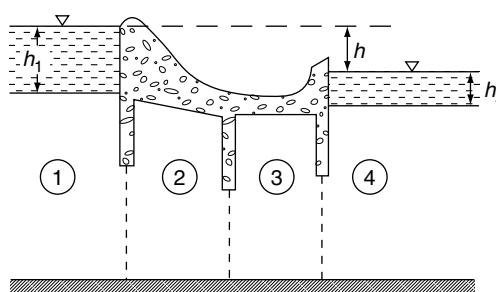


FIGURE 5.22 Four fragments.

Suppose, now, that one computes the discharge in the i th fragment of a structure with m such fragments as

$$Q = \frac{kh_i}{\Phi_i} \quad (5.63)$$

where h_i = head loss through the i th fragment, Φ_i = dimensional form factor in the i th fragment, $\Phi_i = N_e/N_f$ in Equation 5.43. Then, as the discharge through all fragments must be the same,

$$Q = \frac{kh_1}{\Phi_1} = \frac{kh_2}{\Phi_2} = \frac{kh_i}{\Phi_i} = \dots = \frac{kh_m}{\Phi_m}$$

whence

$$\Sigma h_i = \frac{Q}{k} \Sigma \Phi_i$$

and

$$Q = k \frac{h}{\sum_{i=1}^m \Phi_i} \quad (5.64)$$

where h (without subscript) is the total head loss through the section. By similar reasoning, the head loss in the i th fragment can be calculated from

$$h_i = \frac{h\Phi_i}{\Sigma \Phi} \quad (5.65)$$

Thus, the primary task is to implement this method by establishing a catalog of form factors. Following Pavlovsky, the various form factors will be divided into types. The results are summarized in tabular form, Figure 5.23, for easy reference. The derivation of the form factors is well documented in the literature (Harr, 1962, 1977).

Various entrance and emergence conditions for type VIII and IX fragments are shown in Figure 5.24. Briefly, for the entrance condition, when possible, the free surface will intersect the slope at right angles. However, as the elevation of the free surface represents the level of available energy along the uppermost streamline, at no point along the curve can it rise above the level of its source of energy, the headwater elevation. At the point of emergence the free surface will, if possible, exit tangent to the slope (Dachler, 1934). As the equipotential lines are assumed to be vertical, there can be only a single value of the total head along a vertical line, and, hence, the free surface cannot curve back on itself. Thus, where unable to exit tangent to a slope it will emerge vertical.

To determine the pressure distribution on the base of a structure (such as that along C'CC'' in Figure 5.25), Pavlovsky assumed that the head loss within the fragment is linearly distributed along the impervious boundary. Thus, if h_m is the head loss within the fragment, the rate of loss along E'C'CC'E'' will be

$$R = \frac{h_m}{L + s' + s''} \quad (5.66)$$

Once the total head is known at any point, the pressure can easily be determined by subtracting the elevation head, relative to the established (tailwater) datum.

Example 5.15

For the section shown in Figure 5.26a, estimate (a) the discharge and (b) the uplift pressure on the base of the structure.

Fragment type	Illustration	Form factor, Φ (h is head loss through fragment)	Fragment type	Illustration	Form factor, Φ (h is head loss through fragment)
I		$\Phi = \frac{L}{a}$	V		$L \leq 2s: \Phi = 2 \ln(1 + \frac{L}{2a})$ $L \geq 2s: \Phi = 2 \ln(1 + \frac{s}{a}) + \frac{L-2s}{T}$
II		$\Phi = \frac{L}{2} \left(\frac{kh}{Q} \right)$, Fig. 4.17	VI		$L \geq s'+s'': \Phi = \ln[(1 + \frac{s'}{a})(1 + \frac{s''}{a''})] + \frac{L-(s+s'')}{T}$ $L \leq s'+s'': \Phi = \ln[(1 + \frac{b'}{a})(1 + \frac{b''}{a''})]$ where $b' = \frac{L+(s'-s'')}{2}$ $b'' = \frac{L-(s'-s'')}{2}$
III		$\Phi = \frac{L}{2} \left(\frac{kh}{Q} \right)$, Fig. 4.17	VII		$\Phi = \frac{2L}{h_1+h_2}$ $Q = k \frac{h_1-h}{\cot \alpha} \ln \frac{h_d}{h_d-h}$
IV		$b \leq s: \Phi = \ln(1 + \frac{b}{a})$ $b \geq s: \Phi = \ln(1 + \frac{s}{a}) + \frac{b-s}{T}$	VIII		$Q = k \frac{h_1-h}{\cot \alpha} \ln \frac{h_d}{h_d-h}$
			IX		$Q = k \frac{a_2}{\cot \beta} (1 + \ln \frac{a_2+h_2}{a_2})$

FIGURE 5.23 Summary of fragment types and form factors.

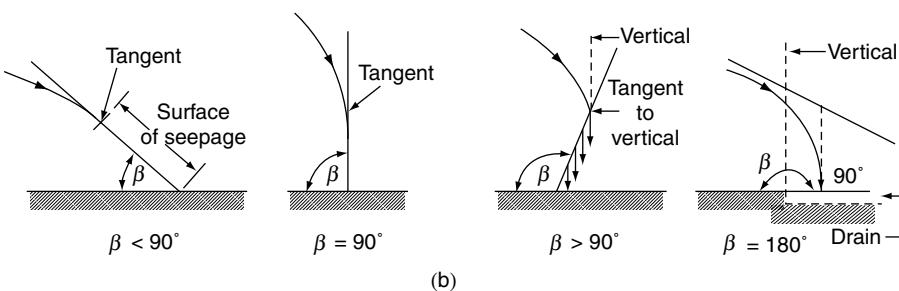
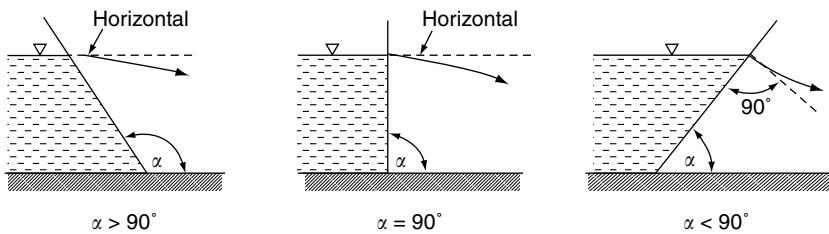
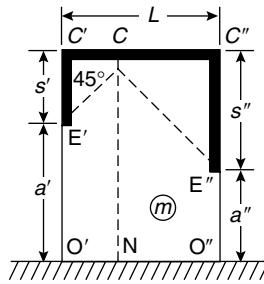
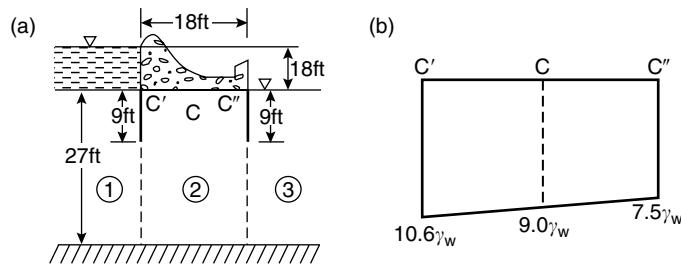


FIGURE 5.24 Entrance and emergence conditions.

**FIGURE 5.25** Illustration of Equation 5.66.**FIGURE 5.26** Example 5.15.

Solution. The division of fragments is shown on the figure. Regions 1 and 3 are both type II fragments, and the middle section is type V with $L = 2s$. For regions 1 and 3, we have, from Figure 5.17, with $b/T = 0$, $\Phi_1 = \Phi_3 = 0.78$.

For region 2, as $L = 2s$, $\Phi_2 = 2 \ln(1 + 18/36) = 0.81$. Thus, the sum of the form factors is

$$\Sigma \Phi = 0.78 + 0.81 + 0.78 = 2.37$$

and the quantity of flow (Equation 5.64) is $Q/k = 18/2.37 = 7.6$ ft.

For the head loss in each of the sections, from Equation 5.64 we find

$$h_1 = h_3 = \frac{0.78}{2.37} (18) = 5.9 \text{ ft}$$

$$h_2 = 6.1 \text{ ft}$$

Hence the head loss rate in region 2 is (Equation 5.65)

$$R = \frac{6.1}{36} = 17\%$$

and the pressure distribution along $C'C'C''$ is shown in Figure 5.26b.

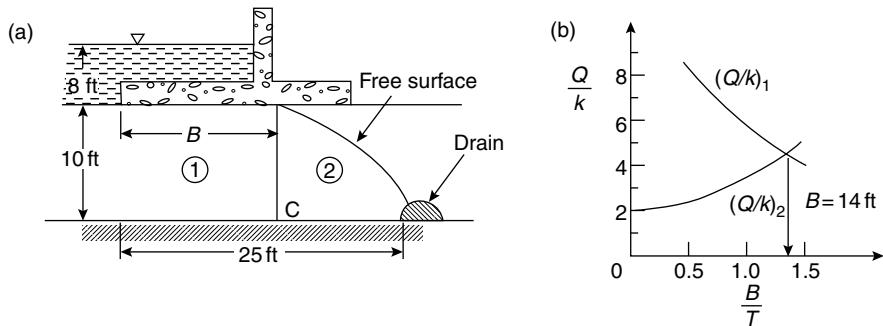


FIGURE 5.27 Example 5.16.

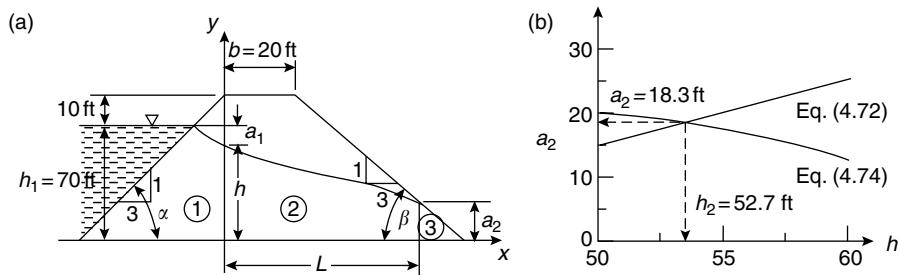


FIGURE 5.28 Example 5.17.

Example 5.16*

Estimate the quantity of discharge per foot of structure and the point where the free surface begins under the structure (point A) for the section shown in Figure 5.27a.

Solution. The vertical line AC in Figure 5.27a is taken as the vertical equipotential line that separates the flow domain into two fragments. Region 1 is a fragment of type III, with the distance B as an unknown quantity. Region 2 is a fragment of type VII with $L = 25 - B$, and $h_1 = 10 \text{ ft}$ and $h_2 = 0$. Thus, we are led to a trial-and-error procedure to find B . In Figure 5.27b are shown plots of Q/k vs. B/T for both regions. The common point is seen to be $B = 14 \text{ ft}$, which yields a quantity of flow of approximately $Q = 100k/22 = 4.5k$.

Example 5.17

Determine the quantity of flow for 100 ft of the earth dam section shown in Figure 5.28a, where $k = 0.002 \text{ ft/min}$.

Solution. For this case, there are three regions. For region 1, a type VIII fragment $h_1 = 70 \text{ ft}$, $\cot \alpha = 3$, $h_d = 80 \text{ ft}$, produces

$$\frac{Q}{k} = \frac{70 - h}{3} \ln \frac{80}{80 - h} \quad (5.67)$$

* For comparisons between analytical and experimental results for mixed fragments (confined and unconfined flow) see Harr and Lewis (1965).

For region 2, a type VII fragment

$$\frac{Q}{k} = \frac{h^2 - a_2^2}{2L} \quad (5.68)$$

With tailwater absent, $h_2 = 0$, the flow in region 3, a type IX fragment, with $\cot \beta = 3$ produces

$$\frac{Q}{k} = \frac{a_2}{3} \quad (5.69)$$

Finally, from the geometry of the section, we have

$$L = 20 + \cot \beta [h_d - a_2] = 20 + 3[80 - a_2] \quad (5.70)$$

The four independent equations contain only the four unknowns, h , a_2 , Q/k , and L , and hence provide a complete, if not explicit solution.

Combining Equations 5.68 and Equation 5.69 and substituting for L in Equation 5.70, we obtain, in general (b = crest width),

$$a_2 = \frac{b}{\cot \beta} + h_d - \sqrt{\left(\frac{b}{\cot \beta} + h_d\right)^2 - h^2} \quad (5.71)$$

and, in particular,

$$a_2 = \frac{20}{3} + 80 - \sqrt{\left(\frac{20}{3} + 80\right)^2 - h^2} \quad (5.72)$$

Likewise, from Equations 5.66 and Equation 5.68, in general,

$$\frac{a_2 \cot \alpha}{\cot \beta} = (h_1 - h) \ln \frac{h_d}{h_d - h} \quad (5.73)$$

and, in particular,

$$a_2 = (70 - h) \ln \frac{80}{80 - h} \quad (5.74)$$

Now, Equations 5.72 and 5.74, and 5.71 and 5.73 in general, contain only two unknowns (a_2 and h), and hence can be solved without difficulty. For selected values of h , resulting values of a_2 are plotted for Equation 5.72 and Equation 5.74 in Figure 5.28b. Thus, $a_2 = 18.3$ ft, $h = 52.7$ ft, and $L = 205.1$ ft. From Equation 5.69, the quantity of flow per 100 ft is

$$Q = 100 \times 2 \times 10^{-3} \times \frac{18.3}{3} = 1.22 \text{ ft}^3/\text{min} = 9.1 \text{ gal}/\text{min}$$

5.7 Flow in Layered Systems

Closed-form solutions for the flow characteristics of even simple structures founded in layered media offer considerable mathematical difficulty. Polubarinova-Kochina (1952) obtained closed-form solutions for the two-layered sections shown in Figure 5.29 (with $d_1 = d_2$). In her solution she found a cluster of parameters that suggested to Harr (1962) an approximate procedure whereby the flow characteristics of structures founded in layered systems can be obtained simply and with a great degree of reliability.

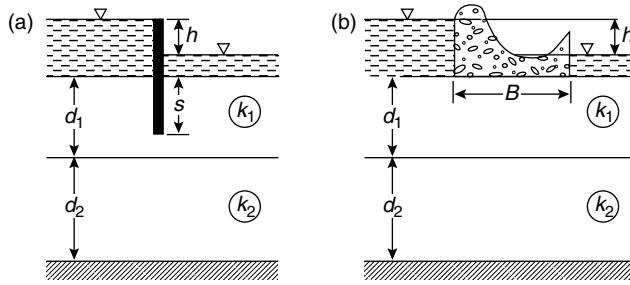


FIGURE 5.29 Example of two-layered systems.

The flow medium in Figure 5.29a consists of two horizontal layers of thickness d_1 and d_2 , underlain by an impervious base. The coefficient of permeability of the upper layer is k_1 and of the lower layer, k_2 . The coefficients of permeability are related to a dimensionless parameter ε by the expression

$$\tan \pi \varepsilon = \sqrt{\frac{k_2}{k_1}} \quad (5.75^*)$$

Thus, as the ratio of the permeabilities varies from 0 to ∞ , ε ranges between 0 and 1/2. We first investigate the structures shown in Figure 5.29a for some special values of ε .

1. $\varepsilon = 0$. If $k_2 = 0$, from Equation 5.75 we have $\varepsilon = 0$. This is equivalent to having the impervious base at depth d_1 . Hence, for this case the flow region is reduced to that of a single homogeneous layer for which the discharge can be obtained directly from Figure 5.17.
2. $\varepsilon = 1/4$. If $k_1 = k_2$, the sections are reduced to a single homogeneous layer of thickness $d_1 + d_2$, for which Figure 5.17 is again applicable.
3. $\varepsilon = 1/2$. If $k_2 = \infty$, $\varepsilon = 1/2$. This represents a condition where there is no resistance to flow in the bottom layer. Hence, the discharge through the total section under steady-state conditions is infinite, or $Q/k_1 h = \infty$. However, of greater significance is the fact that the inverse of this ratio $k_1 h/Q = 0$. It can be shown (Polubarinova-Kochina, 1952), that for $k_2/k_1 \rightarrow \infty$,

$$\frac{Q}{k_1 h} \approx \sqrt{\frac{k_2}{k_1}} \quad (5.76)$$

Thus, we see that for the special values of $\varepsilon = 0$, $\varepsilon = 1/4$, and $\varepsilon = 1/2$ measures of the flow quantities can be easily obtained. The essence of the method then is to plot these values, on a plot of $k_1 h/Q$ vs. ε , and to connect the points with a smooth curve, from which intermediate values can be had.

Example 5.18

In Figure 5.29a, $s = 10$ ft, $d_1 = 15$ ft, $d_2 = 20$ ft, $k_1 = 4k_2 = 1 \times 10^{-3}$ ft/min, $h = 6$ ft. Estimate $Q/k_1 h$.

Solution. For $\varepsilon = 0$, from Figure 5.17 with $s/T = s/d_1 = 2/3$, $b/T = 0$, $Q/k_1 h = 0.39$, $k_1 h/Q = 2.56$.

For $\varepsilon = 1/4$, from Figure 5.17 with $s/T = s/(d_1 + d_2) = 2/7$, $Q/k_1 h = 0.67$, $k_1 h/Q = 1.49$.

For $\varepsilon = 1/2$, $k_1 h/Q = 0$.

The three points are plotted in Figure 5.30, and the required discharge, for $\varepsilon = 1/\Pi \tan^{-1}(1/4)^{1/2} = 0.15$, is $k_1 h/Q = 1.92$ and $Q/k_1 h = 0.52$; whence

$$Q = 0.52 \times 1 \times 10^{-3} \times 6 = 3.1 \times 10^{-3} \text{ ft}^3/(\text{min})(\text{ft})$$

* π is in radian measure.

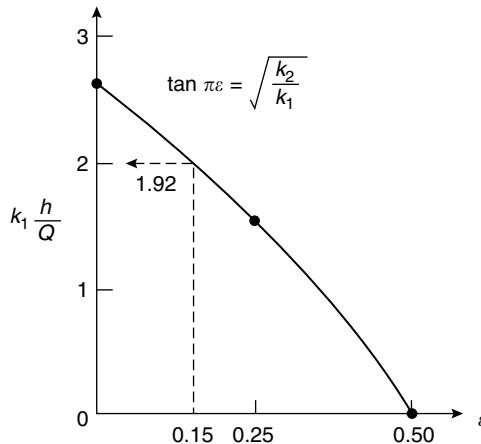


FIGURE 5.30 Example 5.18.

In combination with the method of fragments, approximate solutions can be had for very complicated structures.

Example 5.19

Estimate (a) the discharge through the section shown in Figure 5.31a, $k_2 = 4k_1 = 1 \times 10^{-3}$ ft/day and (b) the pressure in the water at the point P.

Solution. The flow region is shown divided into four fragments. However, the form factors for regions 1 and 4 are the same. In Figure 5.31b are given the resulting form factors for the listed conditions. In Figure 5.31c is given the plot of $k_1 h/Q$ vs. ε . For the required condition ($k_2 = 4k_1$), $\varepsilon = 0.35$, $k_1 h/Q = 1.6$ and hence $Q = (1/1.6) \times 0.25 \times 10^{-3} \times 10 = 1.6 \times 10^{-3}$ ft³/(day)(ft).

The total head at point P is given in Figure 5.31b as Δh for region 4; for $\varepsilon = 0$, $h_p = 2.57$ ft and for $\varepsilon = 1/4$, $h_p = 3.00$. We require h_p for $k_2 = \infty$; theoretically, there is assumed to be no resistance to the flow in the bottom layer for this condition. Hence, the boundary between the two layers (AB) is an equipotential line with an expected value of $h/2$. Thus, $h_p = 10/2 = 5$ for $\varepsilon = 1/2$. In Figure 5.31d is given the plot of h_p vs. ε . For $\varepsilon = 0.35$, $h_p = 2.75$ ft, and the pressure in the water at point P is $(3.75 + 5)\gamma_w = 8.75\gamma_w$. The above procedure may be extended to systems with more than two layers.

Example 5.20

Given the weir section shown in Figure 5.32a. There are four random variables: the soil permeability, the thickness of the pervious layer, and the effective depth of both sheetpiles. The coefficient of variation for the soil permeability will be taken as 100%, with a mean value of 1×10^{-3} cm/sec. A coefficient of variation of 30% will be used for the sand layer thickness. A coefficient of variation of 15% will be assumed for the depth of the sheetpiles. Determine the expected value and standard deviation for the quantity of flow per unit of the section shown.

Solution. AB and CD in Figure 5.32a are taken as equipotential lines; hence, the flow regime is divided into three regions or fragments. In general, with four variables, 16 permutations would be required; however, as the form factors are not functions of soil permeability, only eight form factor terms are necessary for this example.

Region 1 is a type III fragment. This fragment contains both an impervious partial cutoff; say (s_1) and top blanket. The top blanket width of 10 ft is considered to be known in this example (no uncertainty). The form factor is a function of the ratios of b and s_1 to the pervious layer thickness (T). These ratios are entered in Figure 5.17 to determine the term $1/2\phi$ which is then used to calculate ϕ . In the PEM the

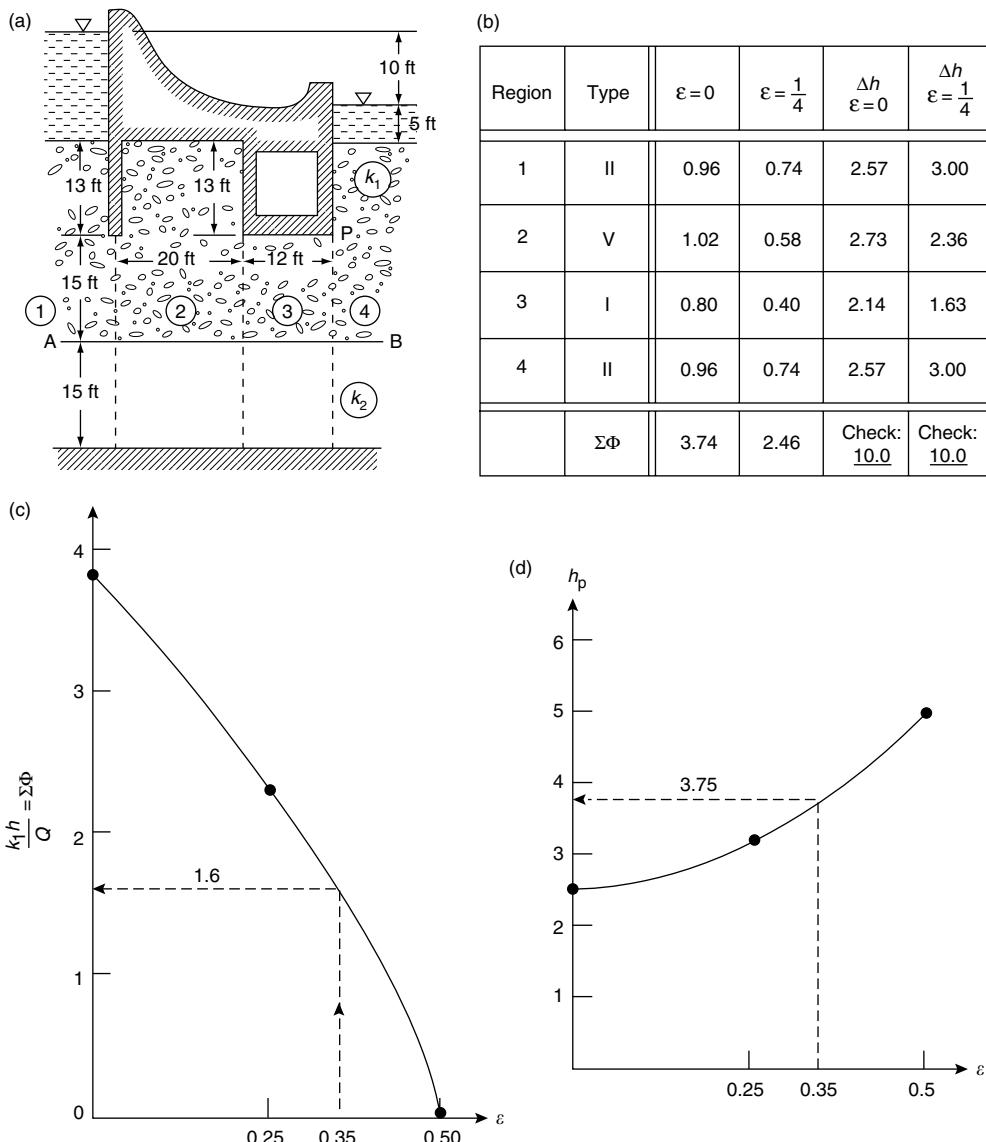
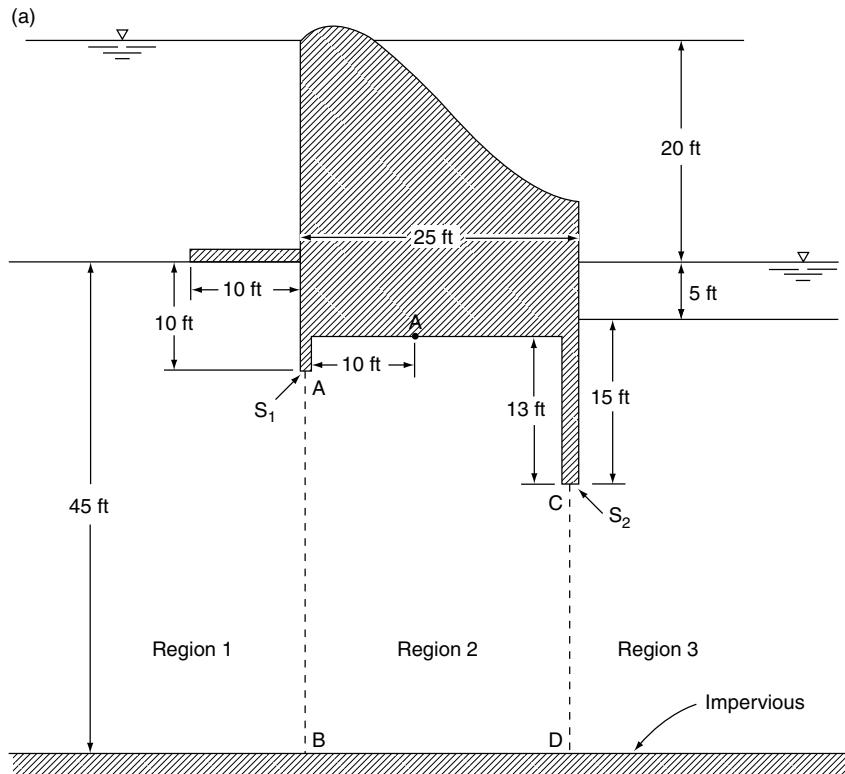


FIGURE 5.31 Example 5.19.

variability in s_1 and T are represented by two plus and minus terms. There are two b/T ratio terms (b/T_+ , and b/T_-) and four s/T ratio terms (s_+/T_+ , s_-/T_+ , s_+/T_- , s_-/T_-) used to determine the four form factors. The corresponding PEM values for Region 1 are shown in Figure 5.32b.

Region 2 is a type VI fragment for which there are three variables. The corresponding PEM plus and minus terms are shown in Figure 5.32c. For this fragment type the equation for the form factor is dependent on the relationship between the length (L) and the sum of the partial cutoffs ($s_1 + s_2$). For the present example the length is greater than the sum of the cutoffs. Thus, the following equation is used (Figure 5.23):

$$\phi = \ln \left[\left(1 + \frac{s_1}{a_1} \right) \left(1 + \frac{s_2}{a_2} \right) \right] + \frac{L - (s_1 + s_2)}{T} \quad (5.77)$$



(b) Type III Fragment

$$b = 10 \text{ ft}$$

$$\bar{s}_1 = 10 \text{ ft}, V_{S_1} = 15\%, \delta_{s_1} = 1.5 \text{ ft}$$

$$S_{1+} = 11.5 \text{ ft } S_{1-} = 8.5 \text{ ft}$$

$$\frac{Q}{kh} = \frac{1}{2\Phi}$$

$$\bar{T} = 45 \text{ ft}, V_T = 30\%, S_T = 13.5 \text{ ft}$$

$$T_+ = 58.5 \text{ ft } T_- = 31.5 \text{ ft}$$

	$\frac{1}{2\Phi}$	$\Phi_{\pm\pm} S_T$
$\frac{b}{T_+} = 0.171$	$\begin{cases} \frac{S_+}{T_+} = \frac{11.5}{58.5} = 0.97 \\ \frac{S_-}{T_+} = 0.145 \end{cases}$.74 0.676 ++ .78 0.641 -+
$\frac{b}{T_-} = 0.317$	$\begin{cases} \frac{S_+}{T_-} = 0.365 \\ \frac{S_-}{T_-} = 0.270 \end{cases}$.52 0.962 +- .57 0.877 --

FIGURE 5.32 (a) Cross section for Example 5.20. (b) Region 1 variables and form factors.

The necessary eight form factors are shown in the figure.

Region 3 is a type II fragment that consists of only a partial cutoff (s_2). The graph in Figure 5.17 is used again to determine the form factor based on the s_2/T ratio. The plus and minus terms of the ratio and the resulting form factors are shown in Figure 5.32d. As there are two random variables, only four form factors are required for this fragment.

(c) Type VI Fragment

$$\begin{aligned}\bar{S}_1 &= 3 \text{ ft}, V_{S_1} = 15\%, \delta_{S_1} = 0.45 \text{ ft} & S_+ &= 3.45 \text{ ft} & S_{1-} &= 2.65 \text{ ft} \\ \bar{a}_1 &= 35 \text{ ft} & a_{1+} &= 34.55 \text{ ft} & a_{1-} &= 35.33 \text{ ft} \\ \bar{S}_2 &= 13 \text{ ft}, V_{S_2} = 15\%, \delta_{S_2} = 1.95 \text{ ft} & S_{2+} &= 14.95 \text{ ft} & S_{2-} &= 11.05 \text{ ft} \\ \bar{a}_2 &= 25 \text{ ft} & a_{2+} &= 23.05 \text{ ft} & a_{2-} &= 26.95 \text{ ft} \\ T &= 38 \text{ ft}, V_T = 30\%, \delta_T = 11.4 \text{ ft} & T_+ &= 49.4 \text{ ft} & T_- &= 26.6 \text{ ft} \\ L &= 25 \text{ ft} \\ L \text{ is always } &> S_1 + S_2\end{aligned}$$

$$\Phi_2 = \ln[(1 + S_1/a_1)(1 + S_2/a_2)] + \frac{L - (S_1 + S_2)}{T}$$

$$\Phi_{+++} = \ln \left[\left(1 + \frac{3.45}{34.55} \right) \left(1 + \frac{14.95}{23.05} \right) \right] + \frac{25 - (3.45 + 14.95)}{49.4} = 0.729$$

$$\Phi_{-++} = 0.722$$

$$\Phi_{+-+} = 0.651$$

$$\Phi_{--+} = 0.645$$

$$\Phi_{++-} = 0.843$$

$$\Phi_{-+-} = 0.851$$

$$\Phi_{+-} = 0.834$$

$$\Phi_{--} = 0.841$$

(d) Type II Fragment

$$\bar{S}_2 = 15 \text{ ft}, V_{S_2} = 15\%, \delta_{S_2} = 2.25 \text{ ft} \quad S_{2+} = 17.25 \text{ ft} \quad S_{2-} = 12.75 \text{ ft}$$

$$\bar{T} = 40 \text{ ft}, V_T = 30\%, \delta_T = 12 \text{ ft} \quad T_+ = 52 \text{ ft} \quad T_- = 28 \text{ ft}$$

$$b = 0 \text{ thus } b/T = 0$$

$$\frac{S_+}{T_+} = 0.332, \frac{1}{2\Phi} = 0.64 \quad \Phi_{++} = 0.781$$

$$\frac{S_-}{T_+} = 0.245, \frac{1}{2\Phi} = 0.73 \quad \Phi_{-+} = 0.685$$

$$\frac{S_+}{T_-} = 0.616, \frac{1}{2\Phi} = 0.41 \quad \Phi_{+-} = 1.220$$

$$\frac{S_-}{T_-} = 0.455, \frac{1}{2\Phi} = 0.53 \quad \Phi_{--} = 0.943$$

FIGURE 5.32 (c) Region 2 variables and form factors. (d) Region 3 variables and form factors.

 $\Sigma\Phi$ Terms

S_1	S_2	T	Φ_1	Φ_2	Φ_3	$\Sigma\Phi_{\pm\pm\pm}$
+	+	+	.676	.729	.781	2.186
-	+	+	.641	.722	.781	2.144
+	-	+	.676	.651	.685	2.012
-	-	+	.641	.645	.685	1.971
+	+	-	.962	.843	1.220	3.025
-	+	-	.877	.851	1.220	2.948
+	-	-	.962	.834	.943	2.739
-	-	-	.877	.841	.943	2.661

FIGURE 5.33 Sum of form factors.

Figure 5.33 shows the summation of the appropriate ϕ terms. Equation 5.64 is then used to calculate the quantity of flow for the sixteen permutations of plus and minus terms required for four random variables. Whereas the coefficient of variation of the soil permeability was taken to be 100%, the plus term is twice the mean value and the minus term is zero (Figure 5.34). Thus, half of the flow values will be zero. The

mean value for the quantity of flow is seen to be 23.62 cubic feet per day with a standard deviation of 24.23 cubic feet per day. This results in a coefficient of variation of 102.6%.

Example 5.21

Obtain the probability distribution of the quantity of flow for the weir in Example 5.20.

Solution. With a coefficient of variation of 102.6%, the normal distribution would lead to the absurdity of negative, $P[Q < 0] = 17\%$, flow quantities! Consequently, the beta distribution will be used with a lower boundary of zero. The upper boundary will be taken to be three standard deviations above the mean. Thus, the bounds for the distribution are 0 to 96.32 cubic feet per day. The resulting distribution is shown in Figure 5.35. This example illustrates the necessity of using the beta distribution rather than the normal distribution when estimating the quantity of flow.

5.8 Piping: Reliability

By virtue of the viscous friction exerted on a liquid flowing through a porous medium, an energy transfer is effected between the liquid and the solid particles. The measure of this energy is the head loss h between the points under consideration. The force corresponding to this energy transfer is called the *seepage force*. It is this seepage force that is responsible for the phenomenon known as *quicksand*, and its assessment is of vital importance in the stability of structures founded in soils and subject to the action of flowing water (seepage).

The first rational approach to the determination of the effects of seepage forces was presented by Terzaghi (1922) and forms the basis for all subsequent studies. Consider all the forces acting on a volume of particulate matter through which a liquid flows.

1. The total weight per unit volume, the mass unit weight is

$$\gamma_m = \frac{\gamma_1(G + e)}{1 + e} \quad (5.78)$$

where e is the void ratio, G is the specific gravity of solids, and γ_1 is the unit weight of the liquid.

2. Invoking Archimedes' principle of buoyancy that *a body submerged in a liquid is buoyed up by force equal to the weight of the liquid displaced*, the effective unit weight of a volume of soil, called the *submerged unit weight*, is

$$\gamma'_m = \gamma_m - \gamma_1 = \frac{\gamma_1(G - 1)}{1 + e} \quad (5.79)$$

To gain a better understanding of the meaning of the submerged unit weight, consider the flow condition shown in Figure 5.36a. If the water column (AB) is held at the same elevation as the discharge face CD($h = 0$), the soil will be in a submerged state and the downward force acting on the screen will be

$$F \downarrow = \gamma'_m LA \quad (5.80)$$

where $\gamma'_m = \gamma_m - \gamma_w$. Now, if the water column is slowly raised (shown dotted to A'B'), water will flow up through the soil. By virtue of this upward flow, work will be done to the soil and the force acting on the screen will be reduced.

3. The change in force through the soil is due to the increased pressure acting over the area or

$$F \uparrow = h \gamma_w A$$

$$Q = \frac{kh}{\Sigma \Phi} \quad \text{Random Variables } k, \Phi(S_1, S_2, T) \\ (\text{thus } 2^4 \text{ or 16 terms})$$

$$h = 20 \text{ ft}$$

$$\bar{k} = 1 \times 10^{-3} \text{ cm/sec} = 2.83 \text{ ft/day}, V_k = 100\%, S_k = 2.83 \text{ ft/day}$$

$$k_+ = 5.66 \text{ ft/day} \quad k_- = 0.0$$

$$E[Q] = \sum_{i=1}^{16} P_{xi} Q_{xi}^2 \quad P_{xi} = \frac{1}{16} \text{ since all terms are taken to be uncorrelated}$$

$$\sigma_Q^2 = \sum_{i=1}^{16} P_{xi} Q_{xi}^2 - (E[Q])^2$$

S_1	S_2	T	k	Q_i	Q_i^2
+	+	+	+	51.950	2698.85
-	+	+	+	52.627	2769.57
+	-	+	+	56.431	3184.43
-	-	+	+	57.258	3278.53
+	+	-	+	37.421	1400.37
-	+	-	+	38.399	1474.48
+	-	-	+	41.329	1708.08
-	-	-	+	42.540	1809.69
+	+	+	-	0.0	0.0
-	+	+	-	0.0	0.0
+	-	+	-	0.0	0.0
-	-	+	-	0.0	0.0
+	+	-	-	0.0	0.0
-	+	-	-	0.0	0.0
+	-	-	-	0.0	0.0
-	-	-	-	0.0	0.0
				$\Sigma = 377.955$	$18324.0 \quad \Sigma/16 = 1145.25$

$$E[Q] = 23.62 \text{ ft}^3/\text{day}$$

$$\sigma_Q^2 = 1145.25 - (23.62)^2 = 587.24$$

$$\sigma_Q = 24.23 \text{ ft}^3/\text{day}$$

$$V_Q = 1.026 \text{ or } 102.6\%$$

FIGURE 5.34 Flow quantities.

Hence, the change in force, granted steady-state conditions, is

$$\Delta F = \gamma'_m LA - h \gamma_w A \quad (5.81)$$

Dividing by the volume AL , the resultant force per unit volume acting at a point within the flow region is

$$R = \gamma'_m - i \gamma_w \quad (5.82)$$

where i is the hydraulic gradient. The quantity $i \gamma_w$ is called the seepage force (force per unit volume). In general, Equation 5.81 is a vector equation, with the seepage force acting in the direction of flow (Figure 5.36b). Of course, for the flow condition shown in Figure 5.36a, the seepage force will be directed vertically upward (Figure 5.36c). If the head h is increased, the resultant force R in Equation 5.81 is seen to decrease. Evidently, should h be increased to the point at which $R = 0$, the material would be at the point of being washed upward. Such a state is said to produce a *quick* (meaning alive) condition. From

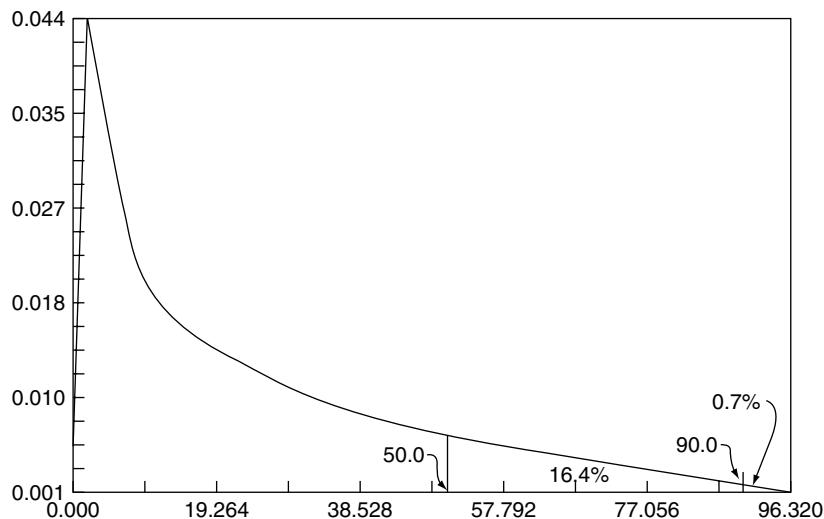


FIGURE 5.35 Beta distribution of flow quantity for Example 5.20.

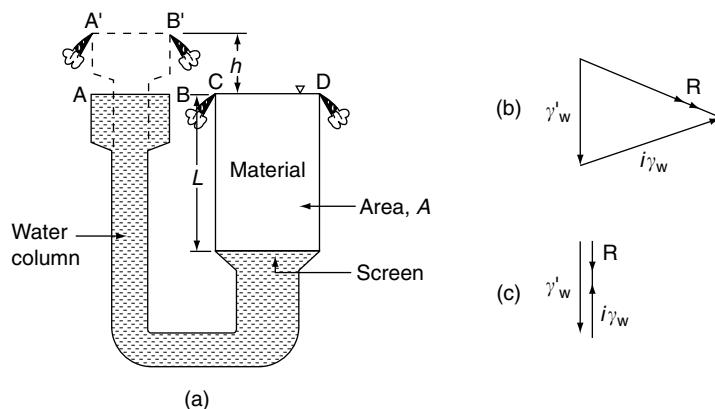


FIGURE 5.36 Development of seepage force concept.

Equation 5.81 it is evident that a quick condition is incipient if

$$i_{cr} = \frac{\gamma'_m}{\gamma_w} = \frac{G - 1}{1 + e} \quad (5.83)$$

Substituting typical values of $G = 2.65$ (quartz sand) and $e = 0.65$ (for sand $0.57 \leq e \leq 0.95$), we see that as an average value the critical gradient can be taken as

$$i_{cr} \approx 1 \quad (5.84)$$

When information is lacking as to the specific gravity and void ratio of the material, the critical gradient is generally taken as unity.

At the critical gradient, there is no interparticle contact ($R = 0$); the medium possesses no intrinsic strength and will exhibit the properties of liquid of unit weight

$$\gamma_q = \left(\frac{G + e}{1 + e} \right) \gamma_1 \quad (5.85)$$

Substituting the above values for G , e , and $\gamma_1 = \gamma_w$, $\gamma_q = 124.8 \text{ lb/ft}^3$. Hence, contrary to popular belief, a person caught in quicksand would not be sucked down, but would find it almost impossible to avoid floating.

Many hydraulic structures, founded on soils, have failed as a result of the initiation of a local quick condition which, in a chain-like manner, led to severe internal erosion called *piping*. This condition occurs when erosion starts at the exit point of a flow line and progresses backward into the flow region, forming pipe-shaped watercourses that may extend to appreciable depths under a structure. It is characteristic of piping that it needs to occur only locally and that once begun it proceeds rapidly, and is often not apparent until structural failure is imminent.

Equation 5.82 and Equation 5.83 provide the basis for assessing the safety of hydraulic structures with respect to failure by piping. In essence, the procedure requires the determination of the maximum hydraulic gradient along a discharge boundary, called the *exit gradient*, which will yield the minimum resultant force (R_{\min}) at this boundary. This can be done analytically, as will be demonstrated below, or from flow nets after a method proposed by Harza (1935).

In the graphical method, the gradients along the discharge boundary are taken as the macrogradient across the contiguous squares of the flow net. As the gradients vary inversely with the distance between adjacent equipotential lines, it is evident that the maximum (exit) gradient is located where the vertical projection of this distance is a minimum, such as at the toe of the structure (point C) in Figure 5.14. For example, the head lost in the final square of Figure 5.14 is one-sixteenth of the total head loss of 16 ft, or 1 ft, and as this loss occurs in a vertical distance of approximately 4 ft, the exit gradient at point C is approximately 0.25. Once the magnitude of the exit gradient has been found, Harza recommended that the factor of safety be ascertained by comparing this gradient with the critical gradient of Equation 5.82 and Equation 5.83. For example, the factor of safety with respect to piping for the flow conditions of Figure 5.14 is $1.0/0.25$ or 4.0. Factors of safety of 4 to 5 are generally considered reasonable for the graphical method of analysis.

The analytical method for determining the exit gradient is based on determining the exit gradient for the type II fragment, at point E in Figure 5.23. The required value can be obtained directly from Figure 5.37 with h_m being the head loss in the fragment.

Example 5.22

Find the exit gradient for the section shown in Figure 5.26a.

Solution. From the results of Example 15.5, the head loss in fragment 3 is $h_m = 5.9 \text{ ft}$. With $s/T = 1/3$, from Figure 5.37 we have $I_E s/h_m = 0.62$; hence,

$$I_E = \frac{0.62 \times 5.9}{9} = 0.41 \quad \text{or} \quad FS = 1/0.41 = 2.44$$

To account for the deviations and uncertainties in nature, Khosla, et al. (1954) recommend that the following factors of safety be applied as critical values of exit gradients; gravel, 4 to 5; coarse sand, 5 to 6; and fine sand, 6 to 7.

The use of reverse filters on the downstream surface, or where required, serves to prevent erosion and to decrease the probability of piping failures. In this regard, the Earth Manual of the U.S. Bureau of Reclamation, Washington, 1974, is particularly recommended.

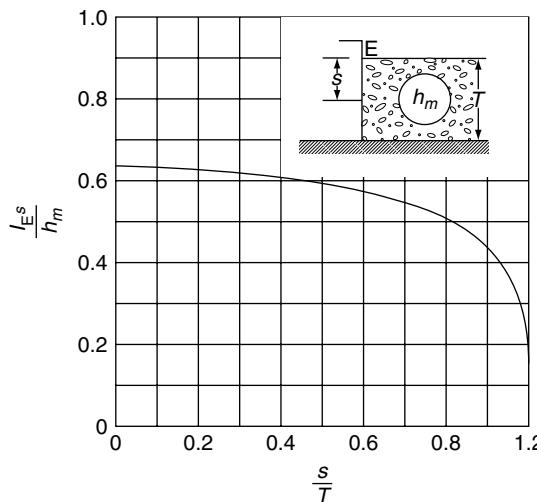


FIGURE 5.37 Exit gradient.

Glossary

Capacity The ability to resist an induced demand; resistance or strength of entity.

Coefficient of Permeability Coefficient of proportionality between Darcy's velocity and the hydraulic gradient.

Correlation Coefficient Measure of the compliance between two variables.

Demand Applied loading or energy.

Expected Value, Expectation Weighted measure of central tendency of a distribution.

Flow Net Trial-and-error graphical procedure for solving seepage problems.

Hydraulic Gradient Space rate of energy dissipation.

Method of Fragments Approximate analytical method for solving seepage problems.

Piping Development of a "pipe" within soil by virtue of internal erosion.

Probability Quantitative measure of a state of knowledge.

Quick Condition Condition when soil "liquefies."

Random Variable An entity whose measure cannot be predicted with certainty.

Regression Means of obtaining a functional relationship among variables.

Reliability Probability of an entity (or system) performing its required function adequately for a specified period of time under stated conditions.

Standard Deviation Square root of variance.

Variance Measure of scatter of variable.

References

- Bear, J. 1972. *Dynamics of Fluids in Porous Media*. American Elsevier Publication Company, New York.
- Bear, J., Zaslavsky, D., and Irmay, S. 1966. *Physical Principles of Water Percolation and Seepage*. Technion, Israel.
- Benjamin, J. R. and Cornell, C. A. 1970. *Probability, Statistics, and Decision for Civil Engineers*. McGraw-Hill, Inc., New York.
- Bercha, F. G. 1978. Application of probabilistic methods in ice mechanics. Preprint 3418, ASCE.
- Borri, A., Ceccotti, A., and Spinelli, P. 1983. Statistical analysis of the influence of timber defects on the static behavior of glue laminated beams, vol. 1, in *4th Int. Conf. Appl. Prob. Soil Struct. Eng.*, Florence, Italy.

- Brahma, S. P. and Harr, M. E. 1962. Transient development of the free surface in a homogeneous earth dam, *Géotechnique*, 12, 4.
- Casagrande, A. 1940. Seepage through dams, "Contributions to Soil Mechanics 1925–1940," Boston Society of Civil Engineering, Boston.
- Cedergren, H. R. 1967. *Seepage, Drainage and Flow Nets*. John Wiley & Sons, New York.
- Dachler, R. 1934. Ueber den Strömungsvorgang bei Hanquellen, *Die Wasserwirtschaft*, 5.
- Darcy, H. 1856. *Les Fontaines Publique de la Ville de Dijon*. Paris.
- Domenico, P. A. and Schwartz, F. W. 1990. *Physical and Chemical Hydrogeology*. John Wiley & Sons, New York.
- Ditlevsen. 1981. *Uncertainty Modeling*. McGraw-Hill, New York.
- Dvinoff, A. H. and Harr, M. E. 1971. Phreatic surface location after drawdown, *J. Soil Mech. Found. Div., ASCE* January.
- Earth Manual*. 1974. Water Resources Technology Publication, 2nd ed., U.S. Department of Interior, Bureau of Reclamation, Washington, DC.
- Edris, E. V. 1990. Probabilistic applications, U.S. Army Corps of Engineers, Waterways Experiment Station, Vicksburg, Mississippi.
- Ellingwood, B., Galambos, T. V., MacGregor, J. G., and Cornell, C. A. 1980. Development of a probability based load criterion for American National Standard A58. *Nat. Bur. Stand. Spec. Publ.* 577, Washington, DC.
- Forchheimer, P. 1930. *Hydraulik*. Teubner Verlagsgesellschaft, mbh, Stuttgart.
- Fredlund, D. G. and Dahlman, A. E. 1972. Statistical geotechnical properties of glacial Lake Edmonton sediments, in *Statistics and Probability in Civil Engineering*. Hong Kong University Press (Hong Kong International Conference), Ed. P. Lumb, distributed by Oxford University Press, London.
- Freeman, H. 1963. *Introduction to Statistical Inference*. John Wiley & Sons, New York.
- Freeze, R. A. and Cherry, J. A. 1979. *Groundwater*. Prentice-Hall, Englewood Cliffs, NJ.
- Galambos, T. V., Ellingwood, B., MacGregor, J. G., and Cornell, C. A. 1982. Probability based load criteria: Assessment of current design practice. *J. Struct. Div., ASCE* 108(ST5).
- Grivas, D. A. and Harr, M. E. 1977. Reliability with respect to bearing capacity failures of structures on ground, in *9th Int. Conf. Soil Mech. Found.* Tokyo, Japan.
- Grivas, D. A. and Harr, M. E. 1979. A reliability approach to the design of soil slopes, in *7th Eur. Conf. on S.M.A.F.E.* Brighton, England.
- Hammitt, G. M. 1966. Statistical analysis of data from a comparative laboratory test program sponsored by ACIL, United States Army Engineering Waterways Experiment Station, Corps of Engineers, Miscellaneous Paper No. 4–785.
- Harr, M. E. 1962. *Groundwater and Seepage*. McGraw-Hill Inc., New York.
- Harr, M. E. and Lewis, K. H. 1965. Seepage around cutoff walls. *RILEM*, Bulletin 29, December.
- Harr, M. E. 1976. Fundamentals of probability theory. *Transp. Res. Rec.* 575.
- Harr, M. E. 1977. *Mechanics of Particulate Media: A Probabilistic Approach*. McGraw-Hill Inc., New York.
- Harr, M. E. 1987. *Reliability-Based Design in Civil Engineering*. McGraw-Hill Inc., New York.
- Harza, L. F. 1935. Uplift and seepage under dams on sand, *Trans. ASCE*, 100.
- Haugen, E. B. 1968. *Probabilistic Approaches in Design*. John Wiley & Sons, New York.
- Khosla, R. B. A. N., Bose, N. K., and Taylor, E. McK. 1954. *Design of Weirs on Permeable Foundations*. Central Board of Irrigation, New Delhi, India.
- Lamb, H. 1945. *Hydrodynamics*. Dover Publications, New York.
- Lipschutz, S. 1965. *Schaum's Outline of Theory and Problems of Probability*. McGraw-Hill Inc., New York.
- Lumb, P. 1966. Variability of Natural Soils, *Can. Geotech. J.*.. vol. 3, May.
- Lumb, P. 1972. Precision and accuracy of soils tests, in *Statistics and Probability in Civil Engineering*, Hong Kong University Press (Hong Kong International Conference), Ed. P. Lumb, distributed by Oxford University Press, London.

- Lumb, P. 1974. Application of statistics in soil mechanics, in *Soil Mechanics —New Horizons*, Ed. I. K. Lee, American Elsevier, New York.
- Muskat, M. 1937. *The Flow of Homogeneous Fluids Through Porous Media*. McGraw-Hill Inc., New York. Reprinted by J. W. Edwards, Ann Arbor, 1946.
- Nielsen, D. R., biggar, J. W., and Erh, K. T. 1973. Spatial Variability of Field Measured Soil-Water Properties, *Hilgardia, J. Agr. Sci.* (Calif. Agr. Experimental stu.), vol. 42, pp. 215–260, Nov.
- Padilla, J. D. and Vanmarcke, E. H. 1974. *Settlement of Structures on Shallow Foundations: A Probabilistic Analysis*, Research Report R74-9, M.I.T.
- Pavlovsky, N. N. 1956. *Collected Works*, Akad. Nauk USSR, Leningrad.
- Pearson, E. S. and Hartley, H. O. 1972. *Biometrika Tables for Statisticians*, Vol. II. Cambridge University Press, London.
- Pearson, K. 1894, 1895. Skew variations in homogeneous material, contributions to the mathematical theory of evolution, *Phil. Trans. R. Soc.*, 185, 186.
- Polubarinova-Kochina, P. Ya. 1941. Concerning seepage in heterogeneous two-layered media, *Inzhenernii Sbornik*, 1, 2.
- Polubarinova-Kochina, P. Ya. 1952. *Theory of the Motion of Ground Water*. Gostekhizdat, Moscow.
- Prasil, F. 1913. *Technische Hydrodynamik*. Springer-Verlag, Berlin.
- Rosenblueth, E. 1975. Point estimates for probability moments, *Proc. Natl Acad. Sci*, USA, 72(10).
- Rosenblueth, E. 1981. Two-point estimates in probabilities, *Appl. Math. Model.*, 5.
- Scheidegger, A. E. 1957. *The Physics of Flow Through Porous Media*. The MacMillan Company, New York.
- Schultze, E. 1972. Frequency distributions and correlations of soil properties, in *Statistics and Probability in Civil Engineering*, Hong Kong University Press (Hong Kong International Conference), Ed. P. Lumb, distributed by Oxford University Press, London.
- Schultze, E. 1975. The General Significance of Statistics for the Civil Engineer, *Proc. 2d Int. Conf. Application of Statistics and Probability in Soil and Structural Engineering* (Aachen, Germany), Sept.
- Terzaghi, K. 1922. Der Grundbruch und Stauwerken und seine Verhütung, *Die Wasserkraft*, 445.

Further Information

- Ang, A. H.-S. and Tang, W. H. 1975. *Probability Concepts in Engineering Planning and Design, Vol. I — Basic Principles*, John Wiley & Sons, New York.
- Guymon, G. L., Harr, M. E., Berg, R. L., and Hromadka, T. V. 1981. A probabilistic-deterministic analysis of one-dimensional ice segregation in a freezing soil column, *Cold Reg. Sci. Technol.* 5, 127–140.
- Hahn, G. J. and Shapiro, S. S. 1967. *Statistical Models in Engineering*. John Wiley & Sons, New York.
- Jayne, E. T. 1978. Where do we stand on maximum entropy? in *The Maximum Entropy Formalism*, Eds. R. D. Levine and M. Tribus, MIT Press, Cambridge, MA.
- Tribus, M. 1969. *Rational Descriptions, Decisions and Designs*, Pergamon Press, New York.
- Whitman, R. Y. 1984. Evaluating calculated risk in geotechnical engineering, *J. Geotech. Eng.*, ASCE 110(2).

6

Soil Properties and Moisture Movement in the Unsaturated Zone

Randel Haverkamp
*Université de Grenoble, Grenoble,
France*

Samuel Debionne
Hydrowide

Pierre Viallet
Hydrowide

Rafael Angulo-Jaramillo
LTHE, CNRS, Grenoble, France

Devaraj de Condappa
IRD, Cotonou, Bénin

6.1	Introduction.....	6-1
6.2	Physical and Hydraulic Soil Properties	6-3
	Soil Phases • Physical Properties of Soils • Hydraulic Properties of Soils • Functional Relationships • Measurement of Soil Characteristics • Estimation Techniques of Soil Characteristics • Spatial Variability of Soil Water Properties	
6.3	Conceptual Aspects of Unsaturated Soil Water Flow..	6-31
	General Flow Equations • Infiltration	
6.4	Analytical Solutions of the Unsaturated Flow Equation	6-34
	Constant Negative (or Zero) Pressure Head Condition at the Soil Surface • Constant Positive Pressure Head Condition at the Soil Surface	
6.5	Scaling Principles	6-40
	Scaling Infiltration • Scaling Evaporation	
6.6	Conclusions and Perspectives on Future Research	6-48
	Glossary and Parameter Notation	6-49
	Acknowledgments.....	6-52
	References	6-52
	Further Reading	6-59

6.1 Introduction

The *unsaturated zone*, sometimes called the *vadose zone*, is the zone between the ground surface and the water table. The term unsaturated zone is somewhat of a misnomer as the capillary fringe above the groundwater or the rain-saturated top soil are portions that are saturated. For this reason some authors (e.g., Bouwer, 1978) prefer the term vadose zone. The unsaturated zone is the hydrological connection between the surface water component of the hydrologic cycle and the groundwater component (Figure 6.1).

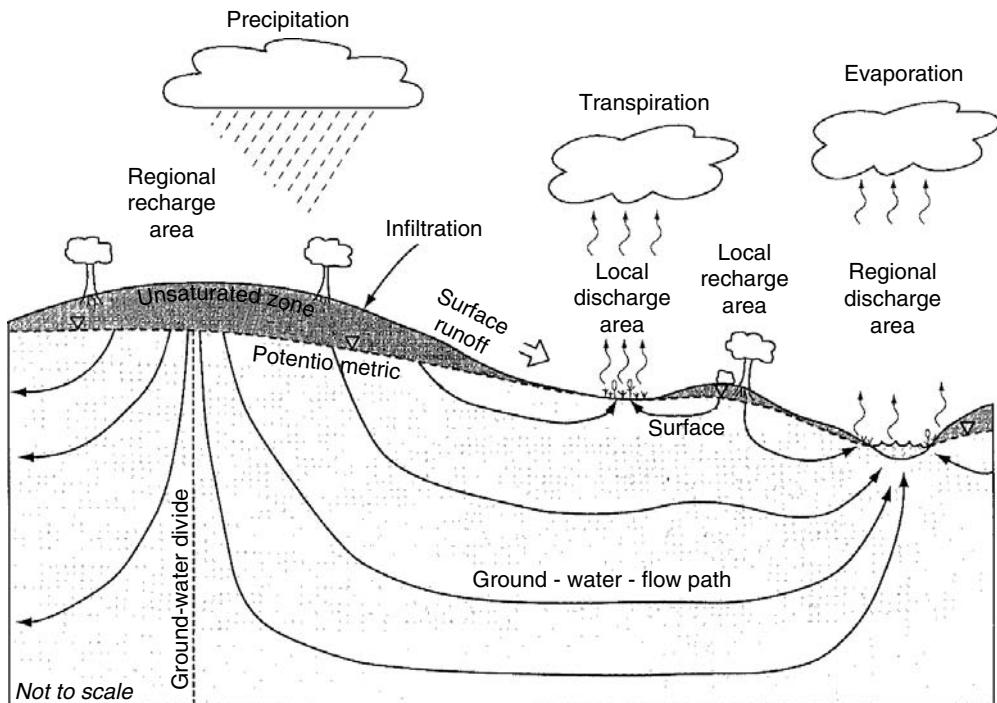


FIGURE 6.1 The unsaturated and saturated zones in the hydrological cycle.

The surface water component includes the precipitation as it reaches the land surface and artificial water application such as irrigation, surface runoff, stream flow, lakes, and artificial impoundments. The rainfall may infiltrate and some of the surface water may percolate through the unsaturated zone. Recharge of the groundwater usually occurs through the unsaturated zone. The unsaturated zone may lose water through evaporation, evapotranspiration, and drainage.

The unsaturated zone plays a crucial role in the transfer of pollutants. Many constituents present in the surface waters eventually find their way into the groundwater through the unsaturated zone. Accident spills of chemicals, application of fertilizers and pesticides on the land surface, leaks from gasoline storage tanks, septic tank drainage, leaching from landfills are examples of anthropogenic activities that contribute to the leaching of contaminants through the unsaturated zone into the groundwater (see Figure 3.22). Chapter 22 develops the multiphase, multi-species equations used to represent the soil, water, and air interactions in the unsaturated zone addressing the principal contaminant sources and pathways in the soil-chemical system.

An introduction to water in the unsaturated zone is given in Section 2.5. Chapter 6 is concerned with the movement of moisture in the unsaturated zone. It covers the conceptual aspects of unsaturated soil water flow related to soil physical properties and soil water characteristics. In Section 6.2 we deal with the soil physical properties, the soil hydraulic properties, and the choice of the functional relationships used to describe the soil hydraulic characteristics. Additionally, we also present some of the measurement and prediction techniques used for the estimation of the soil hydraulic characteristics parameters. We then address the conceptual aspects of the different unsaturated soil water transfer equations used in subsurface hydrology (Section 6.3). An overview of the most significant analytical solutions is presented in Section 6.4; the empirical infiltration laws are discussed in detail in Chapter 7. Section 6.5 sets out the details of a new scaling approach of the water transfer equation (Richards, 1931) and its implications for the soil characteristic scale parameters. Concluding remarks are given in Section 6.6, followed by the references, and a glossary and parameter notation.

6.2 Physical and Hydraulic Soil Properties

6.2.1 Soil Phases

Soils comprise three phases: the solid, the liquid, and the gaseous phases. The *solid phase* includes the mineral particles of gravel, sands, silts, and clays. Gravel, sand, and silt have a dense spherical or angular shape. Clays, on the other hand, have a plate-like structure (with dimensions between 1 and 2 μm) and their capability to adsorb sodium or calcium controls whether they are disperse or flocculent. Particle-size properties are determined from the size distribution of individual particles in a soil sample. Soil particles smaller than 2 mm are divided into the three soil texture groups: sand, silt, and clay. Particles larger than 2 mm are grouped into the gravel class. Figure 6.2 shows the particle-size limits according to the U.S. Department of Agriculture (USDA) (Soil Survey Laboratory Staff, 1992) and the International Soil Science Society (Yong and Warkentin, 1966). Figure 6.3 shows the USDA soil textural triangle that allows for the classification of soils as a function of their particle-size distribution. For example, a soil consisting of 30% sand, 40% silt, and 30% clay by weight is a clay loam. Another grouping is the Unified Soil Classification System (USCS). It is based on particle size, the plasticity index, and the liquid limit and is described in geotechnical engineering texts, such as Lambe and Whitman (1979). The mineral particles, along with the interconnecting pores, form the *soil matrix*. The solid phase may also include organic material such as peat.

The fluids in the soil comprise liquids and gases. The *liquid phase* usually consists of water that can move through the pores of the soil. The water can include numerous dissolved minerals and organic

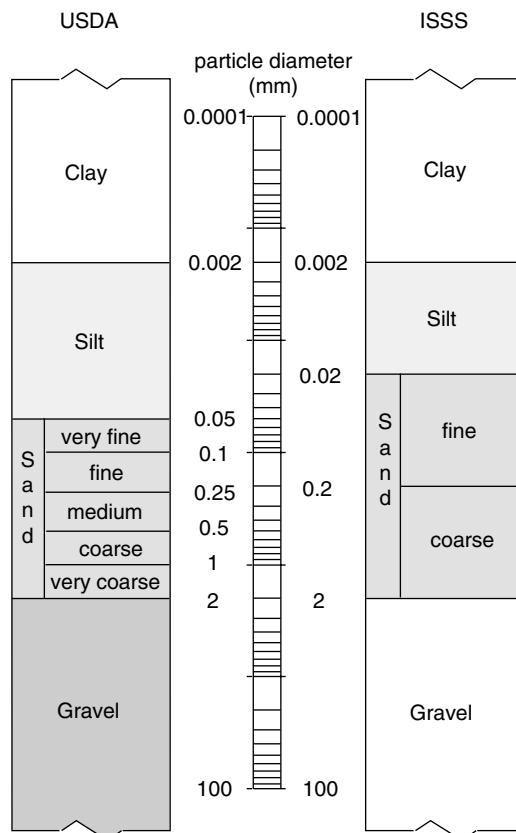


FIGURE 6.2 Particle-size limits according to the USDA and ISSS soil classification schemes.

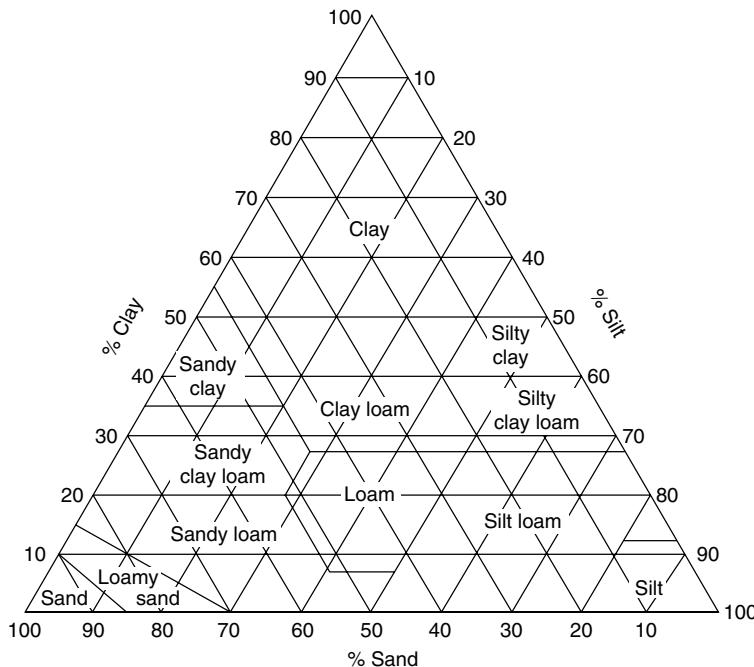


FIGURE 6.3 USDA soil textural classification chart showing the percentages of clay (below 2 μm), silt (2 to 50 μm), and sand (50 to 2000 μm).

compounds. Other liquids may be present. They may be miscible or immiscible in water and generally result from agricultural and industrial activities or accidental spills.

The principal component of the *gaseous phase* is air. Like water, air can also move through the soil and contains other dissolved gases like water vapor and volatile components, usually of pollutants. The principal characteristics of the unsaturated zone depend on the proportion of liquid and gas in the soil matrix. In this three phase system, water and air can flow simultaneously.

6.2.2 Physical Properties of Soils

Let V_s , V_l , V_g , V_t , M_s , M_l , M_g , M_t and W_s , W_l , W_g , W_t be the volumes, masses, and weights of solid, liquid, gas, and total, respectively. Then the density of the solid particles (ρ_s), the dry bulk density (ρ_d), and the porosity (ε) are given by:

$$\rho_s = \frac{M_s}{V_s} = \frac{W_s}{gV_s} \quad (6.1)$$

$$\rho_d = \frac{M_s}{V_t} = \frac{W_s}{gV_t} \quad (6.2)$$

and

$$\varepsilon = \frac{V_l + V_g}{V_t} = \frac{V_t - V_s}{V_t} = 1 - \frac{\rho_d}{\rho_s} \quad (6.3)$$

where g is gravitational acceleration ($\approx 9.82 \text{ m/sec}^2$). The particle density (ρ_s) is normally assumed to be equal to 2.65 g/cm^3 . The void ratio (e) is useful when V_t is not constant due to swelling or shrinking and

is defined as:

$$e = \frac{V_l - V_g}{V_s} = \frac{\varepsilon}{1 - \varepsilon} \quad (6.4)$$

A detailed description of the matrix pore space and associated macropores is given in Chapter 22.

If the definition of Equation 6.3 is applied to a very small elementary volume, the porosity would be equal to one if the element is centered in a pore or would be equal to zero if the element is centered in the solid. The element under consideration has to be equal to or larger than a minimum volume called *representative elementary volume* (REV) to reach a stable value. The continuum approach is used in this chapter. That is to say, an REV is assumed such that the fluid flow can be represented as the product of a transfer coefficient and an energy gradient. For a detailed discussion of the concept of the REV, see Bear and Bachmat (1984).

When small stones and porous coarse fragments (diameter >2 mm) occur in the elementary volume, the whole soil dry bulk density (ρ_{ap}) measured over the whole soil sample is different from the fine soil dry bulk density. The correction formula is:

$$\rho_d = \frac{1 - a - b}{1/\rho_{ap} - a/\rho_s - b/\rho_c} \quad (6.5)$$

where ρ_d is the dry bulk density of the fine soil material; ρ_c is the specific weight of the porous coarse fragments; and a and b are the stone and porous rock weight fractions, respectively, of the whole soil sample. The whole soil porosity (ε_{ap}) and the porosity of the porous coarse fragments (ε_c) are then defined as:

$$\varepsilon_{ap} = 1 - \frac{\rho_{ap}}{\rho_s} \quad (6.6)$$

and

$$\varepsilon_c = 1 - \frac{\rho_c}{\rho_s} \quad (6.7)$$

Dry bulk density and soil porosity are important soil morphological parameters as they are closely related to soil structure. Hence, they influence the soil water transfer properties to a large extent.

6.2.3 Hydraulic Properties of Soils

Regardless of the scales involved, the soil hydraulic properties that affect the flow behavior are incorporated into two fundamental characteristics: (i) the soil water retention curve describing the relation between volumetric soil water content and soil water pressure; and (ii) the hydraulic conductivity function describing the relation between volumetric water content and hydraulic conductivity.

6.2.3.1 Soil Water Content

The soil water content can be expressed by mass (w) or by volume (θ):

$$w = \frac{M_l}{M_s} = \frac{W_l}{W_s} \quad (6.8)$$

and

$$\theta = \frac{V_l}{V_t} = \frac{w\rho_d}{\rho_w} \quad (6.9)$$

where ρ_w is the specific density of water ($\rho_w \approx 1 \text{ g/cm}^3$). In most hydrologic applications, volumetric soil water content is used in nondimensional form:

$$\theta^* = \frac{\theta - \theta_r}{\theta_s - \theta_r} \quad (6.10)$$

Conventionally, θ_s is referred to as the *volumetric soil water content at natural saturation* and θ_r as the *residual volumetric soil water content*. Due to air entrapment, θ_s seldom reaches saturation of the total pore space (ε), for example, Rogowski (1971) gives $\theta_s = 0.9\varepsilon$. Its value is soil-specific and physically well-defined (Haverkamp et al., 2005; Leij et al., 2005). On the contrary, the physical meaning of θ_r is ambiguous. Conceptually, the residual water content may be associated with the immobile water held (by adsorptive forces) within a dry soil profile in films on particle surfaces, in interstices between particles, and within soil pores. In practice, however, its value is generally estimated by fitting the water retention equation to measured data points reducing θ_r to an empirical fitting parameter restricted to the range of data points used. Recently, Haverkamp et al. (2002) explained the residual volumetric water content through the theory of *hysteresis* where θ_r is related to the wetting and drying history prior to the measurement of the water content data points. Most field studies presented in the literature choose a zero θ_r -value (e.g., Kool et al., 1987; van Genuchten et al., 1991; Leij et al., 1996).

When volumetric water content is expressed as the ratio to total pore space (ε), it is referred to as the degree of saturation (S):

$$S = \frac{V_l}{V_l + V_g} = \frac{\theta}{\varepsilon} \quad (6.11)$$

6.2.3.2 Soil Water Pressure

Matric potential or *capillary potential* (ϕ_c) is the measure of the energy status of water retained in a soil by capillarity and surface adsorption (see also Chapter 2, Section 2.4). It is one of the three prevalent components of *total soil water potential* (Φ) given by:

$$\Phi = \phi_g + \phi_c + \phi_o \quad (6.12)$$

where ϕ_g is the *gravitational potential* and ϕ_o is the *osmotic potential* (e.g., van Bavel, 1969). The total soil water potential (Φ) measures the energy per unit quantity (volume, mass, or weight) that is required to move a small volume of water from a pool of water at a datum elevation and at standard atmospheric pressure, to the soil at the point under consideration. The sum of the matric potential and the gravitational potential is often referred to as the *hydraulic potential*. When there is a density gradient of dissolved substances it is necessary to add an *osmotic potential*. Following the International Society of Soil Science (Bolt, 1976), two other components can be included: the *air potential* (ϕ_a), and the *envelope potential* (ϕ_e). The air potential accounts for the difference in pressure that can exist between the air inside the pores and the outside atmospheric pressure over the reference water. When the external pressure due to the overburden is transmitted by the water envelope around the particles, the envelope potential can be significant.

The potential can be expressed per unit volume, per unit mass, or per unit weight. If the potential is expressed in terms of volume, the units are joules per cubic meter or equivalently in units of pressure: Newtons per square meter or Pascals. If the potential is expressed per unit mass or per unit weight, the units are joules per kilogram and meters, respectively. The latter option is conventionally used for most hydrologic studies. When taken per unit weight, the gravitational potential corresponds to the elevation head and the *capillary* or *matric potential* is related to the *soil water pressure head* (h) by:

$$\phi_c = \rho_w gh \quad (6.13)$$

The pressure head h is positive below the water table and negative above it. In the unsaturated zone, the negative of the pressure is also called the *matric suction* or *tension* which, therefore, is expressed as a positive number. A comprehensive description of the different notations is given by Kutilek and Nielsen (1994). In dry soils the pressure head can reach extremely high values; for computational convenience its value is often expressed in the logarithmic mode:

$$pF = \log |h| \quad (6.14)$$

The value of $pF = 4.2$ (i.e., a pressure head of 15,000 cm of water) is the threshold pressure value below which plant transpiration can take place. This point which is referred to as the *wilting point* is sometimes associated with the volumetric soil water content value θ_r (van Genuchten, 1980).

6.2.3.3 Water Retention Characteristic

The variation of volumetric soil water content (θ) with soil water suction (h) is referred to as the *water retention characteristic* ($h(\theta)$). It describes the soil's ability to store or release water. The water retention characteristic is a highly nonlinear S-shaped curve. Figure 6.4 gives an illustration of the typical behavior of the water retention characteristic. It clearly shows that the water retention curve varies with soil structure. For sandy soils the shape of $h(\theta)$ is characterized by a typical step form; for clay soils, on the other hand, the water retention curve is rather steep. Hence, for identical pressure head values clay soils measure generally higher water content values than sandy soils.

The characteristic curves may be different for drying (desorbing) and wetting (absorbing) soils. This phenomenon, referred to as *hysteresis*, can schematically be explained by the "ink bottle effect," that is, the fact that the opening radius r at the top of any pore is generally smaller than the radius R of the main pore. Following Laplace's equation the capillary rise in a tube of radius r is known to be inversely proportional to the radius:

$$h = \frac{2\sigma_{LV} \cos(\omega)}{rg\rho_w} \quad (6.15)$$

where σ_{LV} is the liquid–vapor surface tension of water; and ω is the contact angle between liquid and solid (see also Chapter 2, Equation 2.7). At 20°C, with ρ_w equal to 1 g/cm³ and ω equal to zero, Equation 6.15 is reduced to:

$$h \approx \frac{0.149}{r} \quad (6.16)$$

where soil water suction (h) and pore radius (r) are expressed in cm (e.g., Rose, 1966). Upon drying, the pore remains full until the capillary rise surpasses $0.149/r$. For wetting, the water starts filling the pore when h drops to the lower value $0.149/R$. The water content of the drying pore thus surpasses that of the wetting pore causing the hysteresis. Theoretically, for a given soil geometry, it should be possible to predict hysteretic effects from first principles. However, this problem remains largely unsolved and our rather sketchy understanding of soil structure suggests that only a few soil models will yield to this approach. Instead, hysteresis in soils remains largely based on Poulovassilis' (1962) analysis obtained by applying the independent domain theory to soils (see also Topp, 1971; Mualem, 1974). Parlange (1976, 1980) proposed an alternative theory that has proven to be precise and robust in spite of its operational simplicity. It requires only one boundary curve of the hysteresis envelope to predict the other boundary and all scanning curves in between (see also Hogarth et al., 1988; Liu et al., 1995; Parlange et al., 1999). In a recent study, Haverkamp et al. (2002) proposed a useful extension of the Parlange prediction model based on theory and geometric scaling. Considering the water retention curve to be described by three parameters, where the first defines the shape of the curve and the two others scale the soil water pressure head and volumetric soil water content, the authors derived three geometrical scaling conditions. The first condition determines a shape parameter that is identical for all wetting and drying curves and

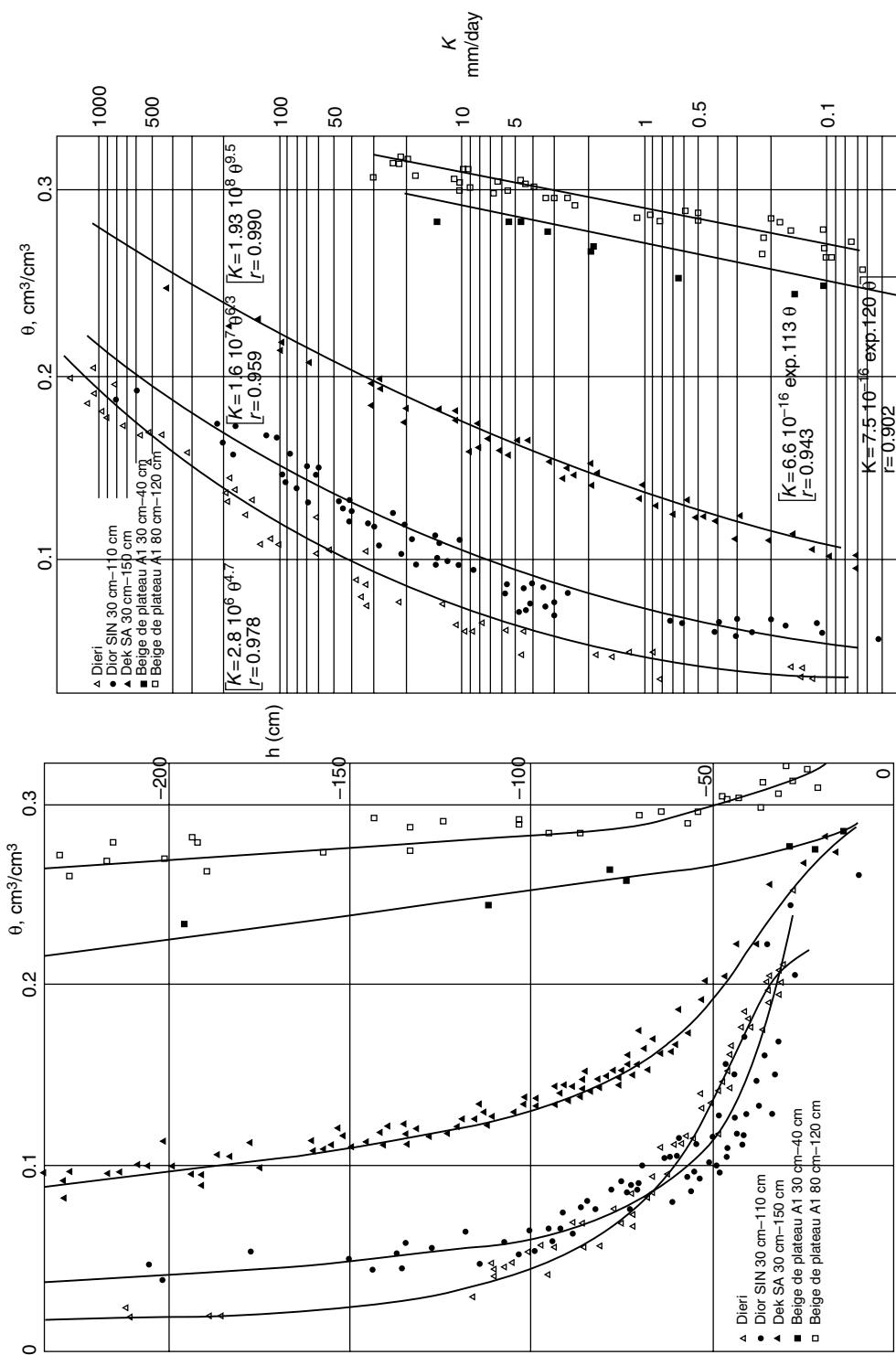


FIGURE 6.4 Water retention and hydraulic conductivity curves of five different soil types measured in Senegal. (Adapted from Hamon, G. 1980. Mise en oeuvre et critique de méthodes de caractérisation hydrodynamique de la zone non-saturée du sol. Applications aux sols de culture du Sénégal. *Thèse Docteur-Ingénieur*, Université Scientifique et Médicale de Grenoble, France. With permission.)

independent of their scanning order; the second defines the relation between the pressure head scale and the water content scale specific to each curve in wetting or drying; and the third condition determines specific water content scale parameters according to the points of departure and arrival of each scanning curve. The authors also derived equations necessary for the calculation of the different scale parameters. Given the saturated water content, all main, primary, and higher order scanning curves can hence be predicted from knowledge of only one curve, although in practice knowledge of a main or primary curve is desirable. Examples of four hysteresis-affected water retention curves measured in the laboratory are shown in Figure 6.5. The predicted main and primary drying curves are compared with measured data.

Hysteresis is more pronounced for sands than for clay soils, particularly for sands with low initial water content profiles prior to wetting (e.g., Vachaud and Thony, 1971). For convenience, hysteresis is usually ignored because its influence under field conditions is often masked by heterogeneities and spatial variability. However, many authors (e.g., Nielsen et al., 1986; Parker and Lenhard, 1987; Russo et al., 1989; Heinen and Raats, 1997; Otten et al., 1997; Whitmore and Heinen, 1999; Si and Kachanoski, 2000) have shown it to be important in simulations of water transfer, solute transport, multiphase flow, and microbial activities, and to disregard it leads to significant errors in predicted fluid distributions with concomitant effects on solute transport and contaminant concentrations (e.g., Hoa et al., 1977; Kool and Parker, 1987; Mitchel and Mayer, 1998).

6.2.3.4 Hydraulic Conductivity

The isothermal hydraulic conductivity was originally introduced by Darcy (1856) for saturated soils and extended to unsaturated soils by Buckingham (1907) and Richards (1931). The hydraulic conductivity is a measure of the ability of the soil to route water. From a theoretical view point $K(\theta)$ can be expressed as:

$$K(\theta) = k \frac{\rho_w g}{\mu_w} k_{rw}(\theta) \quad (6.17)$$

where k is intrinsic permeability; $k_{rw}(\theta)$ is relative water permeability (namely the ratio of the unsaturated to the saturated water permeabilities) that varies from 0 for completely dry soils to 1 for fully saturated soils; and μ_w is the water viscosity. Equation 6.17 shows that the soil conductivity depends on the soil matrix (k), the moving fluid (ρ_w and μ_w) and the fluid content in the soil ($k_{rw}(\theta)$). The dependence of K on θ is unaffected by hysteresis. The hydraulic conductivity at or above saturation ($h \geq 0$) is referred to as the *hydraulic conductivity at natural saturation* (K_s).

6.2.3.5 Hydraulic Conductivity Characteristic

When expressed as a function of the volumetric soil water content, the hydraulic conductivity function $K(\theta)$ is strongly nonlinear. Generally speaking, it behaves like a power function. For decreasing soil water content, the hydraulic conductivity decreases rapidly. Figure 6.4 illustrates the typical behavior of hydraulic conductivity. The hydraulic conductivity function highly depends on soil structure. For sandy soils, the hydraulic conductivity at natural saturation (K_s) is usually higher than that for clay soils, even though the porosity is higher in clay soils.

Besides the water retention and hydraulic conductivity functions, alternative combinations of soil characteristics can be chosen to characterize the soil moisture behavior of soils, such as $K(h)$ (e.g., Gardner, 1958; Rijtema, 1965), $D(\theta)$ (e.g., Childs and Collis-George, 1950), and/or $C(\theta)$ (e.g., Richards, 1931), where D refers to *diffusivity* defined by $D = K(\theta)dh/d\theta$ and C is *specific capacity* defined by $C(\theta) = d\theta/dh$. Both functions, $C(\theta)$ and $D(\theta)$, will be addressed in details in Section 6.3.1. The quintessence is that always two soil characteristics are required to model soil water movement in the vadose zone.

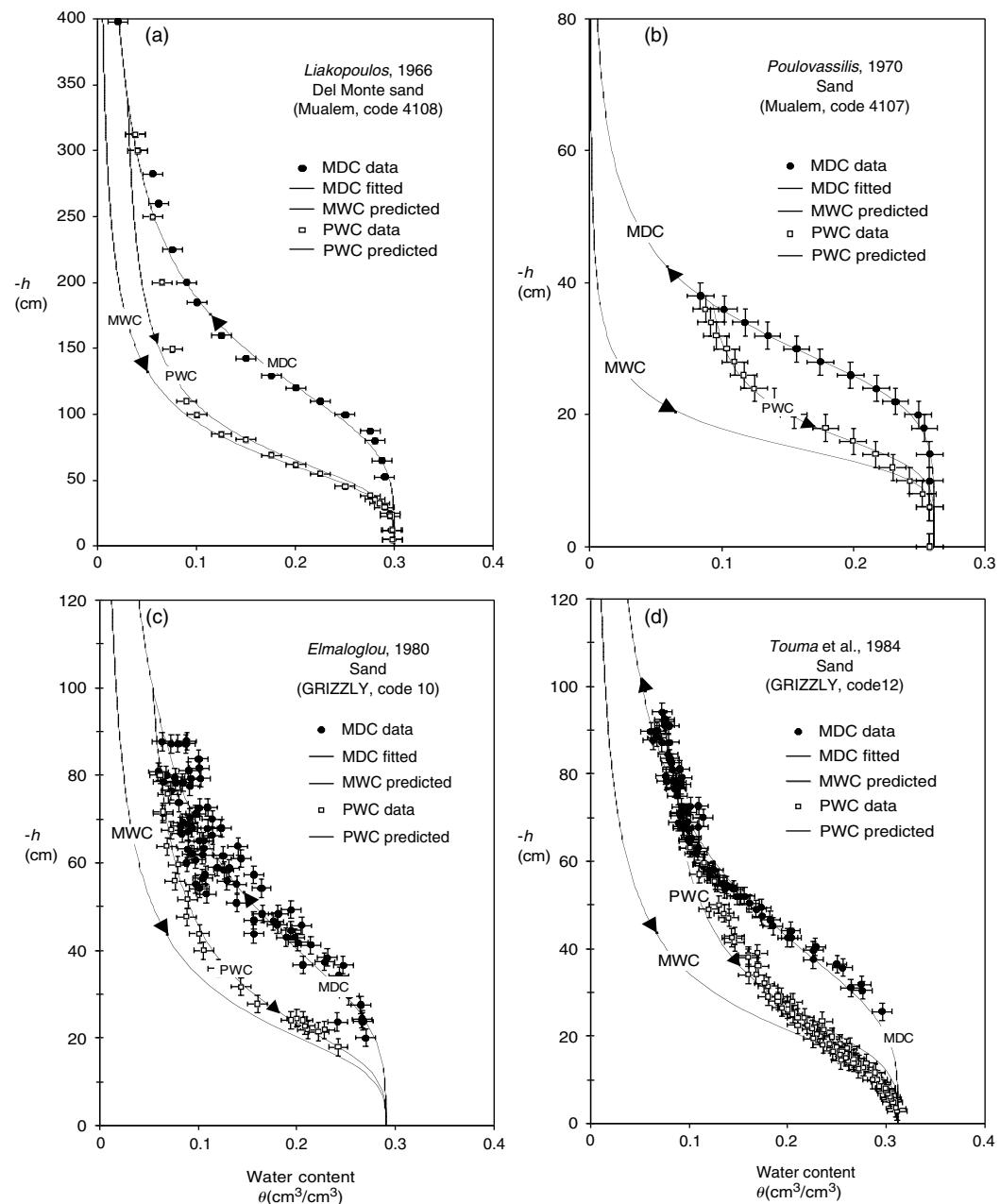


FIGURE 6.5 Main drying curve (MDC) fitted to measured drying data (\bullet) for four sands. The predicted main wetting curve (MWC) and primary wetting curve (PWC) are presented together with experimental primary wetting data (\square). The horizontal and vertical bars correspond to the measurement errors in water content and pressure head (i.e., $\pm 0.01 \text{ cm}^3/\text{cm}^3$ and $\pm 2 \text{ cm}$) associated with each observation. (Adapted from Haverkamp, R., Reggiani, P., Ross, P. J., and Parlange, J.-Y. 2002. Soil water hysteresis prediction model based on theory and geometric scaling. In *Environmental Mechanics: Water, Mass and Energy Transfer in the Biosphere*, Raats, P. A. C., Smiles, D., and Warrick, A. W. [Eds.], Geophysical Monograph 129, American Geophysical Union, Washington, DC, pp. 213–246.)

6.2.4 Functional Relationships

In the past, many different functional relations have been proposed in the literature. However, most refer to the hydraulic properties of soils and only little concern the physical properties of soils. Next, we present the most popular relations used for both groups of physical and hydraulic soil properties.

6.2.4.1 Particle-Size Distribution

The particle-size distribution describes the relation between the distribution frequency (F_p) expressed as a percentage of particle weight per unit sample mass and the effective diameter of soil particles (D_p). According to Haverkamp and Reggiani (2002), the cumulative distribution of particle weight (F_p) is described by an expression similar to that chosen for the water retention curve (see Equation 6.20 and Equation 6.21). Using the van Genuchten (1980) type-equation, $F_p(D_p)$ takes the form:

$$F_p = \left[1 + \left(\frac{D_{pg}}{D_p} \right)^N \right]^{-M} \quad (6.18)$$

where F_p varies from 0 for $D_p = 0 \mu\text{m}$ to 1 for $D_p = 2000 \mu\text{m}$; D_{pg} is a *particle-size scale parameter*; and M and N are the *particle-size shape parameters* assumed to be linked by the expression:

$$M = 1 - \frac{k_M}{N} \quad \text{with } N > k_M \quad (6.19)$$

The integer k_M is chosen identical to that used by the van Genuchten (1980) water retention equation (Equation 6.21). An example of $F_p(D_p)$ determined for a silt loam taken from the GRIZZLY soil database (Haverkamp et al., 1997), is shown in Figure 6.6.

The cumulative particle-size distribution function is of use for the determination of soil hydraulic properties through the use of Pedo-Transfer Functions or PTFs (Bouma and van Lanen, 1987). These

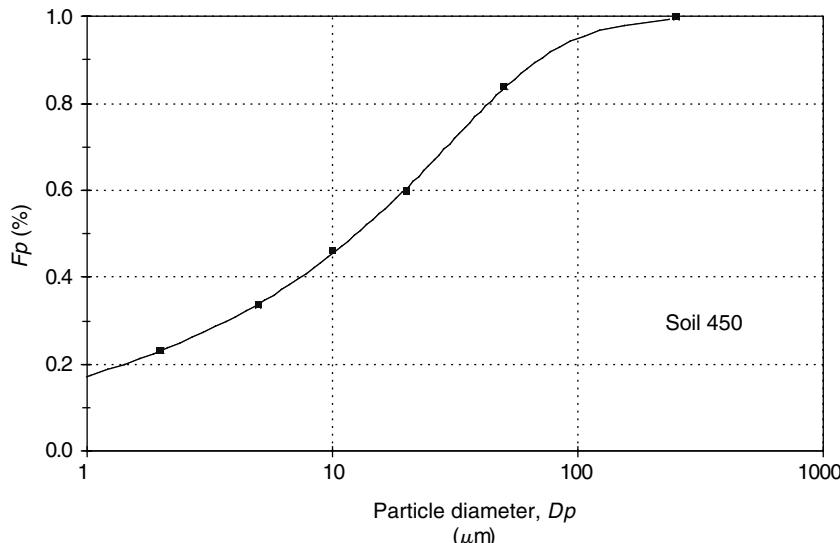


FIGURE 6.6 Cumulative particle-size distribution function $F_p(D_p)$ fitted to measured soil texture data observed for a silt loam. (Adapted from Haverkamp, R., Zammit, C., Bouraoui, F., Rajkai, K., Arrué, J. L., and Heckmann, N. 1997. GRIZZLY, Grenoble Soil Catalogue: Soil survey of field data and description of particle-size, soil water retention and hydraulic conductivity functions. Laboratoire d'Etude des Transferts en Hydrologie et Environnement [LTHE], Grenoble Cedex 9, France. With permission.)

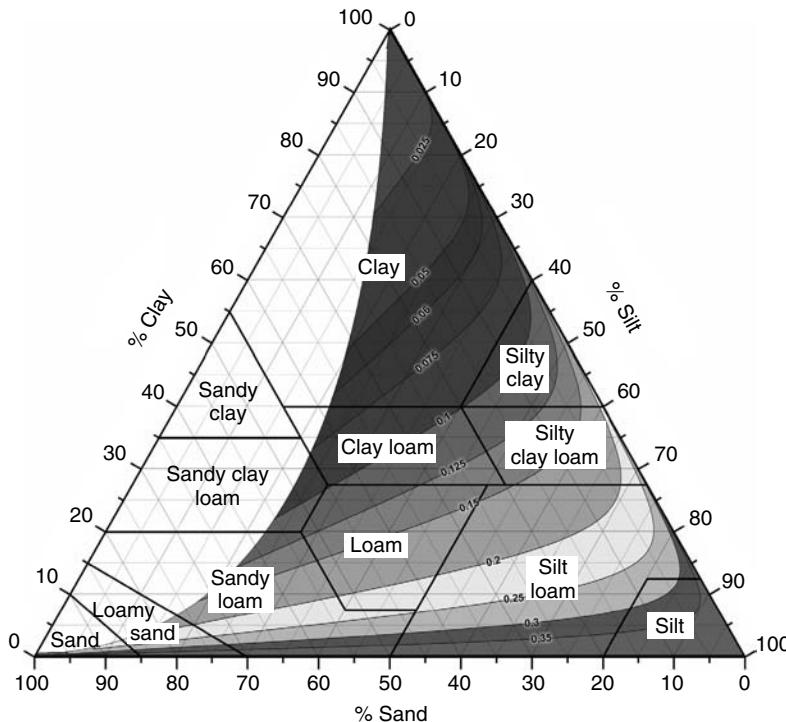


FIGURE 6.7 (See color insert following page 23-52) Contour lines of the particle-size shape parameter M plotted on the basis of the USDA soil texture triangle as a function of the different soil texture classes.

are generally empirical relationships that allow the hydraulic properties of a given soil to be predicted from more widely available data, such as texture (% of sand, silt, and clay) and bulk density (the PTFs are described in detail in Section 6.2.6). Among the PTFs proposed in the literature, some have been developed to predict the parameters of water retention equations of Brooks and Corey (Equation 6.20) or van Genuchten (Equation 6.21). For these PTFs, the direct conversion from particle-size shape parameters (M , N , and D_{pg}) to water retention parameters is a suitable choice. Using the USDA soil texture triangle (Figure 6.3) as classification chart, Figure 6.7 and Figure 6.8 show the contour lines of the particle-size shape (M) and scale (D_{pg}) parameters calculated for a mega number (10^6) of artificially generated soils.¹

As shown by Figure 6.6, Equation 6.18 describes a mono-modal function. This form is traditionally encountered for soils of temperate regions, where the silt fraction is present. As pointed out by Hodnett and Tomasella (2002), the weathering processes in temperate regions which have been going on for long periods of time, uninterrupted by ice ages, will have contributed to the large amounts of silty soils. When plotted on the USDA soil texture triangle, these soils correspond to the area covered by the colored contour plots of Figure 6.7 and Figure 6.8. They represent the bulk of soils of most soil databases such as UNSODA (Leij et al., 1996) and GRIZZLY (Haverkamp et al., 1997) used to derive the classical PTFs (e.g., Rawls and Brakensiek, 1985). However, when soils have low silt content, as is the case for tropical and subtropical soils, then the particle-size frequency curve shows a bimodal behavior (“camel” back behavior). These soils are situated in the blank area of Figure 6.7 and Figure 6.8. Obviously, the use of PTFs derived for the soils of temperate regions are inadequate for tropical soils and specific PTFs should be derived exclusively using data for tropical soils (e.g., Tomasella and Hodnett, 1998; Tomasella et al., 2000; Hodnett and Tomasella, 2002). Even though the complication of bimodal texture behavior has drawn very little attention in the literature, the consequences can be far-reaching. As there is shape similarity between the cumulative

¹The plots are freely provided by HYDROWIDE, expertise in hydroinformatics, www.hydrowide.com

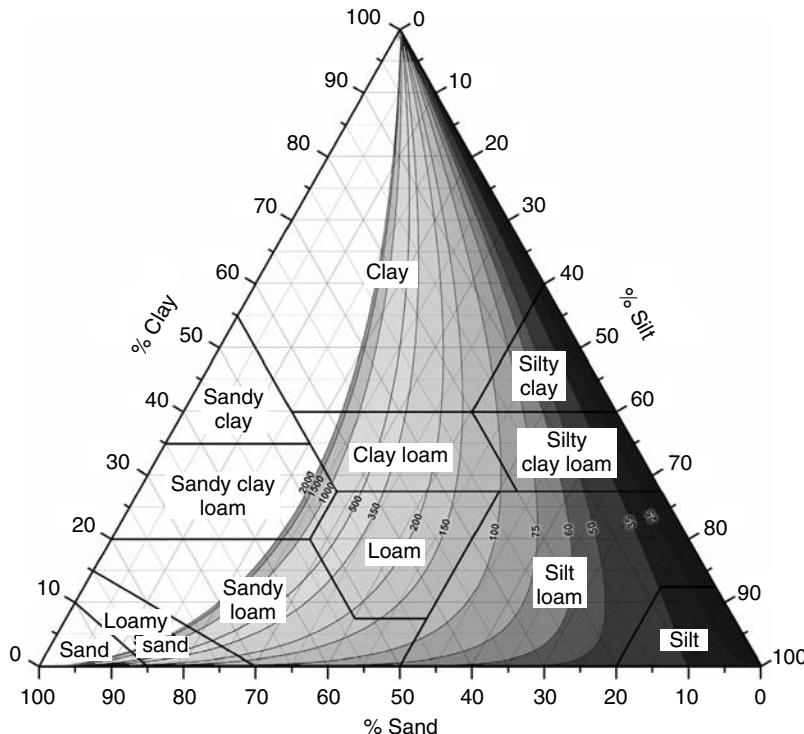


FIGURE 6.8 (See color insert following page 23-52) Contour lines of the particle-size scale parameter D_{p_8} (μm) plotted on the basis of the USDA soil textural triangle as a function of the different soil texture classes.

particle-size distribution and water retention functions (Haverkamp and Parlange, 1986), it follows that most tropical and subtropical soils show bimodal water retention characteristics (de Condappa et al., 2005).

The second category of functional relationships to be dealt with concerns the group of relations used to describe the hydraulic soil properties. Many different functions have been proposed in the literature based on various combinations of the dependent variables θ , h , and K , and a certain number of fitting parameters (e.g., Gardner, 1958; Brooks and Corey, 1964; Brutsaert, 1966; van Genuchten, 1980; Haverkamp and Vauclin, 1981). Independent of the correct physical meaning of the fitting parameters, their values are submitted to constraints imposed by the use of the transfer equations such as the Fokker-Planck or Richards' equation (see Section 6.4). The most frequently used in the literature are the following water retention and hydraulic conductivity expressions.

6.2.4.2 Water Retention Curves

The Brooks and Corey (1964) equation:

$$\begin{cases} \frac{\theta - \theta_r}{\theta_s - \theta_r} = \left[\frac{h_{bc}}{h} \right]^\lambda & \text{for } h \leq h_{bc} \\ \theta = \theta_s & \text{for } h_{bc} \leq h \leq 0 \end{cases} \quad (6.20)$$

and the van Genuchten (1980) water retention equation:

$$\frac{\theta - \theta_r}{\theta_s - \theta_r} = \left[1 + \left(\frac{h}{h_g} \right)^n \right]^{-m} \quad (6.21)$$

where the water pressure head (h) is usually taken to be negative and is expressed in cm of water; h_{bc} is the Brooks and Corey *pressure scale parameter*; h_g is the van Genuchten *pressure scale parameter*; and λ , m , and n are *water retention shape parameters*. The water retention shape parameters m and n are frequently related according to:

$$m = 1 - \frac{k_m}{n} \quad \text{with } n > k_m \quad (6.22)$$

where k_m is an integer value referred to hereafter as the *user index* (Haverkamp et al., 2005) that was initially introduced by van Genuchten (1980) to calculate closed-form analytical expressions for the hydraulic conductivity function when substituted in the predictive conductivity models of Burdine (1953) or Mualem (1976) (see Equation 6.43). For the Mualem theory parameter k_m takes the value $k_m = 1$, and for the Burdine theory $k_m = 2$. The constraint $\theta_r = 0$ is frequently imposed.

The Brooks and Corey equation represents a limiting case of the van Genuchten equation. That is to say, Equation 6.21 reduces to Equation 6.20 for pressure heads very large compared with the scaling pressure (h_g). This explains that both shape parameters λ and mn are often assumed equal when used for converting purposes (e.g., Lenhard et al., 1989; Russo et al., 1991; Rawls et al., 1992; Morel-Seytoux et al., 1996). However, Haverkamp et al. (2005) showed in a recent study that the condition $\lambda = mn$ becomes totally incorrect for large values of λ and mn . They also showed a strong dependence of the shape parameter value on the residual water content θ_r and the user index k_m . This interdependency of parameter values is a potential problem with parameterization of retention data. To facilitate the interpretation, conversion, and optimization of retention parameters, the authors introduced a *water retention shape index* P_{wr} , which constitutes an integral measure of the slope of the water retention curve and which characterizes the retention behavior of a particular soil with a single number. The water retention shape index is constant for a given soil and provides a benchmark measure of texture. Its value which can be associated with the water retention fractal dimension, varies between 0 and 3 which is the maximum expected value for fractal behavior. In an associated study, Leij et al. (2005) applied the shape index concept to derive the conversion equation:

$$P_{wr} = \lambda_0 = \frac{mn_0}{1 + m_0} = \frac{mn_0(k_m + mn_0)}{k_m + 2mn_0} \quad (6.23)$$

where the subscript 0 indicates the constraint $\theta_r = 0$. When a nonzero θ_r -value is applied, the shape parameter values are directly affected. For the example of the Brooks and Corey water retention equation, the shape parameter λ changes according to:

$$P_{wr} = \lambda_0 = \lambda \left[1 + \frac{\theta_r}{\theta_s - \theta_r} \ln \left(\frac{\theta_r}{\theta_s} \right) \right] \quad (6.24)$$

As P_{wr} was shown to be independent of the parameterization, it follows $\lambda \geq \lambda_0$ (Equation 6.23).

The difference in λ and λ_0 compensates for the effect of θ_r . Leij et al. (2005) derived similar equations for the van Genuchten water retention shape factor mn and mn_0 . As an example, Figure 6.9a and b show the influence of θ_r on the shape factor mn (Figure 6.9a) and the water retention shape index P_{wr} (Figure 6.9b) calculated for 253 soil sample taken from the GRIZZLY soil database (Haverkamp et al., 1997) subject to the Burdine condition ($k_m = 2$). Obviously, the interdependency of the different parameter values leads to nonunique parameter sets. When used in soil databases regardless of the different regression constraints, these parameter sets may well lead to rather dubious PTF functions.

The problem of bi-modularity mentioned above for the case of the cumulative particle-size distribution function, also counts for the water retention curve. Apart from Ross and Smettem (1993) and Durner (1994) who use sums of monomodal water retention functions to describe bimodal water retention behavior, only few authors have dealt with the conceptualization of a bimodal water retention model.

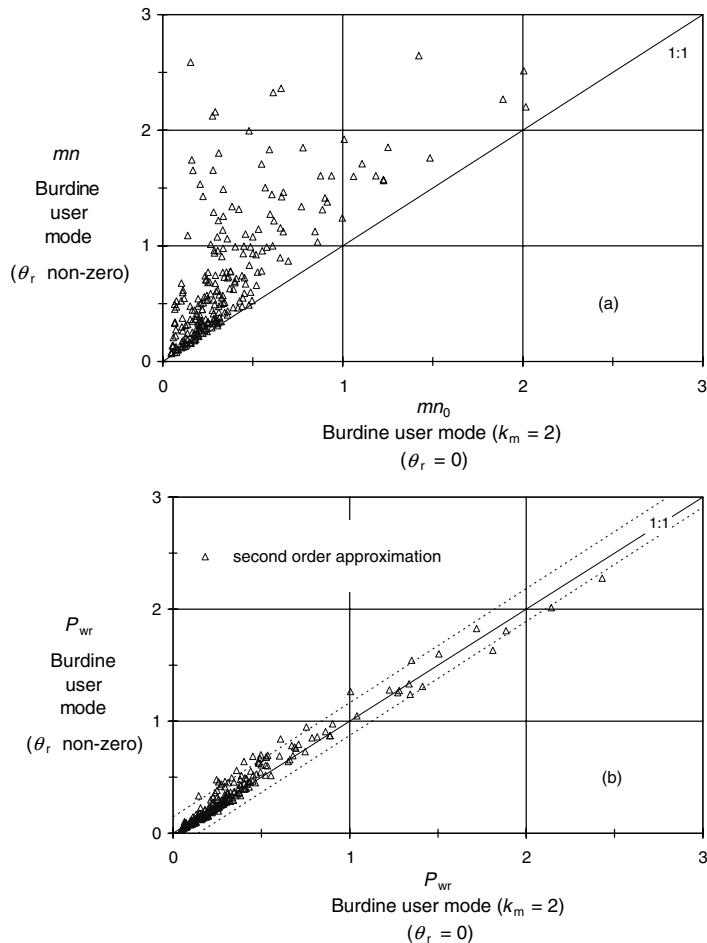


FIGURE 6.9 (a) The shape factor mn with θ_r optimized as a function of mn_0 with $\theta_r = 0$ for the van Genuchten equation for 253 soil samples from GRIZZLY subject to the Burdine condition. (Adapted from Haverkamp, R., Leij, F. J., Fuentes, C., Sciortino, A., and Ross, P. J. 2005. *Soil Sci. Soc. Am. J.*, 69: 1881–1890.) (b) The water retention shape index P_{WR} with θ_r optimized as a function of P_{WR} with $\theta_r = 0$ for the van Genuchten equation for 253 soil samples from GRIZZLY subject to the Burdine condition. (Adapted from Leij, F. J., Haverkamp, R., Fuentes, C., Zatarain, F., and Ross, P. J. 2005. *Soil Sci. Soc. Am. J.*, 69: 1891–1901.)

6.2.4.3 Hydraulic Conductivity Functions

The Brooks and Corey (1964) equation:

$$\frac{K}{K_s} = \left[\frac{\theta - \theta_r}{\theta_s - \theta_r} \right]^\eta \quad (6.25)$$

and the van Genuchten (1980) hydraulic conductivity equation:

$$\frac{K}{K_s} = \left[\frac{\theta - \theta_r}{\theta_s - \theta_r} \right]^{1/2} \left[1 - \left\{ 1 - \left(\frac{\theta - \theta_r}{\theta_s - \theta_r} \right)^{1/m} \right\}^m \right]^2 \quad (6.26)$$

where η is a *conductivity shape parameter*.

Through an extensive study, Fuentes et al. (1992) concluded that only the combination of the van Genuchten water retention equation (Equation 6.21), $h(\theta)$, based on the Burdine theory ($m = 1 - 2/n$)

together with the Brooks and Corey conductivity equation (Equation 6.25) stays valid for all different types of soil encountered in practice without becoming inconsistent with the general water transfer theory. This is due to the rather limiting constraint that exists for shape parameter m when using the Mualem theory: $0.15 \leq m \leq 1$. As mentioned before, the value of θ_r is generally taken equal to zero.

For modeling purposes, the soil characteristic equations are often expressed in dimensionless form (e.g., Haverkamp and Vauclin, 1979; Warrick et al., 1985). The following dimensionless soil variables are used:

$$\theta^* = \frac{\theta}{\theta_s}; \quad h^* = \frac{h}{h_g}; \quad \text{and } K^* = \frac{K}{K_s} \quad (6.27)$$

where the superscript* refers to the nondimensional form of the different variables.

The use of Equations 6.20–6.26 combined with appropriate initial and boundary conditions allows the description of water transfer in the vadose zone of soil in a fully deterministic way (e.g., in the form of Richards [1931] equation, see Section 6.4). However, such a comprehensive approach has its price in complexity, as it requires the determination of five unknown soil parameters:

- The two dimensionless shape parameters (m and η); and
- The three scale parameters θ_s , h_g , and K_s .

Both shape parameters are strongly linked to the textural soil properties, whereas the scale parameters are related to soil structure (Haverkamp and Reggiani, 2002).

Before going to the following section on the measurements of soil characteristics it is important to underline two major difficulties in applying the foregoing formulism to field studies:

- Equations 6.20–6.26 only refer to the capillary water retention and hydraulic conductivity soil properties. Side effects due to *macroporosity*, such as soil cracks, root holes, worm channels, and large pores formed for various biological or mechanical reasons are not incorporated (see rather Chapter 9 and Chapter 22); and
- The soil characteristic equations only give point scale information (i.e., *local scale*) with a critical scale area of 1 m^2 . Consequently, when dealing with the modeling of soil moisture dynamics at field scales, there exists a clear mismatch between the scale at which the functional soil characteristics are defined (and measured!) and the characteristic spatial scale for which models are trying to make predictions.

6.2.5 Measurement of Soil Characteristics

Many experimental works have been devoted over the last decades to the development of measurement techniques for estimating soil hydraulic characteristics. Roughly speaking, two categories of methods can be distinguished for the determination of the unknown soil hydraulic parameters: (1) the measurement techniques (direct or indirect) and (2) the predictive methods. The first category of techniques has been developed mainly for the measurement of the soil hydraulic parameters at a local scale and is difficult to apply over large areas. The second category is more flexible and could possibly be used at larger scales. For that reason we have preferred to address the predictive methods separately in Section 6.2.6.

In general, the measurement techniques rely on precise and time-consuming experimental procedures that can be categorized as being either laboratory- or field-based. While laboratory methods allow accurate measurement of flow processes, they are performed on samples taken from the field, and as a result, their representativity of field conditions can be questioned. The presence of aggregates, stones, fissures, fractures, tension cracks, and root holes, commonly encountered in unsaturated soil profiles, is difficult to represent in small-scale laboratory samples. Field techniques can be more difficult to control, but they have the advantage of estimating *in situ* soil hydraulic properties that are more representative, which is of considerable value in the subsequent use of the hydraulic information. Therefore, it is desirable to aim at field methods that can ease, to some extent, the time-consuming constraints.

6.2.5.1 Soil Water Content

Soil water content is determined using direct or indirect methods. Direct techniques, such as the gravimetric method, involve the measurement of water losses through the process of evaporation, or drainage. Chemical reactions can also be used to displace water. Indirect methods, such as electrical or radiological techniques, estimate the physical properties of some factor known to have an influence on soil water content. Then through a calibration procedure, these properties are related to soil water content. A complete and detailed description of the different methods is given by Gardner (1986).

6.2.5.1.1 Gravimetric Method

This method usually involves the weighing of a wet soil sample, the removal of water by evaporation through oven drying or microwave heating, and then the reweighing of the dry sample. The soil water content is determined as the ratio of the net weights of the water removed to the dry sample. A high accuracy of soil water content can be achieved by drying the soil sample until a constant weight is reached. Even though it is a destructive and time-consuming approach, it is often used routinely because it does not require any calibration and it is easy to perform.

6.2.5.1.2 Electrical Resistance Methods

Electrical resistance methods are based on the strong correlation between the electrical resistance of a porous media and its water content. Two electrodes inserted in a porous block made of a variety of material, resistant to degradation once buried (e.g., gypsum, nylon, fiber glass), are placed into the soil. Once equilibrium is reached (identical matric potential for the porous block and the soil), the water content of the porous block is an indicator of the soil water content. This nondestructive technique is affected by hysteresis; due to technical limitations, only a drying calibration curve can be determined. The use of such techniques is highly limited because of a reduced accuracy, especially for high water contents.

6.2.5.1.3 Neutron Thermalization

The basis of this method is the property of hydrogen atoms to slow down and scatter fast, high energy colliding neutrons. The successive collisions result in a loss of energy until thermal equilibrium with the surrounding atoms is reached (*thermalization*). A source of fast neutrons is lowered in an access tube. Through collisions with the hydrogen atoms, these neutrons become thermalized. A counting system determines the concentration of thermal neutrons by calculating the number of particles emitted by the reaction of the low-energy neutrons and boron trifluoride present in a detector. Through calibration, the counts of the thermal neutrons are transformed to a water content. This method is nondestructive, but it requires calibration and its use is actually limited because of environmental regulations. Another limiting factor is the volume of measurement and its dependency on water content. Detailed information on the error analysis in estimating soil water content from neutron probe measurements are given by Haverkamp et al. (1984) and Vauclin et al. (1984).

6.2.5.1.4 Gamma Rays Attenuation Method

This nondestructive method consists of measuring the attenuation of intensity of a beam of gamma rays when crossing a soil column. The attenuation depends on the soil constituents and bulk density. Measuring the attenuation at two different gamma ray energies yields both the water content (θ) and dry bulk density (ρ_d). The apparatus includes a source of gamma rays and a scintillator detector. An accurate calibration is required.

6.2.5.1.5 Capacitance Method: Time Domain Reflectometry (TDR) Methods

These methods are based on the measurement of the dielectric constant of the soil. As the water has a very large dielectric constant compared to that of the solid phase and that of air, it has a large influence on the dielectric constant of the soil. TDR consists of measuring the transfer time of electromagnetic waves along the parallel metallic rods (two or more) of known length inserted in the soil:

$$C_e = \left[\frac{ct_{tr}}{2l} \right]^2 \quad (6.28)$$

where C_e is the relative dielectric constant of the soil; t_{tr} is the wave travel time from the entrance into the soil to the end of the rods; l is the length of the rods; and c is the light velocity. The next step is to calibrate the C_e vs. a known value of volumetric water content. The technique is described in a comprehensive paper by Topp et al. (1980). Although the authors suggested that the procedure is insensitive to variations in dry bulk density, temperature, mineral composition, and salinity (i.e., one single calibration curve could be applied to nearly all soils), it is evident that the sensitivity of the measurements depends on the dielectric constant of the material between or around the wire probes (e.g., Knight, 1992). Thus, local nonuniformities due to small air gaps or material of different density can cause considerable errors in the measured water content values. It complicates significantly the interpretation of the field water content measurements and often requires site-specific calibration curves. Haverkamp et al. (1996) proposed a procedure to correct the apparent (i.e., TDR) soil water content values for the effects of air gaps, small stones, and water-holding coarse fragments. The advantage of the TDR method is that it is nondestructive. As the use of the neutron probe is limited for environmental reasons, the TDR method is becoming more and more a standard.

6.2.5.1.6 Remote Sensing

Remote sensing involves airborne measurements, usually from satellite or airplanes, of the electromagnetic signature of a body or surface. For inferring soil moisture, measurements are usually done in the visible, infrared, and microwave spectra. Microwave sensors are of increasing interest as they measure the dielectric properties of a surface body without being affected by cloud cover. Two microwave sensors are used: passive sensors, also known as *radiometers*, which measure microwave emission, and active sensors or *radars*, which send a microwave signal and measure its reflection. One major advantage of remote sensing is that it integrates soil moisture over a certain area. However the measurements are limited to the very first centimeters of the soil and are affected by surface roughness and vegetation density. The calibration procedure is extremely difficult because of the temporal and spatial variability of ground measurement of soil moisture.

6.2.5.2 Soil Water Pressure Head

Depending on the pressure range measured, two different devices are most widely used: tensiometers and psychrometers.

6.2.5.2.1 Tensiometers

A tensiometer is composed of a porous cup, a connecting tube or barrel, and a pressure measuring device that can be a water manometer, mercury manometer, pressure transducer, or a vacuum gauge. The porous cup is filled with degased water and has to remain saturated for the whole soil water pressure range. The porous cup and the surrounding soil are in hydrostatic equilibrium, yielding then a measure of the soil water pressure. The use of tensiometers is limited to pressure heads ranging from 0 to 800 cm. For higher pressures, the use of psychrometers is recommended. A detailed description of the different types of tensiometers is given by Cassel and Klute (1986).

6.2.5.2.2 Psychrometer

In this method, the soil water potential is related to the relative humidity defined by the ratio of the water vapor pressure in equilibrium with the liquid phase (p) to the saturated vapor pressure (p_0) as:

$$h = \frac{R_g T_l}{M_{\text{mol}}} \ln \left(\frac{p}{p_0} \right) \quad (6.29)$$

where M_{mol} is the molecular weight of water, R_g is the universal gas constant, and T_l is the temperature (Kelvin) of the liquid phase. Details about the thermocouple psychrometer design and functioning are given by Rawlins and Campbell (1986).

6.2.5.2.3 Other Methods

Additional methods for measuring soil water potential include *electric resistance sensors*, which can be gypsum, nylon, or fiberglass. Once potential equilibrium is reached between the resistance sensor and the soil matrix, the electric conductivity is measured and then transformed to water potential through pre-established calibration curves. Additional methods such as *heat dissipation*, are described by Campbell and Gee (1986).

6.2.5.3 Hydraulic Conductivity

Hydraulic conductivity is a soil characteristic that cannot be measured directly. Unlike the water retention characteristic, which can be considered as a quasistatic soil property, the hydraulic conductivity function is always related to water movement into or through the soil. The transient and steady movement of water in the unsaturated zone of soils depends to a large extent upon the pore network resulting from the assemblage of soil particles and aggregates. This dependence of $K(\theta)$ on the pore-size distribution underlines the complicity that exists between the water retention and hydraulic conductivity functions. It implies at the same time that the optimization of soil water characteristic parameters over measured field data should be carried out simultaneously with a combined objective function. As only very little information is available in the literature on this problem (e.g., Yates et al., 1992), it seems most likely that combined optimization of water retention and hydraulic conductivity parameters is unfortunately seldom used in practice.

In the following, three methods developed for *in situ* determination of hydraulic conductivity properties are presented. The *instantaneous profile method* must be considered as a fixed ground based experiment, while the two other methods, *tension disk infiltrometer* and *pressure ring infiltrometer*, can be looked upon as mobile experiments. The two latter methods are based on inverse techniques. *Inverse methods* consist of procedures that use solutions for a flow process inversely against observations of that process. The hydraulic parameters are then estimated either via reorganization of the solution into an explicit expression, or through numerical optimization (e.g., Simunek et al., 1998). Numerical solutions usually have high computation overheads. A more attractive approach is to use analytical solutions that can yield better appreciation of the role of the different soil parameters interfering in the flow processes, and also require less computation. For a possible successful application of inverse methods, it is crucial to describe the flow process as precisely as possible and to control perfectly well the imposed initial and boundary conditions.

6.2.5.3.1 Instantaneous Profile Method

In field studies, the instantaneous profile method is a transient method used to observe the natural water movement into the soil. During the time interval $\Delta t = t_2 - t_1$, the change per unit surface in the total water storage of the vadose zone (ΔS), from the soil surface to the reference depth z_m (Figure 6.10), is calculated as:

$$\Delta S(z) = \int_0^{z_m} [\theta(z, t_2) - \theta(z, t_1)] dz \quad (6.30)$$

Without any water supply to the soil profile during the period $[t_1, t_2]$, the storage variation corresponds to the water volume drained across the z_m plane or evaporated (i.e., evapotranspiration) through the plane $z = 0$, or even to a combination of both processes simultaneously. The flow direction can be determined from the gradient of the measured profiles of hydraulic head $H(z, t)$. A negative gradient, $dH/dz < 0$, indicates a positive, or downward-oriented water flux in Darcy's law (see also Section 6.3, Equation 6.49). This flux is associated with percolation or drainage. A positive gradient, $dH/dz > 0$, corresponds to a negative or upward-oriented water flux (Figure 6.10), associated with evaporation. A *zero flux plane* z_0 is defined when the hydraulic gradient $dH/dz = 0$. During a redistribution, the zero flux plane moves downward.

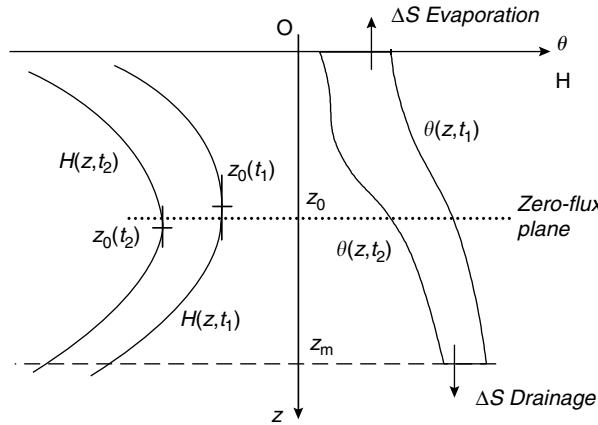


FIGURE 6.10 Instantaneous profile method to estimate unsaturated hydraulic conductivity.

To determine the mean water flux per surface unit, at any soil depth, $q(z)$, it is possible to integrate the mass conservation equation from the zero flux plane z_0 to the chosen depth z as:

$$q(z) = \frac{\Delta S(z_0, z)}{t_2 - t_1} \quad (6.31)$$

where z_0 is the position of the zero flux plane averaged over the period $[t_1, t_2]$. The application of the Darcy's law (Equation 6.49) gives the hydraulic conductivity directly:

$$K(\theta) = -\frac{q(z)}{dH/dz} \quad (6.32)$$

The hydraulic gradient dH/dz in Equation 6.32 is the averaged value for the period Δt calculated at the depth z .

6.2.5.3.2 Tension Disk Infiltrometer

A representation of three tension disk infiltrometers with 250, 80, and 48 mm diameter bases (after the design of Perroux and White, 1988) are shown in Figure 6.11. A graduated reservoir tower provides the water supply; the bubble tower with a moveable air-entry tube (C_1) permits to impose different boundary condition pressure head values at the cloth base (with a mesh of $20 \mu\text{m}$). A thin layer of fine sand is placed over the soil surface to ensure a good hydraulic contact between the disk and the soil. *Cumulative infiltration*, $I(t)$, is recorded by measuring the water level drop in the reservoir tower (for the proper definition of cumulative infiltration see Section 6.3). The transient infiltration flux is then given by the derivative: $q(t) = dI/dt$. By the use of an inverse procedure, these flux data allow the calculation of the spot values of hydraulic conductivity and *sorptivity* valid for the initial and boundary conditions chosen (for the proper definition of sorptivity, see Section 6.4). Under field conditions, the contact layer of sand may be important. When this is the case, the water volume stored during the early stages of infiltration can be significant. Noting I_{sa} as the water volume necessary to fill the sand layer over the time period t_{sa} , the infiltration measurements have to be corrected by the coordinate transformation: $(I - I_{sa})$ and $(t - t_{sa})$.

In most cases, slightly negative supply pressures are applied (Perroux and White, 1988), which allow the determination of hydraulic conductivity values close to saturation representative for the fine soil matrix without being biased by the possible influence of macroporosity. The fact that the initial and boundary conditions are well-controlled makes the disk infiltrometer experiments particularly appropriate for data analysis through inverse procedures. The standard analysis uses Wooding's (1968) solution for

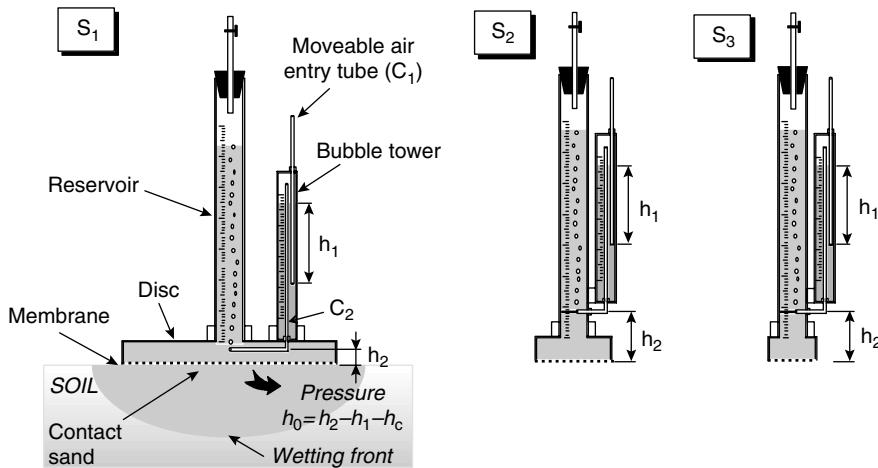


FIGURE 6.11 Tension disk infiltrometer. (Adapted from the design of Perroux, K. M. and White, I. 1988. *Soil Sci. Soc. Am. J.*, 52: 1205–1215. With permission.)

three-dimensional (3D) steady state infiltration valid for infinite time and uniform initial conditions. Unfortunately, these conditions are seldom met in the field. To overcome these limitations, a 3D analytical solution of infiltration has been derived (Haverkamp et al., 1994; Smettem et al., 1994), which allows the description of transient 3D infiltration behavior.

6.2.5.3.3 Pressure Ring Infiltrometer

The pressure ring infiltrometer is formed by a metallic ring that is driven into the soil to a given depth and connected to a reservoir system where the cumulative infiltrated volume could be measured for constant or falling hydraulic head. The reservoir system has to be adapted to the permeability characteristics of each soil. For soils with high permeability, water is supplied to the soil surface at a constant head (H_1) through the sealed top lid from a Mariotte bottle, with a moveable air tube allowing a wide range of different head values H_1 to be applied. For soils with low permeability, water must be supplied from a small capillary tube also acting as a measuring burette. This tube can be positioned either horizontally for water to infiltrate at a constant head, or vertically for water to infiltrate at a continually falling head. The cumulative infiltration, $I(t)$, is obtained from readings of the water supply tube. The saturated hydraulic conductivity is then deduced from the cumulative infiltration measurements by the use of analytical infiltration solutions developed for 1D positive head infiltration subject to uniform initial conditions (Elrick et al., 1995). As these conditions are rarely met under field conditions, the results should be interpreted with caution. The initial condition is determined from undisturbed soil water content samples taken near the single ring.

6.2.5.3.4 Other Methods

Over the last decade the *multistep outflow* method has gained interest in the literature (e.g., Kool et al., 1987; Kool and Parker, 1988; Echings and Hopmans, 1993; van Dam et al., 1994). It concerns a laboratory-based method that allows the estimation of the soil characteristics by inverse modeling of a series of outflow experiments. For a successful application of the inverse optimization procedure it is crucial to ensure the uniqueness of the solution; consequently, various laboratory outflow experiments subject to different initial and boundary conditions are generally required. The multistep outflow method becomes more and more time-consuming as the number of unknown parameters to be determined increases. As for most laboratory methods, the representativity of the results for field conditions is questionable.

6.2.5.4 Combined Water Retention and Hydraulic Conductivity

Recently, a portable field method has been reported in the literature that allows for the simultaneous characterization of both hydraulic soil characteristics, $h(\theta)$ and $K(\theta)$. The method termed as the *Beerkan*

method, was initially pioneered by Haverkamp et al. (1996). It uses the van Genuchten water retention function (Equation 6.21) and the Brooks and Corey hydraulic conductivity function (Equation 6.25) to describe the hydraulic characteristic curves, and estimates their texture-dependent shape parameters (m , n , and η) from simple particle analysis, and their structure-dependent scale parameters (θ_s , h_g , and K_s) from field infiltration experiments at null pressure head. This method earlier described by Haverkamp et al. (1998a) was applied to many field studies (Galle et al., 2001; Braud et al., 2003) and has recently been improved and assessed with simulated data (Braud et al., 2005; Lassabatère et al., 2005).

6.2.5.4.1 Beerkan Method

As for the pressure ring infiltrometer, the Beerkan method uses a simple annular ring. The surface vegetation is removed while the roots remain *in situ*. A soil sample is collected for particle-size analysis and to determine its initial gravimetric water content. Another sample of known volume is extracted to determine its bulk density (ρ_d). Then, the cylinder is positioned at the soil surface and inserted to a depth of about 1 cm to avoid lateral loss of the ponded water at the surface. A fixed volume of water is poured into the cylinder at time zero, and the time elapsed during the infiltration of the known volume of water is measured. When the first volume has completely infiltrated, a second known volume of water is added to the cylinder and the time needed for it to infiltrate is measured (cumulative time). The procedure is repeated for several times until a quasi steady state infiltration flux is observed. At the end of the experiment, the saturated soil is sampled to determine the saturated gravimetric water content (w_s) and thus the saturated volumetric water content (θ_s) from the bulk density (ρ_d) considering the specific density of water (ρ_w) equals 1 g/cm³ (Equation 6.9).

Fitting infiltration experiment data on either numerical or analytical expressions provides estimations of scale parameters and combinations of them, such as sorptivity (described in more detail in Section 6.4 and Section 6.5). Considering an infiltration experiment with zero water pressure on a circular surface above a uniform soil with a uniform initial water content (θ_0), the 3D cumulative infiltration, $I(t)$, and the infiltration rate, $q(t)$, can be approached by the very accurate explicit transient two-term and steady-state expansions given by Haverkamp et al. (1994, 1998a, 1998b). Considering that the conductivity shape parameter (η) and the initial and saturated water content (θ_0 and θ_s) were determined beforehand, the fit of experimental data provides estimations of saturated hydraulic conductivity (K_s) and sorptivity. The latter parameter is then expressed as a function of the soil characteristic scale parameters (θ_s , h_g and K_s) using expressions given by Haverkamp et al. (1998b). It allows for the determination of the pressure scale parameter h_g .

The originality of the *Beerkan method* lies in its inverse procedure which uses a specific identification algorithm based on soil physical constraints. The algorithm noted as BEST (Beerkan Estimation of Soil Transfer parameters) is described in detail by Lassabatère et al. (2005). The identification technique used for the estimation of the water retention and conductivity shape parameters (m , n , and η) from the measured cumulative particle-size distribution (Equation 6.18) is embedded in the BEST algorithm.

6.2.6 Estimation Techniques of Soil Characteristics

As shown in the previous section, direct measurements, either *in situ* or in laboratory, may be very time-consuming and expensive, in particular for large hydrological studies. Furthermore, it is extremely useful to have efficient methods for estimating soil transfer characteristics in areas where the amount of available information is limited. This explains that many attempts have been made in the literature at estimating soil characteristics from readily available data, such as textural soil properties (i.e., particle-size distribution) and porosity, which are the most common measured soil data across the world. These relationships previously referred to (Section 6.2.4.) as PTFs, are particularly enticing as they can either be applied at the local scale using point textural properties or at the watershed scale where textural information has been aggregated (soil maps). However, the PTFs should be used with caution, as they often rely on statistical regression equations which make them site (data) specific.

6.2.6.1 Water Retention Relation

Basically, three different approaches have been developed to predict soil water characteristics from the particle-size distribution: (1) discrete matric potential regression methods, (2) functional regression methods, and (3) semi-physical approaches.

6.2.6.1.1 Discrete Matric Potential Regression Methods

In this approach, a multiple linear regression analysis is conducted to relate specific potential to particle-size distribution, porosity (ε), organic carbon content (CO) and bulk density (ρ_d). These methods make no assumptions concerning the form of the soil water retention curve. Gupta and Larson (1979) developed regression equations between specific matric potential, particle-size distribution, and organic matter content. Rawls and Brakensiek (1982) developed three regression models. The first model uses particle-size distribution, organic matter content, and bulk density as fitting variables. To improve their estimates, Rawls and Brakensiek (1982) introduced soil water content values at -1 , 500 kPa and -33 kPa, respectively, in the second and third model. Tietje and Tapkenhinrichs (1993) as well as Vereecken et al. (1992), who tested the point matric potential approach extensively, concluded that the discrete methods often give poor results as the regression equations are usually based on measurements conducted on disturbed soil samples, limiting their representativeness and applicability.

6.2.6.1.2 Functional Regression Methods

In this approach, a preliminary shape of the soil water retention curve is assumed (e.g., the Brooks and Corey function), and its parameters are derived through fitting (e.g., Clapp and Hornberger, 1978; Bloemen, 1980; Vereecken et al., 1989; Wösten and van Genuchten, 1988). The method is more adapted for unsaturated flow modeling since it gives a continuous functional description of the water retention curve. McCuen et al. (1981) found that the mean and standard deviation of the Brooks and Corey parameters (i.e., λ and h_{bc} of Equation 6.20) vary across soil textural classes. Cosby et al. (1984) extended this work and concluded that the textural soil properties can explain most of the variations of these parameters. This led to the development of regression equations between the mean and standard deviation of the soil hydraulic parameters and soil textural classes (Table 6.1). Rawls and Brakensiek (1985) developed a set of regression equations to estimate the water retention curve from particle-size distribution and porosity (Table 6.2).

All foregoing functional regression models have been developed using extensive databases for soils of temperate regions and, hence, apply to mono-modal water retention functions. As explained before (Section 6.2.4), the tropical and sub-tropical regions often show data outside the range of textural validity of the “temperate soil” PTFs tested. For example, the PTF of Rawls and Brakensiek (1985), illustrated in Table 6.2, is valid for soils with a clay content of 5–60% and a sand content of 5–70%, but some kaolinite tropical soils, particularly *Ferralsols*, can have a clay content of 70–90% (cf. Hodnett and Tomasella, 2002). This might suggest, from a temperate soils viewpoint, that they are heavy clays with a low permeability

TABLE 6.1 Regression Coefficients for the Brooks and Corey Parameters

Coefficients	Mean (μ) $\mu = c_1 + c_2Cl + c_3Sa + c_4Si$				Standard Deviation (σ) $\sigma = c_1 + c_2Cl + c_3Sa + c_4Si$			
	$\log(h_{bc})$	$1/\lambda$	θ_s	$\log(K_s)$	$\log(h_{bc})$	$1/\lambda$	θ_s	$\log(K_s)$
c_1	1.54	3.10	50.5	-0.6000	0.72	0.92	8.23	0.43
c_2	0	0.157	-0.037	-0.0064	0.0012	0.0492	-0.0805	0.0011
c_3	-0.0095	-0.003	-0.142	0.0126	0	0	0	0
c_4	0.0063	0	0	0	-0.0026	0.0144	-0.0070	0.0032

Note: h_{bc} is expressed in cm and K_s in inch/h; Cl, Si, and Sa represent the clay, silt, and sand fractions (%), respectively.

Source: From Cosby, B. J., Hornberger, G. M., Clapp, R. B., and Ginn, T. R. 1984. *Water Resour. Res.*, 20: 682–690. With permission.

TABLE 6.2 Regression Coefficients for the Brooks and Corey parameters

Regression Coefficients	Function Type ^a			
	$\exp(f)$	$\exp(f)$	(f)	$\exp(f)$
c_1	5.3396738	-0.7842831	-0.01824820	-8.968470
c_2	0.1845038	0	0.00513488	-0.028212
c_3	0	0.0177544	0.00087269	0
c_4	-2.48394546	-1.06249800	0.02939286	19.523480
c_5	-0.00213853	-0.00273493	-0.00015395	-0.0094125
c_6	-0.61745089	0	0	0
c_7	0	-0.00005304	0	0.00018107
c_8	-0.04356349	-0.03088295	-0.00108270	0.077718
c_9	0	1.11134946	0	-8.395215
c_{10}	0.50028060	-0.00674491	-0.00235940	0
c_{11}	0.00000540	0	0	-0.0000035
c_{12}	0.00895359	0.00798746	0.00030703	0.0273300
c_{13}	-0.00001282	-0.00000235	0	0.0000173
c_{14}	-0.00855375	-0.00610522	-0.00018233	-0.0194920
c_{15}	-0.00072472	0	0	0.0014340
c_{16}	0.00143598	0.00026587	0	-0.00298

^a Function $f = c_1 + c_2 \text{Cl} + c_3 \text{Sa} + c_4 \varepsilon + c_5 \text{Cl}^2 + c_6 \text{Cl}\varepsilon + c_7 \text{Sa}^2 + c_8 \text{Sae} + c_9 \varepsilon^2 + c_{10} \text{Cl}\varepsilon^2 + c_{11} \text{Cl}^2 \text{Sa} + c_{12} \text{Cl}^2 \varepsilon + c_{13} \text{ClSa}^2 + c_{14} \text{Cl}^2 \varepsilon^2 + c_{15} \text{Sa}^2 \varepsilon + c_{16} \text{Sa}^2 \varepsilon^2$

Note: h_{bc} is expressed in cm and K_s in cm/h; Cl is the clay fraction limited between (5% < Cl < 60%), Sa is the sand fraction limited between (5% < Sa < 70%), and ε is porosity.

Source: From Rawls, W. J. and Brakensiek, D. L. 1985. Prediction of soil water properties for hydrologic modeling. Watershed management in the eighties. *Proc. Irrig. Drain. Div., ASCE*, Denver, Colorado, and from Rawls, W. J. and Brakensiek, D. L. 1989. Estimating soil hydraulic properties. In *Unsaturated Flow in Hydrologic Modeling*, Morel-Seytoux, H. J. (Ed.), Kluwer Academic Publishers, published in cooperation with NATO Scientific Affairs Division, Boston. With permission.

and a moderate to high available water capacity. However, most of these soils have a low bulk density (ρ_d) in the order of 0.9–1.2 g/cm³ which is reflected in the much higher values of porosity (ε) and, hence, θ_s (Equation 6.3). If the pores are large, they will drain readily leading to small values of h_g . Indeed, Hodnett and Tomasella (2002) observed that most of the more clayey tropical soil show smaller mean values of h_g than for the temperate soils. Combined with the high porosity, this results in high values of K_s . Taking once again the example of the *Ferralsols*, it was found (Chauvel et al., 1991) that these soils show a range of large pores leading to small values of h_g , but also have a large proportion of their total porosity in very fine pores. These pores remain water filled at negative pressure values as high as $h = -4000$ kPa suggesting either a high residual water content value in the order of $\theta_r = 0.30 \text{ cm}^3/\text{cm}^3$ (Hodnett and Tomasella, 2002) or a bimodal behavior such as seems clearly imposed by the bimodal shape of the cumulative particle-size distribution function.

For the reasons mentioned above, we have preferred to include an example of a PTF derived exclusively using data for tropical soils. Basically, three tropical PTFs have been reported in the literature, that is, those of Tomasella and Hodnett (1998); Tomasella et al. (2000); and Hodnett and Tomasella (2002). The first study gives multiple regression equations to estimate the Brooks and Corey water retention parameters from texture data, based on a dataset of 614 samples for Amazonian soils. When reported on the USDA soil texture triangle of Figure 6.3 (Tomasella and Hodnett, 1998), they cover mainly the left area where bimodal behavior is expected. The regression coefficients as a function of the different soil textural classes are given in Table 6.3. While the second and third study use perhaps more comprehensive regression analyses, they also include supplementary explicative parameters such as the basic information on coarse and fine sand, organic carbon content (OC), dry bulk density (ρ_d), and pH values.

TABLE 6.3 Regression Coefficients for the Brooks and Corey Parameters According to Equation 6.20

Regression Coefficient	Function Type ^a			
	(f)	exp(f)	(f)	(f)
h_{bc}				
c_1	0.285	1.197	40.61	-2.094
c_2	0	4.17×10^{-3}	0.165	0.047
c_3	0	-4.50×10^{-3}	0.162	0.431
c_4	1.30×10^{-4}	8.94×10^{-4}	0	-8.27×10^{-3}
c_5	7.33×10^{-4}	0	1.37×10^{-3}	0
c_6	3.60×10^{-6}	-1.0×10^{-5}	1.80×10^{-5}	0

^a Function $f = c_1 + c_2 Si + c_3 Cl + c_4 SiCl + c_5 Si^2 + c_6 Si^2 Cl$

Note: h_{bc} is expressed in kPa; θ_s and θ_r are expressed in cm^3/cm^3 . Cl and Si are the percentages of clay and silt, respectively.

Source: From Tomasella, J. and Hodnett, M. G. 1998. *Soil Sci.*, 163: 190–202. With permission.

Notwithstanding that most of foregoing regression models clearly show correlation trends between water retention and textural or structural soil data, the application validity is often restricted to the soils tested for each study, hence, limiting the transportability of the methods. The models certainly present the advantage of easiness of use, but when applied for predictions, they give sometimes hazardous results with important errors on water content, especially in the wet range of $h(\theta)$, which is mostly due to the fact that the models fail to predict correctly the soil structure related water content and pressure scale parameters (θ_s and h_{bc} or h_g) from soil textural information alone.

6.2.6.1.3 Semophysical Methods

In spite of the problem of transportability, the foregoing group of functional regression models has put in evidence the correlation trend between the shape parameters of the water retention and cumulative particle-size distribution functions. This *shape similarity* formed the main hypothesis for the few semophysical models proposed in the literature (e.g., Arya and Paris, 1981; Haverkamp and Parlange, 1986). The shape-similarity hypothesis (different from shape identity) was recently confirmed by an analysis of the soil database GRIZZLY presented by Haverkamp and Reggiani (2002). Similar to the definition of the water retention shape index P_{wr} (Equation 6.23), the shape of the cumulative particle-size distribution functions can be expressed by the particle-size shape index P_{pa} :

$$P_{pa} = \frac{MN(k_M + MN)}{k_M + 2MN} \quad (6.33)$$

where M , N , and k_M are given by Equation 6.18 and Equation 6.19. The positive correlation between P_{wr} and P_{pa} is clearly shown in Figure 6.12a. This result confirms the fact that for a large number of soils the water retention shape parameter depends mainly on texture. A similar analysis for the particle-size scale parameter (D_{pg}) of Equation 6.18 and the pressure scale parameter (h_g) of Equation 6.21 shows no correlation (Figure 6.12b). This observation highlights the danger of predicting the soil structure dependent scale parameters from soil textural information alone.

The two semophysical methods considered here, which are both texture-based, take into account of the soil structural effect through the dry bulk density (ρ_d) embedded in the *tortuosity* parameter (α) for the Arya–Paris model (1981) or through the proportionality factor between pore-size and associated particle diameter for the Haverkamp–Parlange model (1986). An extensive review of both semophysical models is given in the *Soil Science Society Book on “Methods of Soil Analysis”* (Haverkamp and Reggiani, 2002).

6.2.6.1.3.1 *The Arya–Paris Model (1981)* which is the most common used, involves dividing the cumulative particle-size distribution function into a number of fractions, assigning a pore volume and a

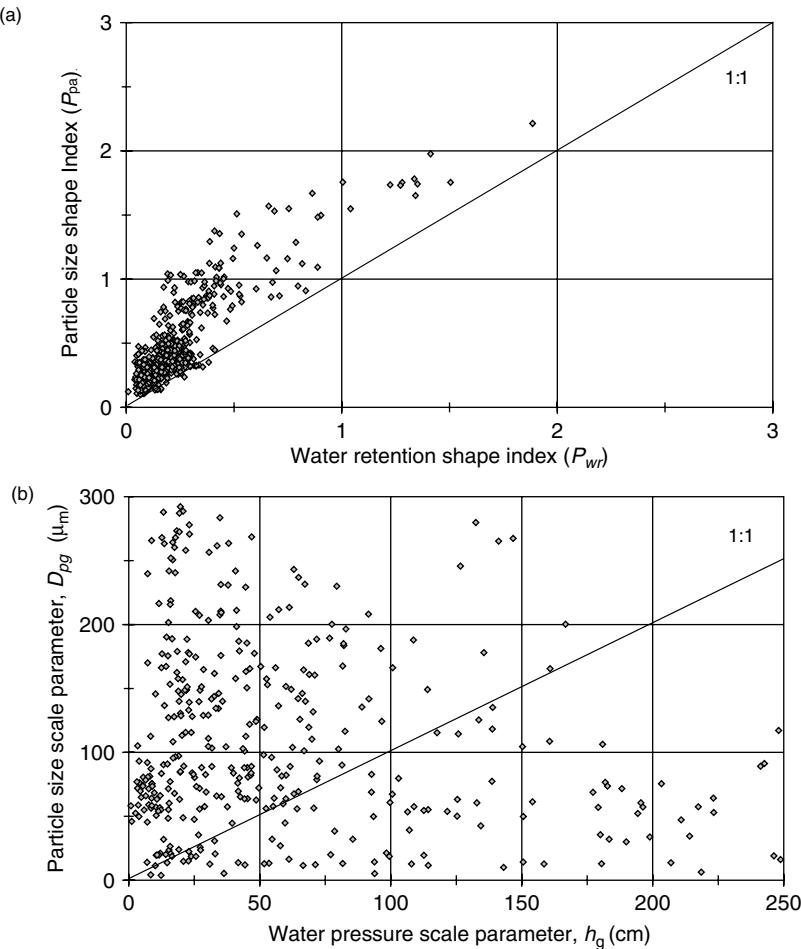


FIGURE 6.12 (a) The particle-size shape index P_{pa} as a function of the water retention shape index P_{wr} calculated for the 660 soils of the GRIZZLY soil database. (b) Pressure scale parameter h_g as a function of the particle-size scale parameter D_{pg} calculated for the 660 soils of the GRIZZLY soil database. (Adapted from Haverkamp, R. and Reggiani, P. 2002. Physical-based water retention prediction models. In *Methods of Soil Analysis, Part 4, Physical Methods*, Dane, J. H., and Clarke, G. C. [Eds.], Soil Science Society of America Book Series, *Soil Sci. Soc. Am. Inc.*, Madison, WI, pp. 762–777.)

volumetric water content to each fraction, and then computing a representative mean pore radius (R_i) and a corresponding water pressure head (h_i) value. They proposed the following nonlinear relationship to relate the pore radius (R_i) to the mean particle radius (D_{pi}):

$$R_i = D_{pi} \left[\frac{2e(nb_i)^{1-\alpha}}{3} \right]^{0.5} \quad (6.34)$$

where e is the void ratio; and nb_i is the number of particles in each size class calculated from:

$$nb_i = \frac{6W_i}{\pi\rho_s(D_{pi})^3} \quad (6.35)$$

W_i is the solid mass per unit sample mass corresponding with D_{pi} chosen in such a way that the sum of W_i is unity; ρ_s is the particle density taken equal to 2.65 g/cm^3 ; and α is an empirical factor referred to

as the *tortuosity factor* fixed at a constant value $\alpha = 1.38$. Subsequently, the pore volume dVp_i associated with each size fraction is defined by:

$$dVp_i = \frac{\pi}{6} e(nb_i)(Dp_i)^3 \quad (6.36)$$

Finally, the volumetric water content values are calculated from the pore volumes dVp_i by:

$$\theta_i = \rho_d \sum_{j=1}^{nb_i} dVp_j \quad (6.37)$$

In doing so, it is assumed that the pore volumes generated by each size fraction are progressively accumulated, from smallest to largest size fraction, and filled with water. The largest cumulative pore-size corresponds to porosity ε . Even though the Arya–Paris model is extremely easy to use, it has some slight drawbacks.

1. The first problem lies in the fact that it considers the porosity ε as the upper limit of the pore space to be filled with water which is rarely the case for field soils. Hence, it does not account for trapped air and systematically overestimates the water content scale parameter. This obviously induces errors in the predicted water content values which are most visible at the wet end of the water retention curve. As indicated by Arya and Paris (1982), this effect could be accounted for in the model at cost of additional soil information. In a later study, Arya et al. (1999) introduced a correction factor θ_s/ε in front of the sum term in Equation 6.37. However, the presence of trapped air translates in a disguised way the problem of tortuosity and, consequently, not only affects the water content scale parameter but also the pressure scale parameter of the water retention curve.
2. The second difficulty lies in the definition of the parameter α , which relates the particle-size to the pore radius. Arya and Paris proposed a constant average α -value ($\alpha = 1.38$) calculated over a limited number of soils. However, several researchers (e.g., Mishra et al., 1989; Gupta and Ewing, 1992) have suggested that predictions of the water retention curve could be improved if α is allowed to vary over the range of particle-sizes. Haverkamp et al. (1998b) have tested the validity of a constant α for more than 1000 soils taken from the UNSODA database (Leij et al., 1996). For the reported soils, both the water retention and particle-size distribution curves are available and the value of α can easily be calculated. Defining the normalized water content as $\theta_i^* = \theta_i/\theta_s$, a series of θ_i^* values was chosen at various intervals between zero and unity. Subsequently, for each soil a value of α was calculated for each θ_i^* . The results shown in Figure 6.13, challenge the validity of a constant α . The values vary not only as a function of soil type, but also as a function of water content. Neither at the wet nor at the dry end of the water retention curve is the constant average value of $\alpha = 1.38$ satisfied. Prompted by the different studies on a non-constant α value, Arya et al. (1999) has proposed an empirical method that allows the estimation of the tortuosity parameter α_i as a function of the number of soil particles (nb_i) and, hence, as a function of water content (θ_i). The method was developed by the use of “real” soil data taken from the UNSODA database (Leij et al., 1996). For five different soil textural classes of the USDA soil texture triangle, namely, sand, sandy loam, loam, silt loam, and clay (see Figure 6.3), about five soils were selected for which the linear relationship between $\log(W_i/Dp_i^3)$ and $\log(Nb_i)$ have been explored (where Nb_i is the number of “real” soil particles calculated from the UNSODA water retention data). Arya et al. (1999) established the following regression equation:

$$c_i = 1 + \left[\frac{c_1 + c_2 \log(W_i/(Dp_i)^3)}{\log(Nb_i)} \right] \quad (6.38)$$

where the parameters c_1 and c_2 are empirical regression parameters with different values for different textural classes (Table 6.4). Haverkamp et al. (2002) have tested the validity of the new formulation

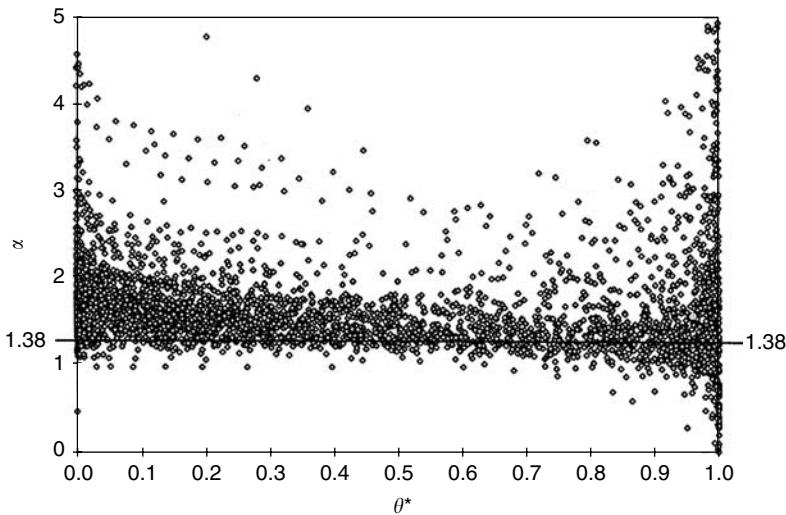


FIGURE 6.13 Evolution of the tortuosity factor α of the Arya and Paris (1981) model as function of θ_i^* calculated for more than 1000 different soils of the UNSODA and GRIZZLY soil databases.

TABLE 6.4 Regression Coefficients According to Equation 6.38

Textural Class	Regression Coefficients	
	c_1	c_2
Sand	-1.415	0.489
Sandy loam	-2.353	0.773
Loam	-0.644	0.395
Silt loam	-1.425	0.353
Clay	-1.559	0.305

Source: From Arya, L. M., Leij, F. J., van Genuchten, M. Th., and Shouse, P. J. 1999. *Soil Sci. Soc. Am. J.*, 63: 510–519. With permission.

of α on the same population of soils as used before. The authors showed that the new formulation of α systematically overestimated the best-fit α -values with roughly 25 to 50%, which is not a serious improvement with respect to the initial definition with a constant α . However, for each particular soil textural class the scatter between the predicted and calculated α -values was considerably reduced putting in evidence that the principles used for the new formulation of α are basically sound but that the regression analysis calculated over a small soil population of roughly five soils is not representative for the soil taken from UNSODA database.

3. The third and last difficulty of the Arya–Paris model (1981) concerns the problem of hysteresis. As the cumulative particle-size distribution is unique for a given soil, only one associated water retention curve can be predicted, either the main wetting curve or the main drying curve. To pass from one main curve to the other and/or along scanning curves, the water retention prediction model should theoretically be coupled with a hysteresis model that does not require any supplementary information. Although the discrete equations used by the Arya–Paris model were not formulated to facilitate such coupling, it is feasible to do so.

In conclusion, the Arya–Paris model is definitely very user-friendly which explains the fact that the model is the most widely used.

6.2.6.1.3.2 The Haverkamp–Parlange Model (1986). The second semi-physical model proposed by Haverkamp and Parlange (1986) is certainly much more complicated and has not the easiness of use of the Arya–Paris model. Their method allows the direct estimation of the parameters of Equation 6.20 for sandy soils without organic matter. The predicted $h(\theta)$ curve was then associated to the boundary wetting curve (BWC). Coupled with the hysteresis model proposed by Parlange (1976) the full family of wetting curves can be predicted. For the sake of simplicity, the authors assumed a linear relationship between the mean pore radius (R) and the corresponding particle radius (D_p):

$$R = \gamma D_p \quad (6.39)$$

where γ is an empirical packing factor to be determined through fitting from textural data.

This method presents the advantage of interpreting the cumulative particle-size distribution function in its continuous form, but it uses the extremely simple relationship between R and D_p (Equation 6.39) which is perhaps valid for structureless soils such as pure sand soils, but which is questionable for non-sandy soils. Moreover, the use of the hysteresis concept obliged the authors to choose different type-equations for $F_p(D_p)$ (Equation 6.18) and $h(\theta)$ (Equation 6.20). This complicated the model description and, hence, made the model more difficult to follow and less user-friendly as compared to the Arya–Paris model addressed before.

6.2.6.1.3.3 Other Methods In the first edition of the *Handbook of Groundwater Engineering*, Chapter 5, Haverkamp et al. (1998b) introduced a method for predicting the van Genuchten water retention shape parameter m (Equation 6.18) as a function of the particle-size shape parameter M (Equation 6.21). Extending this method to the concept of shape indices as given by Equation 6.23 and Equation 6.33, we have:

$$P_{pa} = P_{wr}[1 + \gamma s] \quad (6.40)$$

where γs is a tortuosity factor defined as a function of the fractal dimension s introduced by (Fuentes et al., 1998):

$$\gamma s = \frac{2s - 1}{2s[1 - s]} \quad (6.41)$$

and calculated as the root of the equation:

$$[1 - \varepsilon]^s + \varepsilon^{2s} = 1 \quad (6.42)$$

Parameter ε is the soil porosity defined by Equation 6.3. Application of this method in the algorithm BEST (Lassabatère et al., 2005) gives promising results.

6.2.6.2 Hydraulic Conductivity Relation

Methods used to estimate the hydraulic conductivity curve rely on capillary model hypotheses. Most models proposed in the literature (e.g., Millington and Quirk, 1961; Mualem, 1976) are based on the Burdine's equation (1953) and make use of the water retention equation $h(\theta)$:

$$K^* = [\theta^*]^{p\beta} \left[\frac{\int_0^{\theta^*} 1/(h(\bar{\theta}^*)^2) d\bar{\theta}^*}{\int_0^1 1/(h(\theta^*)^2) d\theta^*} \right] \quad (6.43)$$

where θ^* and K^* are the dimensionless water content and hydraulic conductivity, respectively, given by Equation 6.27; and $p\beta$ is a parameter generally associated with the effect of tortuosity. The differences between the various models arise from the hypotheses introduced to describe the pore structure (i.e., tortuosity) and its interaction with the relative permeability (Equation 6.17). Solving Equation 6.43 allows deriving an expression that relates the hydraulic conductivity shape parameter (i.e., η of Equation 6.25) as function of the water retention shape parameters (i.e., λ or m of Equation 6.20 or Equation 6.21 respectively) and the tortuosity factor:

$$\eta = \frac{2}{\lambda} + 2 + p\beta \quad (6.44)$$

Different values of $p\beta$ can be derived dependent on the capillary model chosen, that is, Childs and Collis George (1950) used $p\beta = 0$; Mualem (1976) $p\beta = 1/2$; Burdine (1953) $p\beta = 1$; and Millington and Quirk (1961) $p\beta = 4/3$. Although these $p\beta$ -values have to be interpreted with considerable caution (as they were mostly based on pure intuition of the different authors combined with some subjective statistical analyses), they definitely have the benefit of indicating the existence of some complicity between the shape parameters of the water retention and hydraulic conductivity functions. The ideal case would be if the $p\beta$ -value could be individualized for each soil.

The scale parameter K_s is strongly related to soil structure. Among the different soil hydraulic characteristic parameters, the saturated hydraulic conductivity is the parameter which is the most influenced by effects such as macropores, stones, fissures, cracks, and other irregularities formed for various biological and mechanical reasons. Hence, it is the parameter which is the most difficult to predict. The models proposed in the literature either give estimations of the capillary conductivity value or are based on site- and soil-specific databases. The results should therefore be considered with caution when applied for field studies.

Mishra and Parker (1990) used the Mualem model (1976) with the van Genuchten water retention function (Equation 6.21) to obtain a closed-form expression of the saturated hydraulic conductivity:

$$K_s = c_1 \frac{[\theta_s - \theta_r]^{2.5}}{h_g^2} \quad (6.45)$$

where c_1 is a constant including the effects of fluid characteristics and the porous media geometric factor; it has a value of $108 \text{ cm}^3/\text{sec}$ when K_s is expressed in cm/sec ; h_g is the van Genuchten (1980) pressure scale parameter. The authors derived a similar predictive equation by the use of the Brooks and Corey (1964) water retention equation (Equation 6.20):

$$K_s = c_1 \frac{[\theta_s - \theta_r]^{2.5}}{h_{bc}^2} \left[\frac{\lambda}{1 + \lambda} \right]^2 \quad (6.46)$$

where λ and h_{bc} are the Brooks and Corey (1964) shape and scale parameters respectively. Ahuja et al. (1985) used the general Kozeny–Carman approach to determine the saturated hydraulic conductivity from the effective porosity ($\varepsilon - \theta_r$):

$$K_s = c_2 [\varepsilon - \theta_r]^{c_3} \quad (6.47)$$

where c_2 is equal to 1058 cm/h when K_s is expressed in cm/h ; and c_3 takes a value of 4 or 5.

Cosby et al. (1984) derived a statistical regression between the saturated hydraulic conductivity and the clay, silt, and sand fractions of soil (Table 6.1). In a similar way Rawls and Brakensiek (1989) presented a statistical model using particle-size distribution and porosity (Table 6.2).

6.2.7 Spatial Variability of Soil Water Properties

The difficulty of parametrization of soil moisture dynamics lies in the characterization and understanding of the different processes involved and the scale at which they take place: microscale, macroscale, and megascale. As underlined by Kabat et al. (1997), water transfer in the vadose zone is characterized by laminar flow with small mixing ratios, whereas atmospheric fluxes occur in the turbulent domain causing large mixing ratios, resulting in different characteristic spatial scales. There is an urgent need to characterize and incorporate the spatial variability into current modeling approaches. The use of distributed models can only represent explicitly heterogeneities of processes with a characteristic length scale larger than the model grid size. This type of approach still uses some lumping at the grid level, and there remains the problem of characterizing the variability at the subgrid level.

Water movement in the unsaturated zone is affected by intrinsic parameters such as soil characteristics and by external factors such as rainfall and management practices. The movement is thus characterized by a high spatial variability. It is important, concerning the concept of spatial variability, to separate the concept of modeling of the unsaturated flow process and the parametrization of the soil hydraulic characteristics. In addition to the obvious lack of theoretical knowledge of the behavior of the flow equation in the unsaturated zone across temporal and spatial scales, the nonlinearity of the flow equations adds also to the complexity of representing spatial variability. The validity of extrapolating concepts developed at the local scale to larger scale may be questioned. A second major problems arises for the model parametrization. Is there a physical meaning for a characteristic curve representing 1 km² grid cell?

As mentioned earlier (Section 6.2.4, Equations 6.20 to 6.26), characterizing soil properties require the determination of five unknown parameters: two shape parameters (i.e., m and η , or λ and η) and three scale parameters (θ_s , K_s and h_g or h_{bc}). The shape parameters as shown by Haverkamp and Reggiani (2002) are strongly linked to soil texture, while the scale parameters depend mainly on the soil structural properties. Textural properties are characterized by a smaller underlying variability than structural properties. This is mainly due to the fact that anthropogenic factors, such as management practices, have a much larger impact on structural properties than on textural properties. Thus, it would be expected to be able to capture and determine the shape parameters λ , m , and η (*static characteristics*) across different scales by the use of pedotransfer functions. However, the scale parameters θ_s and K_s (*dynamic characteristics*) will unlikely be determined from textural information, especially for K_s , which is a driving parameter for modeling surface and subsurface flow. It is furthermore indispensable to note that soil characteristics have a meaning for capillary flow and that this concept will not hold in field conditions where macropore and preferential flow may be present and prevail over capillary flow. Thus the use of pedotransfer functions for estimating the dynamic parameters are doomed to fail. Haverkamp et al. (1998a) proposed an alternate method to determine the scale parameters by coupling the soil characteristics with the unsaturated flow equation. This methodology which will be presented in more detail in a later section of this chapter consists in the upscaling of the transfer equation and the derivation of scale-invariant soil parameters.

Another approach for incorporating spatial variability is the determination of effective parameters which are supposed to be valid over a representative elementary volume (e.g., Woods et al., 1995). However, because of the non linearity between soil water fluxes and the soil properties, and the extrapolation of theories developed at the local scale to large scale, the validity of the concept of effective parameters can be questioned.

6.3 Conceptual Aspects of Unsaturated Soil Water Flow

For the following analysis, soil water movement is supposed to be isothermal and one-dimensional, whereas the influence of swelling and shrinking of the soil porous material is not taken into consideration. For a more detailed description of water movement in swelling soils we refer to work published by Smiles and Rosenthal (1968); Philip (1969a); Smiles (1974, 1995); Groenevelt and Parlange (1974); Sposito (1973, 1975a, 1975b, 1975c); and Gérard-Marchant et al. (1997). Moreover, the case of two-phase flow will not be addressed in this study. Under general field conditions, the effects due to air displacement can be

considered negligible because of the large differences in dynamic viscosity. Only in particular situations (such as flooding, storm rains, or stratified soil profiles) may air not be free to escape, resulting in an increased air pressure ahead of the wetting front and, hence, in a reduced infiltration rate (see also Peck, 1965a, 1965b; Brustkern and Morel-Seytoux, 1970; Phuc and Morel-Seytoux, 1972; Vachaud et al., 1973, 1974; Starr et al., 1978; Touma et al., 1984).

6.3.1 General Flow Equations

For 1D vertical flow the continuity equation takes the form:

$$\frac{\partial \theta}{\partial t} = -\frac{\partial q}{\partial z} \quad (6.48)$$

where θ is volumetric soil water content; q is soil water flux; z is depth taken positive downward; and t is time. Combined with the generalized Darcy's law (1856):

$$q = -K(\theta) \frac{\partial H}{\partial z} \quad (6.49)$$

the soil water transfer equation can be written as:

$$\frac{\partial \theta}{\partial t} = \frac{\partial}{\partial z} \left[K(\theta) \left(\frac{\partial h}{\partial z} - 1 \right) \right] \quad (6.50)$$

where h is soil water pressure head relative to atmospheric pressure ($h \leq 0$); K is hydraulic conductivity as function of θ ; and H is hydraulic head defined as:

$$H = h(\theta) - z \quad (6.51)$$

For non-swelling soils, the hydraulic head (H) represents the energy of soil water per unit of weight at a given depth (z); the possible osmotic component which arises from the presence of solutes in the soil water is neglected (see also Section 6.2.3). Equation 6.50 can be expressed as a θ -dependent equation by introducing the concept of soil water diffusivity (Childs and Collis-George, 1950):

$$\frac{\partial \theta}{\partial t} = \frac{\partial}{\partial z} \left[D(\theta) \frac{\partial \theta}{\partial z} - K(\theta) \right] \quad (6.52)$$

where $D(\theta)$ is diffusivity defined by

$$D(\theta) = K(\theta) \frac{dh}{d\theta} \quad (6.53)$$

Equation 6.52 is generally known as the Fokker–Planck equation.

Similarly Equation 6.50 can be expressed as function of only soil water pressure head (h) by introducing the concept of specific capacity:

$$C(\theta) \frac{\partial h}{\partial t} = \frac{\partial}{\partial z} \left[K(\theta) \left(\frac{\partial h}{\partial z} - 1 \right) \right] \quad (6.54)$$

where C is specific capacity given by $C(\theta) = d\theta/dh$. Equation 6.54 referred to as the Richards (1931) equation, is the 1D, isothermal, unsaturated soil water transfer equation generally used in vadose zone hydrology, especially when coupled saturated and unsaturated flow problems are considered. The use of Equation 6.52 with volumetric water content (θ) as independent variable, without doubt

causes computational difficulties when applied for numerical simulation of water movement in regions close to saturation.

The solution of Equation 6.52 and Equation 6.54 subject to given initial and boundary conditions describes the evolution of the water content profiles $\theta(z, t)$ as a function of depth and time. The initial condition imposed on $\theta(z, t)$ is given by:

$$\theta(z, 0) = \theta_0(z) \quad (6.55)$$

where θ_0 is the initial water content value. The upper boundary condition on $\theta(z, t)$ is either the *Dirichlet* concentration condition:

$$h(0, t) = h_1(t) \quad \text{and/or} \quad \theta(0, t) = \theta_1(t) \quad (6.56)$$

or the *Neumann* flux condition:

$$\left[K - K \frac{\partial h}{\partial z} \right]_{z=0} = q_1(t) \quad (6.57)$$

where h_1 is the surface soil water pressure, θ_1 is the corresponding volumetric soil water content and q_1 is the flux at the soil surface. Both values of h_1 and q_1 may be positive, zero, or negative.

6.3.2 Infiltration

Infiltration concerns the physical process of water entry into the soil through the surface. It is governed by either the concentration type boundary condition (e.g., water ponding at the soil surface) or the flux type surface boundary condition (e.g., rainfall). For a nonuniform initial water content profile, $\theta_0(z)$, the 1D cumulative infiltration (I) expressed in volume per unit surface is defined by the integral of the flux at the soil surface:

$$I(t) = \int_0^t q_1(\bar{t}) d\bar{t} \quad (6.58)$$

For the particular case of infiltration under uniform initial conditions:

$$\theta(z, 0) = \theta_0, \quad t = 0 \text{ for } z \geq 0 \quad (6.59)$$

and constant surface boundary conditions:

$$\begin{aligned} \theta(0, t) &= \theta_1, \quad z = 0 \text{ for } t \geq 0 \\ h(0, t) &= h_1, \quad z = 0 \text{ for } t \geq 0 \end{aligned} \quad (6.60)$$

the cumulative infiltration Equation 6.58 takes the form:

$$I(t) = K_0 t + \int_0^{z_f} [\theta(z, t) - \theta_0] dz \quad (6.61)$$

where $z_f(t)$ is depth of the infiltration wetting front; and K_0 is the hydraulic conductivity at initial water content (θ_0). The constant water pressure value h_1 can be negative or positive equal to the ponded water depth (h_{surf}) imposed at the soil surface ($h_1 = h_{\text{surf}} \geq 0$).

Generally speaking, the time limits of cumulative infiltration are well defined:

$$\lim_{t \rightarrow 0} I(t) = 0 \quad \text{and} \quad \lim_{t \rightarrow \infty} I(t) = \infty \quad (6.62)$$

The infiltration rate, q_1 at the soil surface is defined as $q_1 = dI/dt$ with the time limits:

$$\lim_{t \rightarrow 0} q_1(t) = \infty \quad \text{and} \quad \lim_{t \rightarrow \infty} q_1(t) = K_1 \quad (6.63)$$

where K_1 is the hydraulic conductivity corresponding to the surface boundary condition $\theta = \theta_1$ and/or $h = h_1$; obviously, K_1 becomes equal to $K_1 = K_s$ for $\theta_1 = \theta_s$ and/or $h_1 \geq 0$.

More details on the empirical infiltration equations are given in Chapter 7.

6.4 Analytical Solutions of the Unsaturated Flow Equation

Although the solution of the Richards (and/or Fokker–Planck) equation under general initial and boundary conditions can only be obtained through numerical simulation, exact analytical solutions under idealized conditions exist. As most of the numerical codes are now available from reliable commercialised software packages (e.g., HYDRUS-2D) we rather prefer to explore in this section some aspects of the analytical solutions. Good summaries of most of the numerical problems involved in solving Equation 6.54 can be found in literature. Many different techniques are explored, for example, finite difference, finite element, and/or flux volume methods.

The mathematical difficulties associated with the analytical solutions for $\theta(z, t)$ and/or infiltration, can be best discussed through the initial and boundary conditions given by Equation 6.59 and Equation 6.60. The constant pressure head (h_1) can be chosen negative, zero, or positive. In this section we only consider the problem of constant pressure head infiltration. The flux condition is fully discussed in Chapter 7. The concept of *sorptivity* is addressed in some detail because it represents an important integral variable which links water retention and hydraulic conductivity characteristics. We then give a short discussion of two positive constant head infiltration equations required for the last section dealing with the spatial variability of water flow.

6.4.1 Constant Negative (or Zero) Pressure Head Condition at the Soil Surface

Different analytical solutions have been proposed in the literature (e.g., Philip, 1955, 1957a; Parlange, 1971b, 1975; Philip and Knight, 1974; Babu, 1976; Parlange et al., 1982; Broadbridge, 1990). A good summary can be found in Parlange et al. (1999).

6.4.1.1 Without Gravity Effects

In our discussion we first consider the case where gravity effects can be ignored (such as would be the case for horizontal infiltration). The water movement equation (Equation 6.52) is then reduced to:

$$\frac{\partial \theta}{\partial t} = \frac{\partial}{\partial z} \left[D(\theta) \frac{\partial \theta}{\partial z} \right] \quad (6.64)$$

subject to the conditions:

$$\begin{aligned} \theta(z, 0) &= \theta_0, & t = 0 \text{ for } z \geq 0 \\ \theta(0, t) &= \theta_1, & z = 0 \text{ for } t \geq 0 \end{aligned} \quad (6.65)$$

The partial differential equation (Equation 6.64) along with corresponding initial and boundary conditions can be reduced to an ordinary integral-differential equation by introducing the Boltzmann (1894) similarity variable χ_1 defined by $\chi_1(\theta) = z/t^{1/2}$:

$$\int_{\theta_0}^{\theta} \chi_1(\bar{\theta}) d\bar{\theta} = -2D(\theta) \frac{d\theta}{d\chi_1(\theta)} \quad (6.66)$$

subject to the condition:

$$\theta = \theta_1 \quad \text{for } \chi_1(\theta_1) = 0 \quad (6.67)$$

Equation 6.66 and Equation 6.67 are generally known as the similarity equation of Bruce and Klute (1956). It shows that there exists a fundamental dynamic similarity between z and t associated with the solution of the diffusivity equation (Equation 6.66); it reveals that the advance of the normalized wetting front is proportional to $t^{1/2}$.

Integrating the Boltzmann transform from θ_1 to θ_0 , yields the cumulative infiltration equation:

$$I(t) = S_1(\theta_1, \theta_0) \sqrt{t} \quad (6.68)$$

where $S_1(\theta_1, \theta_0)$ is the sorptivity defined by:

$$S_1(\theta_1, \theta_0) = \int_{\theta_0}^{\theta_1} \chi_1(\theta) d\theta \quad (6.69)$$

The sorptivity S_1 , which is clearly specific for the initial (θ_0) and boundary (θ_1) conditions encountered during each infiltration event, characterizes the ability of the soil to absorb water in the absence of gravity. It was initially presented by Green and Ampt (1911) and later coined by Philip (1957b).

Many solutions of Equation 6.66 have been proposed over time. One of the first reported in the literature is that of Crank and Henry (1949). However, their iterative method suffered from slow convergence and was often found slightly erroneous in the region close to initial water content θ_0 . To overcome this problem Philip (1955) proposed a modified iterative procedure with the value of χ_1 close to θ_0 truncated at a fixed value calculated from an analytical inverse “erfc”-function. Even though the procedure of Philip improved the solution of Crank and Henry (1949) with respect to the convergence criterion, it still remained rather imprecise close to θ_0 , and today this method is considered somewhat tedious and rather imprecise for the calculation of sorptivity.

A different solution of Equation 6.66 was proposed by Parlange (1971a) who started with a slightly modified version of Equation 6.64:

$$\frac{\partial z}{\partial t} + \frac{\partial}{\partial \theta} \left[D(\theta) \frac{\partial \theta}{\partial z} \right] = 0 \quad (6.70)$$

Integration of Equation 6.70 is carried out iteratively in two steps yielding the relation for $\chi_1(\theta)$ and hence $z(\theta, t)$ of the form:

$$\chi_1(\theta) = \frac{z(\theta, t)}{\sqrt{t}} = \int_{\theta}^{\theta_1} \frac{D(\bar{\theta})}{F_{PA}(\bar{\theta})} d\bar{\theta} \quad (6.71)$$

where $F_{PA}(\theta)$ is given by:

$$F_{PA}(\bar{\theta}) = \frac{\int_{\theta_0}^{\bar{\theta}} \left[\int_{\hat{\theta}}^{\theta_1} D(\tilde{\theta}) d\tilde{\theta} \right] d\hat{\theta}}{\left[2 \int_{\theta_0}^{\theta_1} [\theta - \theta_0] D(\theta) d\theta \right]^{1/2}} \quad (6.72)$$

Parameters $\bar{\theta}$, $\hat{\theta}$, and $\tilde{\theta}$ are simply the integral variables. Equation 6.71 gives the moisture content profile $\chi_1(\theta)$ directly for any soil once $D(\theta)$ is known. Obviously, the approach can easily be extended to the case where gravity is not negligible (Parlange, 1971b).

In a later article Philip (1973) used this iterative concept by reformulating the F -function as:

$$F(\theta) = \frac{q(z, t)}{q_1(t)} = \frac{\int_{\theta_0}^{\theta} \chi_1(\bar{\theta}) d\bar{\theta}}{\int_{\theta_0}^{\theta_1} \chi_1(\theta) d\theta} \quad (6.73)$$

where $F(\theta)$ was referred to as the *flux-concentration* relation. Integration of Equation 6.70 after substitution of Equation 6.73 yields the well-known equation of $\chi_1(\theta)$:

$$\chi_1(\theta) = \frac{2}{S_1(\theta_1, \theta_0)} \int_{\theta}^{\theta_1} \frac{D(\bar{\theta})}{F(\bar{\theta})} d\bar{\theta} \quad (6.74)$$

The substitution of Equation 6.74 into Equation 6.69 allows us to redefine the sorptivity expression as:

$$S_1^2(\theta_1, \theta_0) = 2 \int_{\theta_0}^{\theta_1} \frac{[\theta - \theta_0] D(\theta)}{F(\theta)} d\theta \quad (6.75)$$

Probably the best understanding of the structure of any solution of Equation 6.66 can be based on the analytical expansion technique of Heaslet-Alksne (1961) which was applied by Parlange et al. (1992). A good compromise for the F -function which balances accuracy with simplicity (Elrick and Robin, 1981) was given by Parlange (1975):

$$F(\theta) = \frac{2[\theta - \theta_0]}{[\theta_1 + \theta - 2\theta_0]} \quad (6.76)$$

resulting in the sorptivity equation used for many studies (e.g., Ross et al., 1996):

$$S_1^2(\theta_1, \theta_0) = \int_{\theta_0}^{\theta_1} [\theta_1 + \theta - 2\theta_0] D(\theta) d\theta \quad (6.77)$$

6.4.1.2 With Gravity Effects

For the case where the gravity effects cannot be neglected the analytical solutions are mostly expressed through the Fokker–Planck equation (Equation 6.52). The infiltration solution of Philip (1957a, 1957b)

which originally was developed for the surface boundary condition $h_1 = 0$ and $\theta_1 = \theta_s$, is written in terms of a series expansion in powers of $t^{1/2}$ as:

$$z(\theta, t) = \chi_1(\theta)t^{1/2} + \chi_2(\theta)t + \chi_3(\theta)t^{3/2} + \dots \quad (6.78)$$

where the first term, $\chi_1(\theta)$ embodies the influence of the capillary forces on the flow process, and the following terms, for example, $\chi_2(\theta)$ and $\chi_3(\theta)$, reflect the gravity effect on infiltration. Integration of Equation 6.78 gives the cumulative infiltration equation, $I(t)$:

$$I(t) = S_1(\theta_1, \theta_0)t^{1/2} + [K_0 + S_2(\theta_1, \theta_0)]t + O(t^{3/2}) + \dots \quad (6.79)$$

where the definition of $S_1(\theta_1, \theta_0)$ is given above by Equation 6.69 and:

$$S_2(\theta_1, \theta_0) = \int_{\theta_0}^{\theta_1} \chi_2(\theta)d\theta \quad (6.80)$$

Like the sorptivity ($S_1(\theta_1, \theta_0)$) the second term ($S_2(\theta_1, \theta_0)$) is specific for the initial and boundary conditions encountered during each particular infiltration event. The third and higher terms of Equation 6.79 can easily be calculated by transforming the time series expansion into a recurrence series as a function of the first and second term (Haverkamp et al., 1990).

In theory, the Philip infiltration equation is physically based which makes it suitable for the simulation (*a posteriori*) of infiltration events once the soil characteristics are fully known. However, in practice, when the solution is used for prediction purposes (*a priori*) some ambiguity exists over the correct interpretation of Equation 6.79:

1. Though an accurate and simple estimation of sorptivity has been developed by Parlange (1975) in the form of Equation 6.77, no simple approximation exists for the second integral term. Generally, the value of $S_2(\theta_1, \theta_0)$ is associated to a constant A_2 calculated from curve fitting with:

$$I(t) = A_1\sqrt{t} + [K_0 + A_2]t \quad (6.81)$$

over experimental cumulative infiltration data where A_2 is confined over the interval (Youngs, 1968; Philip, 1969b; Talsma and Parlange, 1972):

$$\frac{K_1 - K_0}{3} \leq A_2 \leq 2\frac{K_1 - K_0}{3} \quad (6.82)$$

As Equation 6.81 neglects deliberately all higher order terms ($O(t^{3/2})$) of the original time series expansion (Equation 6.79), it is evident that the optimized values of A_1 and A_2 reflect to some extent this truncation effect. Consequently, A_1 and A_2 are just specific ad hoc constants which become rather delicate to compare with the physically defined integral parameters $S_1(\theta_1, \theta_0)$ and $S_2(\theta_1, \theta_0)$ (Haverkamp et al., 1988).

2. Then, the definitions of $S_1(\theta_1, \theta_0)$ and $S_2(\theta_1, \theta_0)$ were originally developed for the case of infiltration with negative or zero head surface boundary condition ($h_1 \leq 0$ and $\theta_1 \leq \theta_s$). For the case of positive head infiltration ($h_1 \geq 0$ and $\theta_1 = \theta_s$) the integrals S_1 and S_2 have to be reformulated. This can easily be done for the sorptivity definition of S_1 :

$$S_1^2(h_{\text{surf}}, h_0) = S_1^2(\theta_s, \theta_0) + 2K_s h_{\text{surf}}[\theta_s - \theta_0] \quad (6.83)$$

where h_0 stands for the initial soil water pressure $h(\theta_0)$ and h_{surf} is the ponded water depth imposed at the soil surface. However, for the second integral term (S_2) it is far more difficult to adjust Equation 6.80.

3. Finally, the main problem for the application of the Philip (1957a) solution lies in the fact that the time series of Equation 6.78 becomes divergent for large times no matter how many terms are developed. Thus, the solution is only valid for a limited time range. The time limit is mostly set at t_{grav} :

$$t_{\text{grav}} = \left[\frac{S_1(\theta_s, \theta_0)}{K_s - K_0} \right]^2 \quad (6.84)$$

a time for which the gravity forces are supposed to become predominant over the capillary forces during the infiltration process (Philip, 1969b).

To overcome this time limit problem Parlange et al. (1982) developed a quasi-exact infiltration solution of the Fokker–Planck equation (Equation 6.52) valid over the entire time range $t \in [0, \infty]$. This work made use of a special double integration procedure of the water transfer equation. In later studies by Parlange et al. (1985a) and Haverkamp et al. (1990), this infiltration solution which originally was developed for a zero or negative surface head condition, was extended to a positive head boundary condition ($h_1 \geq 0$). For the sake of brevity, we will report more details on this solution in Section 6.4.2 dealing with positive head infiltration. To transform the positive head solution to a negative head solution the value of h_{surf} has simply to be set equal to zero and θ_s has to be replaced by θ_1 when necessary.

6.4.2 Constant Positive Pressure Head Condition at the Soil Surface

Before discussing the quasi-exact positive head solution of Parlange et al. (1985a), it is preferred to address the somewhat simpler infiltration equation of Green and Ampt (1911); the latter solution which is widely used for hydrologic studies will be used in Section 6.5 to illustrate the scaling principles.

The hypotheses underlying the Green and Ampt approach for positive head infiltration assume a nominal wetting front in the form of a step function with a one-point hydraulic conductivity function ($K(\theta) = K_s$) behind the wetting front and a constant driving pressure ($h_{\text{surf}} - h_f$) pulling the wetting front downward through the unsaturated zone; h_f is a constant equivalent water pressure head at the wetting front. The corresponding water retention curve takes the form of a step function with $\theta = \theta_0$ for $h \leq h_f$ and $\theta = \theta_s$ for $h_f \leq h \leq 0$. In the original infiltration equation of Green and Ampt (1911), the initial value of hydraulic conductivity (K_0) was considered equal to zero. However, for hydrologic model studies this condition is quite often not verified. Hence, we choose to present here the generalized Green and Ampt infiltration equation with $K_0 \neq 0$ (e.g., Swartzendruber, 1987; Ross et al., 1996):

$$I(t) = K_s t + \frac{[\theta_s - \theta_0][h_{\text{surf}} - h_f(\theta_s, \theta_0)]K_s}{[K_s - K_0]} \ln \left\{ 1 + \frac{[I(t) - K_0 t][K_s - K_0]}{[\theta_s - \theta_0][h_{\text{surf}} - h_f(\theta_s, \theta_0)]K_s} \right\} \quad (6.85)$$

The corresponding infiltration flux (q_1) at the soil surface can be solved by:

$$t = \frac{[\theta_s - \theta_0][h_{\text{surf}} - h_f(\theta_s, \theta_0)]K_s}{[K_s - K_0]^2} \left\{ \frac{K_s - K_0}{q_1 - K_s} - \ln \left[1 + \frac{K_s - K_0}{q_1 - K_s} \right] \right\} \quad (6.86)$$

Note that Equation 6.85 and Equation 6.86 take the form of the classical Green and Ampt equations given in Chapter 7 (Equation 7.10) when K_0 is set equal to zero. When nonpositive constant head infiltration is involved ($h_1 \leq 0$), Equation 6.85 and Equation 6.86 can be applied equally well simply by replacing the values of θ_s and K_s by θ_1 and K_1 and setting h_{surf} equal to zero. Further details on the derivation of the Green and Ampt equation can be found in Chapter 7.

Although Equation 6.85 and Equation 6.86 are exact solutions of the Richards equation (Equation 6.54) for soils that correspond to the Green and Ampt configuration (hereafter referred to as “Green and Ampt”

soils), these conditions are seldom met under field conditions. For that reason, emphasis is given in the literature to relating the value of h_f in some theoretical way to more realistic soil hydraulic properties (e.g., Bouwer, 1964; Mein and Larson, 1973; Neuman, 1976). Even though the value of h_f changes slightly with time during infiltration (Haverkamp et al., 1988), a good estimation of h_f can be obtained from its short time limit (Neuman, 1976) derived from the second-order sorptivity definition as given by Parlange (1975):

$$h_f(\theta_1, \theta_0) = \frac{1}{2} \int_{h(\theta_1)}^{h(\theta_0)} \left[\frac{\theta_1 + \theta(h) - 2\theta_0}{\theta_1 - \theta_0} \right] \frac{K(h)}{K_1} dh \quad (6.87)$$

Parameter $h_f(\theta_s, 0)$ must be seen as the pressure scale parameter of the Green and Ampt water retention curve. Equation 6.87 clearly shows the dependence of $h_f(\theta_1, \theta_0)$ on initial conditions (θ_0) and boundary conditions (θ_1) encountered during infiltration; hence, all estimations that relate $h_f(\theta_1, \theta_0)$ in a statistical way to purely textural soil data (such as used by most PTFs [e.g., Rawls and Brakensiek, 1983]), lead necessarily to erroneous results.

As the integral parameter $h_f(\theta_1, \theta_0)$ is properly defined only for Green and Ampt soils, it is preferable to use the sorptivity concept instead (Equation 6.77 and Equation 6.83) when dealing with general field soils. For a negative head surface boundary condition ($h_1 \leq 0$) the relation between h_f and sorptivity is given by:

$$S_1^2(\theta_1, \theta_0) = 2K_s[\theta_1 - \theta_0]h_f(\theta_1, \theta_0) \quad (6.88)$$

and for the positive head surface boundary condition ($h_{\text{surf}} = h_1 > 0$) by:

$$S_1^2(h_{\text{surf}}, \theta_0) = 2K_s[h_{\text{surf}} - h_f(\theta_s, \theta_0)][\theta_s - \theta_0] \quad (6.89)$$

Substitution of Equation 6.88 and Equation 6.89 into Equation 6.85 and Equation 6.86 allows the Green and Ampt infiltration equation to be expressed in terms of sorptivity.

Parlange et al. (1982) developed a quasi-exact infiltration solution of the Fokker–Planck equation (Equation 6.52) valid over the entire time range $t \in [0, \infty]$. This work was concerned with a zero or negative surface head condition ($h_1 \leq 0$). In later studies by Parlange et al. (1985a) and Haverkamp et al. (1990), this infiltration solution was extended to a positive head boundary condition ($h_1 = h_{\text{surf}} \geq 0$). Without going in to too much detail we will report here a concise description of the approach.

Considering the case of infiltration into a soil at a uniform initial soil water pressure, $h_0 = h(\theta_0)$ (Equation 6.59) from a surface boundary source at either a constant negative head ($h_1 \leq 0$) or a constant positive head ($h_1 = h_{\text{surf}} \geq 0$), double integration of Equation 6.54 allows the expression of the water content profiles, $z(\theta, t)$, in the rigorous form:

$$z(\theta, t) = z_s(t) + \int_{\theta}^{\theta_s} \frac{D(\bar{\theta})}{F(\bar{\theta}, t)[q_1(t) - K_0] - [K(\bar{\theta}) - K_0]} d\bar{\theta} \quad (6.90)$$

where $\bar{\theta}$ is the integration variable; z_s is the depth of soil considered to be saturated when $h_1 \geq 0$; and $F(\theta, t)$ is the flux concentration relation defined for diffusivity and gravity driven infiltration (see also Equation 6.73). Obviously, $z_s = 0$ and $\theta_s = \theta_1$ when $h_1 \leq 0$. Integrating z by parts from θ_0 to θ_s yields then:

$$[I(t) - K_0 t] = h_{\text{surf}} \frac{[\theta_s - \theta_0]K_s}{q_1(t) - K_s} + \int_{\theta_0}^{\theta_s} \frac{[\theta - \theta_0]D(\theta)}{F(\theta, t)[q_1(t) - K_0] - [K(\theta) - K_0]} d\theta \quad (6.91)$$

For the integration of Equation 6.91, Parlange et al., (1982) introduce an integral soil parameter β expressed as a function of conductivity and diffusivity. The relation was later slightly generalized by Haverkamp et al., (1990) to:

$$\frac{[K(\theta) - K_0]}{[K_s - K_0]} = f(\theta) \left[1 - \frac{2\beta(\theta_s, \theta_0)}{S_1^2(\theta_s, \theta_0)} \int_{\theta}^{\theta_s} \frac{[\bar{\theta} - \theta_0]}{f(\bar{\theta})} D(\bar{\theta}) d\bar{\theta} \right] \quad (6.92)$$

where $f(\theta)$ is the purely diffusivity-driven flux-concentration relation, $F(\theta, 0)$, given by Equation 6.76. Obviously β is defined over the interval $0 \leq \beta \leq 1$ and is function of the soil type and governing initial and boundary conditions. The lower limit $\beta = 0$ corresponds to the Green and Ampt type of soil in which $dK/d\theta$ increases much less rapidly with θ than the diffusivity, whereas the upper limit $\beta = 1$ corresponds to soils in which $dK/d\theta$ and the diffusivity behave similarly. Following Equation 6.92, parameter β is defined as function of volumetric water content ($\theta_0 \leq \theta \leq \theta_s$); as a result, β is only influenced by the surface boundary condition when $\theta_1 \leq \theta_s$ or $h_1 \leq 0$. Ross et al. (1996) showed that β is slightly affected by changes in the surface boundary condition θ_1 , especially when θ_1 stays close to θ_s ($\theta_1 \geq 0.75\theta_s$). The effect of the changing initial condition θ_0 is even more straightforward: obviously, β equals 1 for $\theta_0 = 0$ (Equation 6.92); it remains close to 1 for $\theta_0 \leq 0.25\theta_s$ and decreases then substantially at the initial water content $\theta_0 = 0.25\theta_s$ (Ross et al., 1996).

The use of Equation 6.92 allows the solution of the integral term of Equation 6.91 resulting in the following cumulative infiltration equation (Haverkamp et al., 1990):

$$I(q_1, t) = K_0 t + h_{\text{surf}} [\theta_s - \theta_0] \frac{K_s}{q_1 - K_s} + \frac{S_1^2(\theta_s, \theta_0)}{2[K_s - K_0]} \ln \left[1 + \frac{K_s - K_0}{q_1 - K_s} \right] \quad (6.93)$$

However, the cumulative infiltration (I) can only be calculated from Equation 6.93 once the corresponding surface infiltration flux (q_1) is known. The derivation in time of Equation 6.93 yields the necessary flux/time equation, $q_1(t)$:

$$2[K_s - K_0]t = \frac{2h_{\text{surf}} K_s [\theta_s - \theta_0]}{q_1 - K_s} + \frac{S_1^2(\theta_s, \theta_0)}{\beta[1 - \beta][K_s - K_0]} \ln \left[1 + \beta \frac{K_s - K_0}{q_1 - K_s} \right] - \frac{S_1^2(\theta_s, \theta_0) + 2[1 - \beta]h_{\text{surf}} K_s [\theta_s - \theta_0]}{[1 - \beta][K_s - K_0]} \ln \left[1 + \frac{K_s - K_0}{q_1 - K_s} \right] \quad (6.94)$$

The infiltration solution given by the combined use of Equation 6.93 and Equation 6.94 is valid over the entire time range $t \in [0, \infty]$ and is applicable for both positive and negative head surface boundary conditions. Although this solution is proven to be extremely precise, it is also admittedly cumbersome to apply. For that reason Barry et al. (1995) presented an extension of foregoing infiltration solution that improved the applicability of the approach by transforming the implicit combination of Equation 6.93 and Equation 6.94 into an explicit infiltration equation without affecting the precision.

6.5 Scaling Principles

In addition to the obvious lack of theoretical knowledge of the behavior of the flow equation in the unsaturated zone across temporal and spatial scales, the nonlinearity of the flow equations adds also to the complexity of representing spatial variability. There exists an acute disparity between the current theoretical knowledge, characterization, and measurement of water movement in the vadose zone and the characteristic spatial scales for which surface flux models such as SVATs and GCMs are trying to make predictions.

Hence, there is an urgent need to assess the adequacy and usefulness of scale matching between the different flow phenomena involved in the interactions between atmosphere, land surface, vadose zone, and aquifer. The correct specification of soil hydrological processes at small or large scales is directly conditioned by the possibility of parameterizing soil water fluxes for the scales compatible with the grid size of field scales involved. The efficiency of soil characteristic parameterization at different scales depends on the clear definition of the functional relationships and parameters to be measured. The use of actual technologies to characterize surface fluxes gives only local information and provides only an indication of the underlying spatial variability of the different processes. For these multiple reasons, the input parameters used in different modeling approaches bear little resemblance to those measured. Spatial variability can be incorporated by the determination of effective parameters which are supposed to be valid over an REV (e.g., Woods et al., 1995). Some of these methods are based on the use of PTFs that employ textural soil properties as principal input data. As mentioned earlier (Section 6.2.7), textural properties are characterized by a smaller underlying variability than the structural properties. The soil characteristic shape parameters (such as the water retention shape parameter m and the hydraulic conductivity shape parameter η , Equation 6.21 and Equation 6.25), which are strongly related to soil textural properties, could be determined from PTFs. However, the description of water flow in the unsaturated zone requires not just the knowledge of the shape parameters; the information about the soil hydraulic scale parameters (such as the water content scale parameter θ_s , the water pressure scale parameter h_g , and the hydraulic conductivity scale parameter K_s , is even more important. The scale parameters that imbed the dynamical properties of soils cannot be determined just from soil textural properties. Any attempt to determine the soil scale parameters at different spatial grid scales has to incorporate, in one way or another, the flow behavior of the soil at the chosen grid scale. The nonlinearity between soil water fluxes and soil properties, and the extrapolation of theories developed at the local scale to large scales introduce inherent limitations. Hence, methods that simply try to upscale the soil hydraulic characteristic scale parameters and then describe the flow behavior with those aggregated soil characteristics do not take into account these nonlinearities and will generally fail to predict the unsaturated flow behavior in a correct way. As the knowledge of the unsaturated flow behavior of a soil is of major importance for vadose zone hydrology, rather than the knowledge of the hydraulic soil characteristics (which are nothing else than intermediate mathematical functional relationships used to calculate the flow behavior), it would be more obvious to upscale the transfer equation itself.

Basically, two categories of scaling techniques can be distinguished: physically based techniques such as *dimensional* and *inspectional* analyses, and empirically based techniques such as the *functional normalization* technique. A complete discussion of the different methods is given by Langhaar (1951) and Kline (1965). As stated by Tillotson and Nielsen (1984), the former category allows one “to convert a set of physically interrelated dimensional quantities into a set of nondimensional quantities which conserve the original interrelationship for a system that manifests geometric, kinematic or dynamic similarity.” The scaling factors obtained through such dimensional analysis have definite physical meaning in terms of the system being studied. On the other hand, the second category (i.e., functional normalization analyses) derives scaling factors relating properties of the two systems in some empirical way. Most of these methods (e.g., Miller and Miller 1955a, 1955b, 1956; Reichardt et al., 1972; Warrick et al., 1977; Simmons et al., 1979; Warrick and Nielsen, 1980; Nielsen et al., 1983) attempt to combine sets of soil water characteristic curves into one single reference curve on the base of simple geometric similarity and, hence, fail *a priori* to scale the vadose zone flow behavior in a general way. For the sake of generality, it is evident that the first category of dimensional and inspectional analyses is the most desirable and appropriate for tackling the problem of scaling.

6.5.1 Scaling Infiltration

Dynamical analysis rests on the fundamental postulate that the *invariance* of a physical law under a series of scale transformations implies the invariance of all consequences of the law under the same transformations (Birkhoff, 1960). As the solution of the cumulative infiltration problem is simply a consequence of the

flow equation for a particular set of initial and boundary conditions, this postulate guarantees that the scaling factors obtained for the cumulative infiltration law apply equally for the flow equation.

The following set of scale transformations is chosen:

$$\begin{aligned} z &= \alpha_z z^*, \quad t = \alpha_t t^*, \quad \theta = \alpha_\theta \theta^* \\ h &= \alpha_h h^*, \quad K = \alpha_K K^*, \quad C = \alpha_C C^* \end{aligned} \quad (6.95)$$

where α_z , α_t , α_θ , α_h , α_K , and $\alpha_C = \alpha_\theta/\alpha_h$ are constant parameters which define, respectively, the non-dimensional variables z^* , t^* , θ^* , h^* , K^* , and C^* . Substitution of these scaling factors into Equations 6.54 to 6.57 allows us to express the flow equation (e.g., Richards' equation) as a non-dimensional boundary-value problem. The particular values of the scaling factors α should reflect the effect of the initial and boundary conditions on the values of the independent variable.

The first step of inspectional analysis is to select a quasi-exact solution of the infiltration problem. The generalized positive head infiltration solution (Equation 6.93 and Equation 6.94) of Parlange et al. (1985a) is a quasi-exact 1D analytical solution that is valid for any soil type and over the complete time range: $t \in [0, \infty]$. The obvious choice would be to apply the scaling analysis to this infiltration solution (see Haverkamp et al., 1998a). However, for the sake of brevity we choose here to illustrate the principles of the dynamical scaling analysis on the Green and Ampt (1911) solution (Equation 6.85 and Equation 6.86) subject to initial and boundary conditions given by Equation 6.59 and Equation 6.60 with $h_1 = h_{\text{surf}} > 0$. Even though it is perhaps the least realistic infiltration equation among the many different analytical solutions of 1D infiltration presented in the literature, it is still an exact solution of the flow equation for the particular case where soils are considered to behave like Green and Ampt soils. The scaling factors which will be derived for the Green and Ampt equations (Equation 6.85 and Equation 6.86) apply equally well for the generalized infiltration solution of Parlange et al. (1985a) (Haverkamp et al., 1998a).

The second step of inspectional analysis (Sposito and Jury, 1985) consists of trying to define the scaling factors α of Equation 6.95. To do so, the infiltration solution represented by Equation 6.85 and Equation 6.86 is expressed in a dimensionless form through the use of the following dimensionless variables:

$$[I - K_0 t] = \frac{S_1^2(h_{\text{surf}}, \theta_0)}{c_p [K_s - K_0]} \Delta I^* \quad (6.96)$$

and

$$t = \frac{S_1^2(h_{\text{surf}}, \theta_0)}{c_p [K_s - K_0]^2} t^* \quad (6.97)$$

where ΔI^* is the dimensionless net cumulative infiltration; and c_p represents a proportionality constant that depends on the functional relationships chosen to describe the soil hydraulic characteristics. For the case of the Green and Ampt configuration where the soil characteristics are simply defined as functions of the soil structure-dependent scale parameters θ_s , h_f , and K_s , the value of c_p is equal to 2. From Equation 6.96 and Equation 6.97, the dimensionless infiltration flux at the soil surface is defined as:

$$\Delta q_1^* = \frac{d\Delta I^*}{dt^*} = \frac{q_1 - K_0}{K_s - K_0} \quad (6.98)$$

Substitution of Equations 6.96 to 6.98 into Equation 6.85 and Equation 6.86 allows us to express the Green and Ampt infiltration solution in the following dimensionless form:

$$\Delta I^* = t^* + \frac{c_p}{2} \ln \left[1 + \frac{2}{c_p} \Delta I^* \right] \quad (6.99)$$

and

$$\frac{2}{c_p} t^* = \frac{1}{\Delta q_1^* - 1} - \ln \left[1 + \frac{1}{\Delta q_1^* - 1} \right] \quad (6.100)$$

The simultaneous use of Equation 6.99 and Equation 6.100 reduces the infiltration problem to a dimensionless boundary-value problem valid for all soil types (obviously, still in the context of the Green and Ampt concept), initial and boundary conditions. This dimensionless solution allows the straightforward determination of the specific scaling factors for cumulative infiltration (α_I) and infiltration rate (α_q). Choosing the scaling factors α_I and α_q in a similar way to the general scale transformations defined previously for Equation 6.95:

$$[I - K_0 t] = \alpha_I \Delta I^* \quad \text{and} \quad [q_1 - K_0] = \alpha_q \Delta q_1^* \quad (6.101)$$

the scaling factors α_I and α_t can directly be calculated from Equation 6.96 and Equation 6.97:

$$\alpha_I = \frac{S_1^2(h_{\text{surf}}, \theta_0)}{c_p [K_s - K_0]} \quad (6.102)$$

and

$$\alpha_t = \frac{S_1^2(h_{\text{surf}}, \theta_0)}{c_p [K_s - K_0]^2} \quad (6.103)$$

The infiltration rate scaling factor α_q can then be determined from Equation 6.98:

$$\alpha_q = \frac{\alpha_I}{\alpha_t} = [K_s - K_0] \quad (6.104)$$

The key point illustrated by foregoing analysis is that for any Green and Ampt soil the infiltration behavior (cumulative and infiltration flux) is fully determined by only two scaling factors: α_I (or α_q) and α_t . Both parameters embody the effect of soil type, initial and boundary conditions. The dimensionless invariant cumulative infiltration curve ($\Delta I^*(t^*)$) fixes the only *dynamical similarity class* for which the physical system can be said to be macroscopically similar.

So far, the scaling factors α_I and α_q still exhibit the specific character related to the type of flow problem analyzed (cumulative constant head infiltration). To bypass this specificity it is necessary to go down one step in the hierarchy of the water movement problematic by establishing the set of basic scaling factors that reduce the Richards' equation ($\theta(z, t)$) rather than those reduce the cumulative infiltration equation $I(t)$. This can be pursued by decomposing the specific infiltration scaling factors α_I and α_q into the set of scaling factors α_θ , α_h , α_K , and α_z . The determination of the conductivity scaling factor α_K is straightforward:

$$\alpha_K = \alpha_q = [K_s - K_0] \quad (6.105)$$

The introduction of the scaling factors α_t and α_K into the Richards' equation (Equation 6.54) gives the identities:

$$\alpha_h = \alpha_z \quad \text{and} \quad \alpha_I = \alpha_\theta \alpha_h \quad (6.106)$$

By choosing the definition for the scaling factor α_θ as:

$$\alpha_\theta = [\theta_s - \theta_0] \quad (6.107)$$

the pressure scaling factor (α_h) and thus the space scaling factor (α_z) are defined by:

$$\alpha_h = \alpha_z = \frac{S_1^2(h_{\text{surf}}, \theta_0)}{c_p[\theta_s - \theta_0][K_s - K_0]} \quad (6.108)$$

As the time and space scaling factors (i.e., α_t and α_z) are fully determined by the soil scaling factors α_θ , α_h , and α_K , the soil water transfer equation can be transformed simply into a non-dimensional form by the use of only three scaling factors; these factors take into account the effect of soil type, initial and boundary conditions giving the dimensionless form of the Richards' equation (Equation 6.54) the flexibility to encompass the water flow behavior of diverse physical systems through the adjustment of experimentally controllable parameters.

To make this inspectional analysis of the water transfer and infiltration equation useful for practical application, the soil scale factors should be expressed in terms of the classically used soil characteristic parameters contained in the expressions of $h(\theta)$ and $K(\theta)$ such as Equation 6.21 and Equation 6.25. To do so, the scaling factors α_θ , α_h , and α_K have to be stripped of initial and boundary condition effects. Letting λ stand for the fraction of α , which is purely related to the soil characteristic, the different soil scaling factors can be redefined as functions of the governing surface boundary condition ($h_1 = h_{\text{surf}}$):

$$\alpha_\theta = \lambda_\theta \left\{ \frac{\theta_s - \theta_0}{\theta_s} \right\} \quad \text{with } \lambda_\theta = \theta_s \quad (6.109)$$

$$\alpha_K = \lambda_K \left\{ \frac{K_s - K_0}{K_s} \right\} \quad \text{with } \lambda_K = K_s \quad (6.110)$$

and

$$\alpha_h = \lambda_h \left\{ \frac{S_1^2(h_{\text{surf}}, 0)}{S_1^2(\theta_s, 0)} \right\} \left\{ \frac{\theta_s}{\theta_s - \theta_0} \frac{K_s}{K_s - K_0} \frac{S_1^2(h_{\text{surf}}, \theta_0)}{S_1^2(h_{\text{surf}}, 0)} \right\} \quad (6.111)$$

where

$$\lambda_h = \frac{S_1^2(\theta_s, 0)}{c_p \theta_s K_s} \quad (6.112)$$

For the definitions given above, the residual water content θ_r is chosen equal to zero (Kool et al., 1987). For the case of positive head infiltration, the effect of the surface boundary condition enters only through the definition of the pressure scaling factor α_h . The second and third factors on the right-hand side of Equation 6.111 represent, respectively, the correction for the surface boundary condition effects and the initial condition effects.

Although the different sorptivity ratios which enter in the scaling factor α_h (Equation 6.111) are well-defined by Equation 6.77 and Equation 6.83, it is preferable to simplify the definitions of α_h to a maximum by expressing the sorptivity ratios in terms of the scaling factors α_K and α_θ . A precise estimation of the sorptivity ratio $S_1^2(\theta_s, \theta_0)/S_1^2(\theta_s, 0)$ is given by:

$$\frac{S_1^2(\theta_s, \theta_0)}{S_1^2(\theta_s, 0)} \cong \frac{\theta_s - \theta_0}{\theta_s} \frac{K_s - K_0}{K_s} \quad (6.113)$$

Using Equation 6.83 and Equation 6.113, the expression for the positive head pressure scale parameter (α_h) can be simplified considerably:

$$\alpha_h \cong \lambda_h + \frac{2}{c_p} \frac{\lambda_K}{\alpha_K} h_{\text{surf}} \quad (6.114)$$

The foregoing equations (Equation 6.109 to Equation 6.114) are valid not only for the Green and Ampt (1911) infiltration solution (Equation 6.85 and Equation 6.86), but for the generalized analytical infiltration solution of Parlange et al. (1985a), (Equation 6.93 and Equation 6.94) as well. For the particular case of Green and Ampt soils, the scaling factors α_h and λ_h take the form:

$$\alpha_h = \lambda_h + \frac{K_s}{K_s - K_0} h_{\text{surf}} \quad \text{with } \lambda_h = -h_f \quad (6.115)$$

where h_f is taken as a negative pressure head value.

The essence of this example on the Green and Ampt infiltration equation is that there exists a unique dynamical similarity in the behavior of unsaturated soil water movement in Green and Ampt soils when governed by the Richards' equation. The infiltration behavior is defined by an invariant curve. However, when the scaling analysis is applied on the generalized constant head infiltration solution (Equation 6.93 and Equation 6.94) there exists, theoretically, no unique dynamical similarity in the soil water behavior (Haverkamp et al., 1998a). Instead there is a multitude of dynamical similarity classes depending on the combination of soil type, initial and boundary conditions. In its most general form, the positive or negative head infiltration behavior is defined by three infiltration scaling factors embodying the effect of soil type, initial and boundary conditions. However, the non-uniqueness of the different similarity classes was found to be very small implying that for practical purposes the hypothesis of unique dynamical similarity can be applied once the Green and Ampt equation is adapted to more general field conditions.

The purely soil-related scaling factors λ_θ , λ_h , and λ_K represent the soil characteristic scale parameters used by the soil hydraulic characteristic curves such as Equations 6.20, 6.21, 6.25, and 6.26. Considering the water retention curve of Brooks and Corey (1964) substitution of λ_θ and λ_h into Equation 6.20 gives:

$$\begin{cases} \frac{\theta}{\lambda_\theta} = \left[\frac{h}{\lambda_h} \right]^{-\lambda} & \text{for } h \leq \lambda_h \\ \theta = \lambda_\theta & \text{for } \lambda_h \leq h \leq 0 \end{cases} \quad (6.116)$$

and for the van Genuchten (1980) water retention curve:

$$\frac{\theta}{\lambda_\theta} = \left[1 + \left(\frac{h}{\lambda_h} \right)^n \right]^{-m} \quad (6.117)$$

Substitution of λ_θ and λ_K in the hydraulic conductivity equation (Equation 6.25) of Brooks and Corey (1964) yields:

$$\frac{K}{\lambda_K} = \left[\frac{\theta}{\lambda_\theta} \right]^\eta \quad (6.118)$$

When describing unsaturated water flow subject to a given set of initial and boundary conditions, the water flow behavior of the soil should be independent of the choice of the soil hydraulic functional relationships. This can be guaranteed by applying the sorptivity criterion that imposes the sorptivity, $S_1(\theta_s, 0)$ (Equation 6.77) calculated by the use of the redefined water retention and hydraulic conductivity equations such as given by Equation 6.116 to Equation 6.118, and the sorptivity, $S_1(\theta_s, 0)$, expressed as function of λ_h (Equation 6.112) to be equal. Establishing this equality yields the correct c_p value which is specific for the functional relationship of $h(\theta)$ chosen and guarantees the sorptivity criterion.

The values of c_p calculated for the Brooks and Corey (1964) and van Genuchten (1980) water retention equations are respectively:

$$c_p = \frac{2\lambda\eta[\lambda\eta - 1] + \lambda[2\lambda\eta - 1]}{[\lambda\eta - 1][\lambda\eta - 1 + \lambda]} \quad (6.119)$$

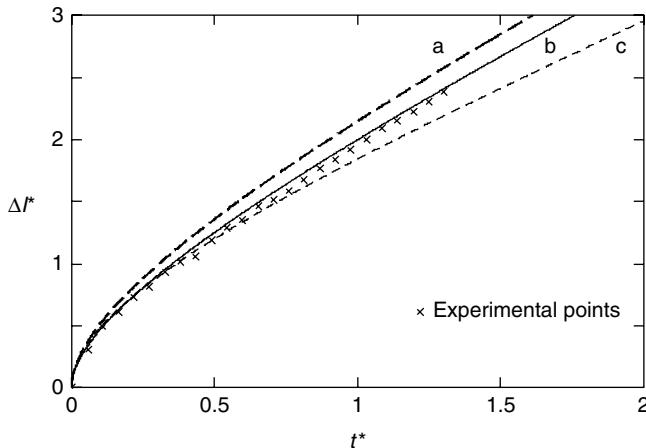


FIGURE 6.14 Experimental points (Touma et al., 1984) and predicted dimensionless cumulative infiltration curves: (a) the classical Green and Ampt (1911) equation, (b) the modified Green and Ampt equation, and (c) the Talsma and Parlange (1972) equation.

and

$$c_p = \Gamma \left(1 + \frac{1}{n} \right) \left\{ \frac{\Gamma(m\eta - 1/n)}{\Gamma(m\eta)} + \frac{\Gamma(m\eta + m - 1/n)}{\Gamma(m\eta + m)} \right\} \quad (6.120)$$

where Γ stands for the usual gamma function. Obviously, for Green and Ampt soils with the sorptivity given by Equation 6.88, the particular value of c_p is equal to 2. When dealing with general field soils, c_p is purely defined by textural soil parameters (Equation 6.119 and Equation 6.120). The values of m and η or λ and η can be determined by the use of PTFs (see Section 6.2.6) allowing the values of c_p to be easily calculated.

The use of c_p together with the Green and Ampt equation (Equation 6.99) has the advantage of being an invariant dimensionless infiltration curve valid for general field soils without being limited by the hypothesis of the Green and Ampt configuration. The fact that the proportionality constant c_p is pinned down on the sorptivity expression that is valid for realistic soil hydraulic characteristics relations (such as the van Genuchten and Brooks and Corey curves), guarantees a realistic description of the infiltration flow behavior as textural soil properties are taken into account. This result is illustrated by the dimensionless cumulative infiltration curves ($\Delta I^*(t^*)$) shown in Figure 6.14. The envelope of possible similarity classes bounded by the two invariant solutions of the classical Green and Ampt equation ($c_p = 2$) and the Talsma and Parlange (1972) equation (valid for the Gardner [1958] soil) are presented, together with the experimental constant head infiltration data given by Touma et al. (1984). The Green and Ampt equation (Equation 6.99) with c_p calculated from Equation 6.120 gives a good agreement with the measured infiltration data.

The experimental cumulative infiltration data together with the surface boundary condition (h_{surf}) are the only data supposed to be known from field measurements. The identification procedure aims to determine the scaling factors α_I and α_t from these input data. Generally speaking, the optimization of two parameters on a monotone increasing function (such as the cumulative infiltration curve), guarantees *a priori* the uniqueness of the solution. Once the values of α_I and α_t are determined, the purely soil-related scaling parameters λ_θ , λ_h , and λ_K can be calculated (details of the identification procedure are given by Haverkamp et al., 1998a). Some of the results are shown in Figure 6.15. In the context of the European international project on Climate and Hydrological Interactions between the Vegetation, the Atmosphere, and the Land surface (ECHIVAL) an infiltration field experiment was carried out in the desertification-threatened area of La Mancha (Spain). One hundred cumulative infiltration curves were measured over

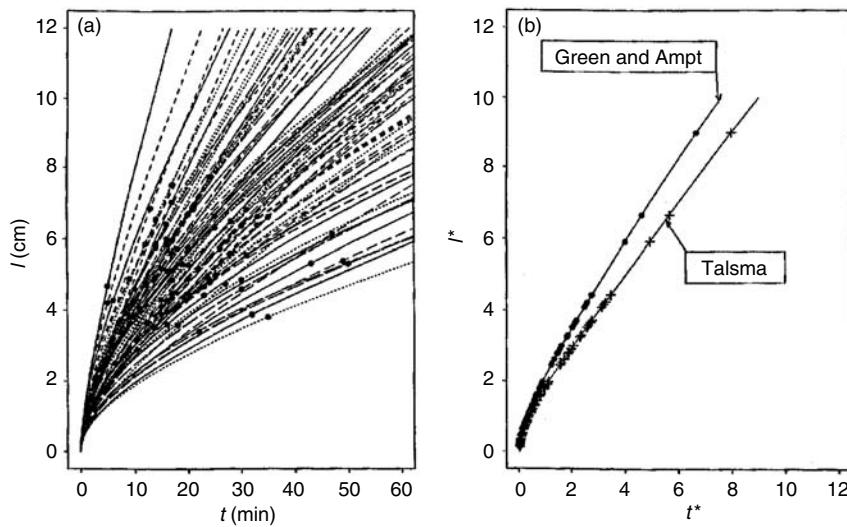


FIGURE 6.15 Fan of 100 different cumulative infiltration curves measured over an area of 100 km² before scaling (a) and after scaling (b). (Adapted from Haverkamp, R., Parlange, J.-Y., Cuenca, R., Ross, P. J., and Steenhuis, T. S. 1998a. Scaling of the Richards' equation and its application to watershed modeling. In *Scale Invariance and Scale Dependence in Hydrology*, Sposito, G. [Ed.], Cambridge University Press, New York, pp. 190–223. With permission.)

an area of 100 km² with a grid mesh of 1 × 1 km. The results are given in Figure 6.15a. For each of the 100 curves the parameters α_1 and α_t were determined by the use of the Green and Ampt equation (Equation 6.99) with $c_p = 2$ and the Talsma and Parlange (1972) solution. The scaled results, $I^*(t^*)$ are given in Figure 6.15b showing the unique similarity which exists for both infiltration curves. All the 100 points fit each of the two infiltration curves perfectly well. It showed that with a minimum of measurement points for $I(t)$ precise information could be obtained on local soil hydraulic characteristics and their possible variations (upper and lower limits). Large areas were surveyed in short times. For example, the entire cumulative infiltration measurement campaign over the area of 100 km² did not take more than three weeks.

The foregoing scaling concept is valid for any spatial grid scale provided that suitable scaling factors can be measured and the soil is considered equivalent homogeneous for the grid scale involved. The concept can be used for any flow equation. Although the details of the approach were discussed through the infiltration law as flow equation, there are inherent limitations for using the infiltration phenomenon. The characteristic grid size of infiltration is still rather small (in the order of 1 m²). When dealing with hydrologic model studies the ideal choice would be to use the evaporation flux as flow phenomenon as its characteristic grid size is much larger (in the order of 1 km²).

6.5.2 Scaling Evaporation

An identical approach to that developed previously can be applied for evaporation using Salvucci's (1997) approximate exfiltration equation based on moisture profile similarity:

$$[K(\theta_0)]^2 t = \frac{S_d(\theta_1, \theta_0)}{2} \left[\frac{K(\theta_0)}{f_e} - \ln \left(1 + \frac{K(\theta_0)}{f_e} \right) \right] \quad (6.121)$$

subject to initial and boundary conditions:

$$\begin{aligned} \theta &= \theta_0 & \text{at } t = 0 & \text{and } z \geq 0 \\ \theta &= \theta_1 & \text{at } z = 0 & \text{and } t > 0 \end{aligned} \quad (6.122)$$

where f_e is the exfiltration rate; and the *desorptivity* $S_d(\theta_1, \theta_0)$ is estimated by analogy with sorptivity (Parlange et al., 1985b) from:

$$S_d(\theta_1, \theta_0) = \frac{8}{3} \int_{h_0}^{h_1} [\theta - \theta_0] K dh \quad (6.123)$$

Note that for evaporation $\theta_1 < \theta_0$. Even though Equation 6.121 does not assume a sharp water content front in the drying soil, it is similar to the Green and Ampt infiltration (1911) equation. Considering, for the sake of brevity, $K_0 = 0$ and $c_p = 2$, the following dimensionless variables can be defined:

$$f_e^* = \frac{K(\theta_0)}{f_e^*} \quad \text{and} \quad t = \frac{S_d^2(\theta_1, \theta_0)}{2[K(\theta_0)]^2} t^* \quad (6.124)$$

Equation 6.121 can then be written in the dimensionless form:

$$f_e^* = t^* + \ln[1 + f_e^*] \quad (6.125)$$

As Equation 6.125 is identical to the Green and Ampt equation (Equation 6.99) with $c_p = 2$, the scaling analysis explained previously can be applied to determine the invariant exfiltration curve. This approach is extremely interesting as there are actually methods, such as remote sensing, that provide areal averages of evaporation. It would then be possible from cumulative exfiltration (by analogy with cumulative infiltration) to evaluate the exfiltration scaling factors. Once the desorptivity is converted to sorptivity, the areal soil hydraulic scale parameters consistent with the flow process can be determined.

As new technologies such as Geographic Information Systems (GIS) and remote sensing are increasingly becoming standard tools for evaluating flow processes at various scales, this scaling methodology proposes compatibility rules for combining different parameters and data sets, which is an aspect that is often neglected in GIS operations (Band and Moore, 1995).

6.6 Conclusions and Perspectives on Future Research

This chapter is mainly concerned with the water transfer in the unsaturated or vadose zone. The conceptual aspects of unsaturated soil water flow related to soil physical properties and soil water characteristics have been presented. Different functional relationships to represent soil hydraulic characteristics and different measurement and estimation techniques have been addressed. An overview of the most significant analytical solutions and the most important physical infiltration laws are given. It has been shown that the textural soil properties could give information regarding the shape parameters of the soil hydraulic characteristics, and the scale parameters that imbed the dynamic characteristics of the soil could only be determined if a flow equation is somehow considered. In this perspective, a scaling approach of the water transfer equation in the unsaturated zone has been given.

This new theory is based on the scaling of the unsaturated flow equation rather than on the scaling of the soil hydraulic characteristics. Following the postulate of invariance of a physical law and its consequences under a series of transformations, it is shown that in general there exists no unique dynamical similarity in the behavior of unsaturated soil water movement. A unique dynamical similarity is only found for two particular soil descriptions: the Green and Ampt (1911) soil and the Gardner (1958) soil. Both limit the envelope of possible similarity classes that exist for general field soils. However, when modifying the Green and Ampt configuration by taking into account the effects of soil texture, an invariant infiltration law can be determined that is realistic for general field soils.

This approach offers a real alternative to classic estimation techniques of soil hydraulic characteristics, which are virtually impossible to perform over a great number of grid locations over a catchment, with a

realistic use of resources in terms of time and costs. It is possible through quick infiltration experiments to determine the dynamic characteristics of soils that are far more representative and accurate than any estimate given from textural information such as PTFs. Furthermore, the effects of preferential flow can be taken into account. The second major advantage of the scaling approach is that it is compatible with the scale of measurement. As this methodology can be applied to any flow equation such as evaporation, it is possible through the use of remote sensing, which provide realistic aggregated values of water fluxes, to tackle the mismatch that exists between the scale on which current knowledge of vadose zone water transfer processes is based and the scale for which the water budget models are making predictions.

Glossary and Parameter Notation

a	stone weight fraction
b	porous rock weight fraction
A_1	first term of Philip's truncated infiltration equation ($\text{m/sec}^{1/2}$)
A_2	second term of Philip's truncated infiltration equation (m/sec)
$c_{1,\dots,n}$	regression coefficients
c	light velocity (m/sec)
c_p	proportionality constant
C	specific capacity ($1/\text{m}$)
C_e	dielectric constant
Cl	clay fraction (%)
dVp_i	pore volume associated with each size class ' i ' of the Arya and Paris (1981) model (m^3)
D	diffusivity (m^2/sec)
Dpg	particle-size scale parameter (m)
Dp	effective diameter of soil particle (m)
Dp_i	particle radius of size class ' i ' of the Arya and Paris (1981) model (m)
e	void ratio (m^3/m^3)
f_p	particle-size distribution frequency expressed as a percentage of particle weight per unit sample mass
f_e	exfiltration rate (m/sec)
f_e^*	dimensionless exfiltration rate
F	flux-concentration relation
F_{PA}	flux-concentration relation defined by Parlange (1971a)
F_p	cumulative distribution of particle weight
g	gravitational acceleration, 9.82 m/sec^2
h	soil water pressure head (m)
h_0	soil water pressure head at initial water content (m)
h_1	surface soil water pressure head (m)
h_{bc}	Brooks and Corey (1964) soil water retention scale parameter (m)
h_g	van Genuchten (1980) soil water retention scale parameter (m)
h_f	Green and Ampt wetting front pressure head value (m)
h_{surf}	ponded water depth at soil surface (m)
h^*	dimensionless soil water pressure head
H	hydraulic head (m)
I	one-dimensional cumulative infiltration (m^3/m^2)
k	intrinsic permeability (m^2)
k_m	parameter relating van Genuchten (1980) shape parameters m and n
k_M	parameter relating particle-size shape parameters M and N
k_{rw}	relative water permeability
K	hydraulic conductivity (m/sec)
K_s	hydraulic conductivity at natural saturation (m/sec)

K_0	hydraulic conductivity at initial water content θ_0 (m/sec)
K_1	hydraulic conductivity at water content θ_1 (m/sec)
K^*	dimensionless hydraulic conductivity
l	rod length of TDR probe (m)
m	van Genuchten (1980) water retention shape parameter
M	particle-size shape parameter
M_g	mass of gas phase of REV (kg)
M_l	mass of liquid phase of REV (kg)
M_{mol}	molecular weight
M_s	mass of solid phase of REV (kg)
M_t	total mass of REV (kg)
n	van Genuchten (1980) water retention shape parameter
N	particle-size shape parameter
nb_i	number of particles in size class ' i ' of the Arya and Paris (1981) model
OC	percentage of organic carbon content
p	water vapor pressure
p_0	saturated vapor pressure
$p\beta$	tortuosity factor
pF	logarithmic mode of pressure head
P_{wr}	water retention shape index
P_{pa}	particle-size shape index
PTF	pedo-transfer function
q	soil water flux (m/sec)
q_1	one-dimensional infiltration flux at the soil surface (m/sec)
r	radius pore opening (m)
R	radius of the main pore (m)
R_i	pore radius of size class ' i ' of the Arya and Paris (1981) model (m)
R_g	universal gas constant
REV	representative elementary soil volume (m^3)
s	fractal dimension
S	degree of saturation (m^3/m^3)
Sa	sand fraction (%)
S_d	desorptivity ($\text{m/sec}^{1/2}$)
Si	silt fraction (%)
S_l	sorptivity ($\text{m/sec}^{1/2}$)
S_2	second integral term of Philip's analytical infiltration equation
t	time (sec)
t_{tr}	travel time of electromagnetic TDR wave (sec)
t^*	dimensionless time
T_l	temperature liquid phase (K)
V_g	volume gas phase of REV (m^3)
V_l	volume liquid phase of REV (m^3)
V_s	volume solid phase of REV (m^3)
V_t	total volume of REV (m^3)
w	gravimetric soil water content (kg/kg)
W_g	weight of gas phase of REV (kg)
W_l	weight of liquid phase of REV (kg)
W_s	weight of solid phase of REV (kg)
W_t	total weight of REV (kg)
x	horizontal space coordinate (m)
z	depth, positive downward (m)

z_m	reference depth of instantaneous profile method (m)
z_0	zero flux plan (m)
z^*	dimensionless depth
z_f	depth of infiltration wetting front (m)
α	tortuosity factor ($\alpha = 1.38$) of the Arya and Paris (1981) model
α_C	specific capacity scaling factor (1/m)
α_h	pressure head scaling factor (m)
α_I	cumulative infiltration scaling factor (m)
α_K	hydraulic conductivity scaling factor (m/sec)
α_q	flux scaling factor (m/sec)
α_t	time scaling factor (sec)
α_z	depth scaling factor (m)
α_θ	water content scaling factor (m^3/m^3)
β	Parlange et al. (1982) conductivity shape parameter
γ	packing coefficient of the Haverkamp and Parlange (1986) model
γ_s	tortuosity factor defined by the water retention shape index
ΔS	change in water storage (m)
Δt	time step (sec)
ΔI^*	dimensionless net cumulative infiltration
Δq_I^*	dimensionless net infiltration flux
ε	soil porosity (m^3/m^3)
ε_{ap}	whole soil porosity (m^3/m^3)
ε_c	porosity of porous coarse fragments (m^3/m^3)
λ	Brooks and Corey (1964) shape parameter
λ_h	purely soil-related pressure head scaling factor (m)
λ_K	purely soil-related hydraulic conductivity scaling factor (m/sec)
λ_θ	purely soil-related water content scaling factor (m^3/m^3)
η	Brooks and Corey (1964) hydraulic conductivity shape parameter
θ	volumetric soil water content (m^3/m^3)
θ_r	residual volumetric soil water content (m^3/m^3)
θ_s	volumetric soil water content at natural saturation (m^3/m^3)
θ_0	initial volumetric soil water content (m^3/m^3)
θ_1	surface volumetric soil water content (m^3/m^3)
θ^*	dimensionless volumetric soil water content: θ/θ_s
μ_w	water viscosity (kg/m/sec)
ρ_{ap}	whole soil dry bulk density (kg/m ³)
ρ_c	specific weight of porous coarse fragments (kg/m ³)
ρ_d	dry bulk density (kg/m ³)
ρ_s	specific weight of soil particles: 2.65 (g/cm ³)
ρ_w	soil water specific weight (kg/m ³)
σ_{LV}	liquid-vapor surface tension of water (kg/m/sec ²)
ϕ_a	air potential (joule/m ³)
ϕ_c	capillary potential (joule/m ³)
ϕ_e	envelope potential (joule/m ³)
ϕ_g	gravitational potential (joule/m ³)
ϕ_o	osmotic potential (joule/m ³)
Φ	total soil water potential (joule/m ³)
χ_1	Boltzmann similarity variable (m/sec ^{1/2})
χ_2	second term of Philip's time series solution (m/sec)
χ_3	third term of Philip's time series solution (m/sec ^{3/2})
ω	contact angle between liquid and solid phase (rad)

Acknowledgments

This work was partially supported by the French National Research Council (CNRS) and the IRD (Institut de Recherche pour le Développement) at Benin.

References

- Ahuja, L. R., Naney, J. W., and Williams, R. D. 1985. Estimating soil water characteristics from simpler properties or limited data. *Soil Sci. Soc. Am. J.*, 49: 1100–1105.
- Arya, L. M. and Paris, J. F. 1981. A physico-empirical model to predict the soil moisture characteristic from particle-size distribution and bulk density data. *Soil Sci. Soc. Am. J.*, 45: 1023–1030.
- Arya, L. M. and Paris, J. F. 1982. Reply to “Comments on a physico-empirical model to predict the soil moisture characteristic from particle-size distribution and bulk density data.” *Soil Sci. Soc. Am. J.*, 46: 1348–1349.
- Arya, L. M., Leij, F. J., van Genuchten, M. Th., and Shouse, P. J. 1999. Scaling parameter to predict the soil water characteristic from particle-size distribution data. *Soil Sci. Soc. Am. J.*, 63: 510–519.
- Babu, D. K. 1976. Infiltration analysis and perturbation methods, 1. Absorption with exponential diffusivity. *Water Resour. Res.*, 12: 89–93.
- Band, L. E. and Moore, I. D. 1995. Scale: landscape attributes and geographical information systems. *Hydrol. Process.*, 9: 401–422.
- Barry, D. A., Parlange, J.-Y., Haverkamp, R., and Ross, P. J. 1995. Infiltration under ponded conditions: 4. An explicit predictive infiltration formula. *Soil Sci.*, 160: 8–17.
- Bear, J. and Bachmat, Y. 1984. Transport phenomena in porous media — basic equations. In *Fundamentals of Transport in Porous Media*, Bear, J. and Corapcioglu, M. Y. (Eds.), Martinus Nijhoff, pp. 3–61.
- Birkhoff, G. 1960. *Hydrodynamics*, 2nd ed. Princeton University Press. Princeton, NJ.
- Bloemen, G. W. 1980. Calculation of hydraulic conductivities of soils from texture and organic matter content. *Z. Pflanzenernährung und Bodenkunde*, 143: 581–615.
- Bolt, G. H. 1976. Soil Physics Terminology. *Int. Soc. Soil Sci., Bull.* 49: 16–22.
- Boltzmann, L. E. 1894. Zur Integration der Diffusionsgleichung bei variablen Diffusions-coefficienten. *Ann. Physik (Leipzig)*, 53: 959–964.
- Bouma, J. and van Lanen, J. A. J. 1987. Transfer functions and threshold values; from soil characteristics to land qualities. In *Quantified Land Evaluation*, Beek K. J. et al. (Eds.), International Institute Aerospace Survey Earth Science, ITC Publishing, Enschede, The Netherlands, pp. 106–110.
- Bouwer, H. 1964. Unsaturated flow in groundwater hydraulics. *J. Hydraul. Div. Am. Soc. Civ. Eng.*, 90HY5: 121–127.
- Bouwer, H. 1978. *Groundwater Hydrology*. McGraw-Hill Book Company, New York, NY.
- Braud, I., Haverkamp, R., Arrué, J. L., and López, M. V. 2003. Spatial variability of soil surface properties and consequences for the annual and monthly water balance of a semiarid environment (EFEDA Experiment), *J. Hydrometeorol.*, 4: 121–137.
- Braud, I., de Condappa, D., Soria-Ugalde, J. M., Haverkamp, R., Angulo-Jaramillo, R., Galle, S., and Vauclin, M. 2005. Used of scaled forms of the infiltration equation of unsaturated soil hydraulic properties (the *Beerkan method*), *Eur. J. Soil Sci.*, 56: 361–374.
- Broadbridge, P. 1990. Solution of a nonlinear absorption model of mixed saturated–unsaturated flow. *Water Resour. Res.*, 26: 2435–2444.
- Brooks, R. H. and Corey, C. T. 1964. *Hydraulic Properties of Porous Media*. Hydrol. Paper 3. Colorado State University, Fort Collins.
- Bruce, R. R. and Klute, A. 1956. The measurement of soil water diffusivity. *Soil Sci. Soc. Am. Proc.*, 20: 458–462.
- Brustkern, R. L. and Morel-Seytoux, H. J. 1970. Analytical treatment of two-phase infiltration. *J. Hydraul. Div. ASCE*, 96: 2535–2548.
- Brutsaert, W. 1966. Probability laws for pore-size distributions. *Soil Sci.*, 101: 85–92.

- Buckingham, E. 1907. Studies on the movement of soil moisture. *U.S. Dept. Agr. Bur. Soils Bull.*, 38.
- Burdine, N. T. 1953. Relative permeability calculations from pore-size distribution data. *Petr. Trans. Am. Inst. Mining Metall. Eng.*, 198: 71–77.
- Campbell, G. S. and Gee, G. W. 1986. Water potential: miscellaneous methods. In *Methods of Soil Analysis, Part 1, Physical and Mineralogical Methods*, Klute, A. (Ed.), Monograph 9, 2nd ed. Soil Science Society of America Book Series, *Soil Sci. So. Am. Inc.* Madison, WI.
- Cassel, D. K. and Klute, A. 1986. Water potential: tensiometry. In *Methods of Soil Analysis, Part 1, Physical and Mineralogical Methods*, Klute, A. (Ed.), Agronomy Monograph 9, 2nd ed. Madison, WI.
- Chauvel, A., Grimaldi, M., and Tessier, D. 1991. Changes in pore space distribution following deforestation and revegetation: an example from the Central Amazon Basin, Brazil. *For. Ecol. Manag.*, 38: 259–271.
- Childs, E. C. and Collis-George, C. 1950. The permeability of porous materials. *Proc. R. Soc. Lond. Ser. 201 A*, 392–405.
- Clapp, R. B. and Hornberger, G. M. 1978. Empirical equations for some hydraulic properties. *Water Resour. Res.*, 14: 601–604.
- Cosby, B. J., Hornberger, G. M., Clapp, R. B., and Ginn, T. R. 1984. A statistical exploration of the relationship of soil moisture characteristics to the physical properties of soils. *Water Resour. Res.*, 20: 682–690.
- Crank, J. and Henry, M. E. 1949. Diffusion in media with variable properties, II. *Trans. Faraday Soc.*, 45: 1119–1128.
- Darcy, H. 1856. *Les fontaines publiques de la ville de Dijon*. Dalmont, Paris.
- de Condappa, D., Lachassagne, P., Debionne, S., Viallet, P., and Haverkamp, R. 2005. The bimodal area of the soil textural triangle: specific for tropical and subtropical soils. General assembly of the European Geophysical Society (XXX Congress), Vienna, Austria. EGU05-A-04619; HS24-1TH4O-004 (April 24–29).
- Durner, W. 1994. Hydraulic conductivity estimation for soils with heterogeneous pore structure. *Water Resour. Res.*, 30: 211–223.
- Echings, S. O. and Hopmans, J. W. 1993. Optimization of hydraulic functions from transient outflow and soil water pressure data. *Soil Sci. Soc. Am. J.*, 57: 1167–1175.
- Elmaloglou, S. 1980. Effets des stratifications sur les transferts de matières dans les sols, *Thèse Docteur-Ingénieur*, Université Scientifique et Médicale de Grenoble, France.
- Elrick, D. E. and Robin, M. J. 1981. Estimating sorptivity of soils. *Soil Sci.*, 132: 127–133.
- Elrick, D. E., Parkin, G. W., Reynolds, W. D., and Fallow, D. J. 1995. Analysis of early time and steady state single-ring infiltration under falling head conditions. *Water Resour. Res.*, 31: 1883–1895.
- Fuentes, C., Haverkamp, R., and Parlange, J.-Y. 1992. Parameter constraints on closed-form soil water relationships. *J. Hydrol.*, 134: 117–142.
- Fuentes, C., Vauclin, M., Parlange, J.-Y., and Haverkamp, R. 1998. Soil water conductivity of a fractal soil. In *Fractals in Soil Scienc*, Baveye, P. et al. (Eds.), CRC Press, Boca Raton, FL, pp. 333–340.
- Galle, S., Angulo-Jaramillo, R., Braud, I., Boubkraoui, S., Bouchez, J.-M., de Condappa, D., Derive, D., Gohoungossou, A., Haverkamp, R., Reggiani, P., and Soria-Ugalde, J. M. 2001. *Estimation of soil hydraulic properties of the Donga watershed (CATCH-Bénin)*. Proceedings of the GEWEX 4th International Conference, Paris, 10–14 September, Institut Pierre Simon Laplace, p. 136.
- Gardner, W. R. 1958. Some steady state solutions of the unsaturated moisture flow equation with application to evaporation from a water table. *Soil Sci.*, 85: 228–232.
- Gardner, W. H. 1986. Water content. In *Methods of Soil Analysis, Part 1, Physical and Mineralogical Methods*, Klute, A. (Ed.), Monograph 9, 2nd ed. Soil Science Society of America Book Series, *Soil Sci. So. Am. Inc.* Madison, WI.
- Gérard-Marchant, P., Angulo-Jaramillo, R., Haverkamp, R., Vauclin, M., Groenevelt, P., and Elrick, D. E. 1997. Estimating the hydraulic conductivity of slowly permeable and swelling materials from single ring experiments. *Water Resour. Res.*, 33: 1375–1382.
- Green, W. H. and Ampt, G. A. 1911. Studies in soil physic: I. The flow of air and water through soils. *J. Agric. Sci.*, 4: 1–24.

- Groenevelt, P. H. and Parlange, J.-Y. 1974. Thermodynamic stability of swelling soils. *Soil Sci.*, 118: 1–5.
- Gupta, S. C. and Larson, W. E. 1979. Estimating soil water retention characteristics from particle-size distribution, organic matter percent and bulk density. *Water Resour. Res.*, 15: 1633–1635.
- Gupta, S. C. and Ewing, R. P. 1992. Modeling water retention characteristics and surface roughness of tilled soils. In *Indirect Methods for Estimating the Hydraulic Properties of Unsaturated Soils*, van Genuchten, M. Th., Leij, F. J., and Lund, L. J. (Eds.), Dept. of Environmental Sciences, University of California, Riverside, CA, pp. 379–388.
- Hamon, G. 1980. Mise en oeuvre et critique de méthodes de caractérisation hydrodynamique de la zone non-saturée du sol. Applications aux sols de culture du Sénégal. *Thèse Docteur-Ingénieur*, Université Scientifique et Médicale de Grenoble, France.
- Haverkamp, R. and Vauclin, M. 1979. A note on estimating finite difference interblock hydraulic conductivity values for transient unsaturated flow problems. *Water Resour. Res.*, 15: 181–187.
- Haverkamp, R. and Vauclin, M. 1981. A comparative study of three forms of the Richards' equation used for predicting one-dimensional infiltration in unsaturated soils. *Soil Sci. Soc. Am. J.*, 45: 13–20.
- Haverkamp, R., Vauclin, M., and Vachaud, G. 1984. Error analysis in estimating soil water content from neutron probe measurements; I. Local standpoint. *Soil Sci.*, 137: 78–90.
- Haverkamp, R. and Parlange, J.-Y. 1986. Predicting the water retention curve from particle-size distribution: I. Sandy soils without organic matter. *Soil Sci.*, 142: 325–339.
- Haverkamp, R., Kutilek, M., Parlange, J.-Y., Rendon, L., and Krejca, M. 1988. Infiltration under ponded conditions: 2. Infiltration equations tested for parameter time dependence and predictive use. *Soil Sci.*, 145: 317–329.
- Haverkamp, R., Parlange, J.-Y., Starr, J. L., Schmitz, G., and Fuentes, C. 1990. Infiltration under ponded conditions: 3. A predictive equation based on physical parameters. *Soil Sci.*, 149: 292–300.
- Haverkamp, R., Ross, P. J., Smettem, K. R. J., and Parlange, J.-Y. 1994. Three-dimensional analysis of infiltration from the disc infiltrometer: 2. Physically based infiltration equation. *Water Resour. Res.*, 30: 2931–2935.
- Haverkamp, R., Arrúe, J. L., Vandervaere, J.-P., Braud, I., Boulet, G., Laurent, J.-P., Taha, A., Ross, P. J., and Angulo-Jaramillo, R. 1996. Hydrological and thermal behavior of the vadose zone in the area of Barrax and Tomelloso (Spain): experimental study, analysis and modeling. Project UE n° EV5C-CT 92 00 90.
- Haverkamp, R., Zammit, C., Bouraoui, F., Rajkai, K., Arrúe, J. L., and Heckmann, N. 1997. GRIZZLY, Grenoble Soil Catalogue: Soil survey of field data and description of particle-size, soil water retention and hydraulic conductivity functions. Laboratoire d'Etude des Transferts en Hydrologie et Environnement (LTHE), Grenoble Cedex 9, France.
- Haverkamp, R., Parlange, J.-Y., Cuenca, R., Ross, P. J., and Steenhuis, T. S. 1998a. Scaling of the Richards' equation and its application to watershed modeling. In *Scale Invariance and Scale Dependence in Hydrology*, Sposito, G. (Ed.), Cambridge University Press, New York, pp. 190–223.
- Haverkamp, R., Bouraoui, F., Zammit, C., and Angulo-Jaramillo, R. 1998b. Movement of moisture in the unsaturated zone. In *Groundwater Engineering Handbook*, Delleur, J. W. (Ed.), CRC Press LLC, Boca Raton, FL, pp. 5.1–5.50.
- Haverkamp, R. and Reggiani, P. 2002. Physical based water retention prediction models. In *Methods of Soil Analysis, Part 4, Physical Methods*, Dane, J. H. and Clarke, G. C. (Eds.), Soil Science Society of America Book Series, *Soil Sci. Soc. Am. Inc.*, Madison, WI, pp. 762–777.
- Haverkamp, R., Reggiani, P., Ross, P. J., and Parlange, J.-Y. 2002. Soil water hysteresis prediction model based on theory and geometric scaling. In *Environmental Mechanics: Water, Mass and Energy Transfer in the Biosphere*, Raats, P. A. C., Smiles, D., and Warrick, A. W. (Eds.), Geophysical Monograph 129, American Geophysical Union, Washington, DC, pp. 213–246.
- Haverkamp, R., Leij, F. J., Fuentes, C., Sciortino, A., and Ross, P. J. 2005. Soil water retention: I. Introduction of a shape index. *Soil Sci. Soc. Am. J.*, 69: 1881–1890.
- Heaslet, M. A. and Alksne, A. 1961. Diffusion from a fixed surface with a concentration-dependent coefficient. *J. Soc. Ind. Appl. Math.*, 9: 584–596.

- Heinen, M. and Raats, P. A. C. 1997. Hysteretic hydraulic properties of a coarse sand horticultural substrate. In *Characterization and Measurement of the Hydraulic Properties of Unsaturated Porous Media I*, van Genuchten, M. Th., Leij, F. J., and Wu, L. (Eds.), Dept. of Environmental Sciences, University of California, Riverside, CA, pp. 467–476.
- Hillel, D. 1980a. *Fundamentals of Soil Physics*. Academic Press, New York.
- Hillel, D. 1980b. *Applications of Soil Physics*. Academic Press, New York.
- Hoa, N. T., Gaudu, R., and Thirriot, C. 1977. Influence of hysteresis effects on transient flows in saturated–unsaturated porous media. *Water Resour. Res.*, 13: 992–996.
- Hodnett, M. G. and Tomasella, J. 2002. Marked differences between van Genuchten soil water retention parameters for temperate and tropical soils: a new water retention pedo-transfer function developed for tropical soils. *Geoderma*, 108: 155–180.
- Hogarth, W. L., Hoppmans, J., Parlange, J.-Y., and Haverkamp, R. 1988. Application of a simple soil water hysteresis model. *J. Hydrol.*, 98: 21–29.
- Kabat, P., Hutjes, R. W. A., and Feddes, R. A. 1997. The scaling characteristics of soil parameters: from plotscale heterogeneity to subgrid parametrization. *J. Hydrol.*, 190: 363–396.
- Kline, S. J. 1965. *Similitude and Approximation Theory*. McGraw-Hill Inc., New York.
- Knight, J. H. 1992. Sensitivity of time domain reflectometry measurements to lateral variations in soil water content. *Water Resour. Res.*, 28: 2345–2352.
- Kool, J. B. and Parker, J. C. 1987. Development and evaluation of closed-form expressions for hysteretic soil hydraulic properties. *Water Resour. Res.*, 23: 105–114.
- Kool, J. B., Parker, J. C., and van Genuchten, M. Th. 1987. Parameter estimation for unsaturated flow and transport models — A review. *J. Hydrol.*, 91: 255–293.
- Kool, J. B. and Parker, J. C. 1988. Analysis of the inverse problem for transient flow. *Water Resour. Res.*, 24: 817–830.
- Kutilek, M. and Nielsen, D. R. 1994. *Soil Hydrology*. Catena Verlag, Cremlingen-Desteds, Germany.
- Lambe, T. W. and Whitman, R. V. 1979. *Soil Mechanics, SI Version*. John Wiley & Sons, New York.
- Langhaar, H. L. 1951. *Dimensional Analysis and Theory of Models*. John Wiley & Sons Inc., New York.
- Lassabatère, L., Angulo-Jaramillo, R., Sorie-Ugalde, J. M., Cuenca, R., Braud, I., and Haverkamp, R. 2005. Beerkan estimation of soil transfer parameters through infiltration experiments — BEST. *Soil Sci. Soc. Am. J.*, 69: 1931–1942.
- Leij, F. J., Alves, W. J., van Genuchten, M. Th., and Williams, J. R. 1996. The UNSODA — Unsaturated Soil Hydraulic Database — User's Manual Version 1.0. *Report EPA/600/R-96/095*. National Risk Management Research Laboratory, Office of Research Development, U.S. Environmental Protection Agency, Cincinnati, OH, 45268: 1–103.
- Leij, F. J., Haverkamp, R., Fuentes, C., Zatarain, F., and Ross, P. J. 2005. Soil water retention: II. Derivation and application of shape index. *Soil Sci. Soc. Am. J.*, 69: 1891–1901.
- Lenhard, R. J., Parker, J. C., and Mishra, S. 1989. On the correspondence between Brooks–Corey and van Genuchten models. *J. Irrig. Drain. Eng.*, 115: 744–751.
- Liakopoulos, A. C. 1966. Theoretical approach of the infiltration problem, *Bull. IASH*, XII: 69–110.
- Liu, Y., Parlange, J.-Y., and Steenhuis, T. S. 1995. A soil water hysteresis model for fingered flow data. *Water Resour. Res.*, 31: 2263–2266.
- McCuen, R. H., Rawls, W. J., and Brakensiek, D. L. 1981. Statistical analysis of the Brooks and Corey and Green and Ampt parameters across soil texture. *Water Resour. Res.*, 17: 1005–1013.
- Mein, R. G. and Larson, C. L. 1973. Modeling infiltration during a steady rain. *Water Resour. Res.*, 9: 384–394.
- Miller, E. E. and Miller, R. D. 1955a. Theory of capillary flow: I. Practical implications. *Soil Sci. Soc. Am. J.*, 19: 267–271.
- Miller, E. E. and Miller, R. D. 1955b. Theory of capillary flow: II. Experimental information. *Soil Sci. Soc. Am. J.*, 19: 271–275.
- Miller, E. E. and Miller, R. D. 1956. Physical theory for capillary flow phenomena. *J. Appl. Phys.*, 27: 324–332.

- Millington, R. J. and Quirk, J. P. 1961. Permeability of porous solids. *Trans. Faraday Soc.*, 57: 1200–1206.
- Mishra, S. J., Parker, J. C., and Singhal, N. 1989. Estimation of soil hydraulic properties and their uncertainty from particle-size distribution data. *J. Hydrol.*, 108: 1–18.
- Mishra, S. J. and Parker, J. C. 1990. On the relation between saturated hydraulic conductivity and capillary retention characteristics. *Ground Water*, 28: 775–777.
- Mitchel, R. J. and Mayer, A. S. 1998. The significance of hysteresis in modelling solute transport in unsaturated porous media. *Soil Sci. Soc. Am. J.*, 62: 1506–1512.
- Morel-Seytoux, H. J., Meyer, P. D., Nachabe, M., Touma, J., van Genuchten M. Th., and Lenhard, R. J. 1996. Parameter equivalence for the Brooks–Corey and van Genuchten soil characteristics: preserving the effective capillary drive. *Water Resour. Res.*, 32: 1251–1258.
- Mualem, Y. 1974. A conceptual model of hysteresis. *Water Resour. Res.*, 10: 514–520.
- Mualem, Y. 1976. A new model for predicting the hydraulic conductivity of unsaturated porous media. *Water Resour. Res.*, 12: 513–522.
- Neuman, S. P. 1976. Wetting front pressure head in the infiltration model of Green and Ampt. *Water Resour. Res.*, 12: 564–566.
- Nielsen, D. R., Tillotson, P. M., and Vieira, S. R. 1983. Analyzing field-measured soil–water properties. *Agric. Water Manage.*, 6: 93–109.
- Nielsen, D.R., van Genuchten, M. T., and Biggar, J. W. 1986. Water flow and solute transport processes in the unsaturated zone. *Water Resour. Res.*, 22: 89–108.
- Otten, W., Raats, P. A. C., and Kabat, P. 1997. Hydraulic properties of root zone substrates used in greenhouse horticulture. In *Characterization and Measurement of the Hydraulic Properties of Unsaturated Porous Media I*, van Genuchten, M. Th., Leij, F. J., and Wu, L. (Eds.), Dept. of Environmental Sciences, University of California, Riverside, CA, pp. 477–488.
- Parker, J. C. and Lenhard, R. J. 1987. A model of hysteretic constitutive relations governing multiphase flow. 1. Saturation–pressure relations. *Water Resour. Res.*, 23: 2187–2196.
- Parlange, J.-Y. 1971a. Theory of water movement in soils: 1. One-dimensional absorption. *Soil Sci.*, 111: 134–137.
- Parlange, J.-Y. 1971b. Theory of water movement in soils: 2. One-dimensional infiltration. *Soil Sci.*, 111: 170–174.
- Parlange, J.-Y. 1975. On solving the flow equation in unsaturated soils by optimization: horizontal infiltration. *Soil Sci. Soc. Am. Proc.*, 39: 415–418.
- Parlange, J.-Y. 1976. Capillary hysteresis and relationship between drying and wetting curves. *Water Resour. Res.*, 12: 224–228.
- Parlange, J.-Y. 1980. Water transport in soils. *Ann. Rev. Fluid Mech.*, 12: 77–102.
- Parlange, J.-Y., Lisle, I., Braddock, R. D., and Smith, R. E. 1982. The three parameter infiltration equation. *Soil Sci.*, 133: 337–341.
- Parlange, J.-Y., Haverkamp, R., and Touma, J. 1985a. Infiltration under ponded conditions: 1. Optimal analytical solution and comparison with experimental observations. *Soil Sci.*, 139: 305–311.
- Parlange, J.-Y., Vauclin, M., Haverkamp, R., and Lisle, I. 1985b. Note: the relation between desorptivity and soil–water diffusivity. *Soil Sci.*, 139: 458–461.
- Parlange, M. B., Prasad, S. N., Parlange, J.-Y., and Römkens, M. J. M. 1992. Extension of the Heaslet–Alksne technique to arbitrary soil water diffusivities. *Water Resour. Res.*, 28: 2793–2797.
- Parlange, J.-Y., Steenhuis, T. S., Haverkamp, R., Barry, D. A., Culligan, P. J., Hogarth, W. L., Parlange, M. B., and Ross, P. J. 1999. Soil properties and water movement. In *Vadose Zone Hydrology*, Parlange, M. B. and Hopmans, J. (Eds.), Oxford University Press, New York, pp. 99–129.
- Peck, A. J. 1965a. Moisture profile development and air compression during water uptake by bounded bodies. 2. Horizontal columns. *Soil Sci.*, 99: 327–334.
- Peck, A. J. 1965b. Moisture profile development and air compression during water uptake by bounded bodies. 3. Vertical columns. *Soil Sci.*, 100: 44–51.
- Perroux, K. M. and White, I. 1988. Design of disc permeameters. *Soil Sci. Soc. Am. J.*, 52: 1205–1215.
- Philip, J. R. 1955. Numerical solution of equations of the diffusion type with diffusivity concentration-dependent: 1. *Trans. Faraday Soc.*, 51: 885–892.

- Philip, J. R. 1957a. The theory of infiltration: 1. The infiltration equation and its solution. *Soil Sci.*, 83: 345–357.
- Philip, J. R. 1957b. The theory of infiltration: 4. Sorptivity and algebraic infiltration equations. *Soil Sci.*, 84: 257–264.
- Philip, J. R. 1969a. Hydrostatics and hydrodynamics in swelling soils. *Water Resour. Res.*, 5: 1070–1077.
- Philip, J. R. 1969b. The theory of infiltration. *Adv. Hydrosci.*, 5: 215–305.
- Philip, J. R. 1973. On solving the unsaturated flow equation: 1. The flux–concentration relation. *Soil Sci.*, 116: 328–335.
- Philip, J. R. and Knight, J. H. 1974. On solving the unsaturated flow equation: 3. New quasi-analytical technique. *Soil Sci.*, 117: 1–13.
- Phuc, Le Van and Morel-Seytoux, H. J. 1972. Effect of soil air movement and compressibility on infiltration rates. *Soil Sci. Soc. Am. Proc.*, 36: 237–241.
- Poulvassilis, A. 1962. Hysteresis in pore water and application of the concept of independent domains. *Soil Sci.*, 93: 405–412.
- Poulvassilis, A. 1970. Hysteresis in pore water in granular porous bodies. *Soil Sci.*, 109: 5–12.
- Rawlins, S. L. and Campbell, G. S. 1986. Water potential: thermocouple psychrometry. In *Methods of Soil Analysis, Part 1, Physical and Mineralogical Methods*, Klute, A. (Ed.), Monograph 9, 2nd ed. Soil Science Society of America Book Series, *Soil Sci. So. Am. Inc.* Madison, WI.
- Rawls, W. J. and Brakensiek, D. L. 1982. Estimating soil water retention from soil properties. *J. Irrig. Drain. Div.*, ASCE, 108: 166–171.
- Rawls, W. J. and Brakensiek, D. L. 1983. Green-ampt infiltration parameters from soil data. *J. Hydraul. Div. Am. Soc. Civ. Eng.*, 109 HY1: 62–70.
- Rawls, W. J. and Brakensiek, D. L. 1985. Prediction of soil water properties for hydrologic modeling. Watershed Management in the Eighties. *Proc. Irrig. Drain. Div.*, ASCE, Denver, Colorado.
- Rawls, W. J. and Brakensiek, D. L. 1989. Estimating soil hydraulic properties. In *Unsaturated Flow in Hydrologic Modeling*, Morel-Seytoux, H. J. (Ed.), Kluwer Academic Publishers, published in cooperation with NATO Scientific Affairs Division, Boston.
- Rawls, W. J., Ahuja, L. R., and Brakensiek, D. L. 1992. Estimating soil hydraulic properties from soils data. In *Indirect Methods for Estimating the Hydraulic Properties of Unsaturated Soils*, van Genuchten, M. Th., Leij, F. J., and Lund, L. J. (Eds.), Dept. of Environmental Sciences, University of California, Riverside, CA, pp. 329–340.
- Reichardt, K., Nielsen, D. R., and Bigger, J. W. 1972. Scaling of horizontal infiltration into homogeneous soils. *Soil Sci. Soc. Am. Proc.*, 36: 241–245.
- Richards, L. A. 1931. Capillary conduction of liquids through porous media. *Physics* 1: 318–333.
- Rijtema, P. E. 1965. An analysis of actual evapotranspiration. *Agric. Res. Rep.* 659, Center for Agricultural Publications and Documentation, Wageningen, The Netherlands.
- Rogowski, A. S. 1971. Watershed physics: model of soil moisture characteristics. *Water Resour. Res.*, 7: 1575–1582.
- Rose, C. W. 1966. *Agricultural Physics*. Pergamon Press, New York.
- Ross, P. J. and Smettem, K. R. J. 1993. Describing soil hydraulic properties with sums of simple functions. *Soil. Sci. Soc. Am. J.*, 57: 26–29.
- Ross, P. J., Haverkamp, R., and Parlange, J.-Y. 1996. Calculating parameters of infiltration equations from soil hydraulic functions. *Transp. Porous Media*, 24: 315–339.
- Russo, D., Jury, W. A., and Butters, G. L. 1989. Numerical analysis of solute transport during transient irrigation: 1. The effect of hysteresis and profile heterogeneity. *Water Resour. Res.*, 25: 2109–2118.
- Russo, D., Bresler, E., Shani, U., and Parker, J. C. 1991. Analyses of infiltration events in relation to determining soil hydraulic properties by inverse problem methodology. *Water Resour. Res.*, 27: 1361–1373.
- Salvucci, G. D. 1997. Soil and moisture independent estimation of stage-two evaporation from potential evaporation and albedo or surface temperature. *Water Resour. Res.*, 33: 111–122.
- Si, B. C. and Kachanoski, R. G. 2000. Unified solution for infiltration and drainage with hysteresis: theory and field test. *Soil Sci. Soc. Am. J.*, 64: 30–36.

- Simmons, C. R., Nielsen, D. R., and Biggar, J. W. 1979. Scaling of field-measured soil–water properties: I. Methodology. *Hilgardia*, 47: 77–102.
- Simunek, J., Angulo-Jaramillo, R., Schaap, M. G., Vandervaere, J.-P., and van Genuchten, M. Th. 1998. Using an inverse method to estimate the hydraulic properties of crusted soils from tension-disc infiltrometer data. *Geoderma*, 86: 61–81.
- Smettem, K. R. J., Parlange, J.-Y., Ross, P. J., and Haverkamp, R. 1994. Three-dimensional analysis of infiltration from the disc infiltrometer: 1. A capillary-based theory. *Water Resour. Res.*, 30: 2925–2929.
- Smiles, D. E. and Rosenthal, M. J. 1968. The movement of water in swelling material. *Aust. J. Soil Res.*, 6: 237–248.
- Smiles, D. E. 1974. Infiltration into a swelling material. *Soil Sci.*, 117: 140–147.
- Smiles, D. E. 1995. Liquid flow in swelling soils. *Soil Sci. Soc. Am. J.*, 59: 313–318.
- Soil Survey Laboratory Staff. 1992. Soil survey laboratory methods manual. *Soil Surv. Invest. Rep. 42 Version 2.0*, USDA-SCS U.S. Government Printing Office, Washington, DC.
- Sposito, G. 1973. Volume changes in swelling clays. *Soil Sci.*, 115: 315–320.
- Sposito, G. 1975a. Steady vertical flows in swelling soils. *Water Resour. Res.*, 11: 461–464.
- Sposito, G. 1975b. A thermodynamic integral equation for the equilibrium moisture profile in swelling soils. *Water Resour. Res.*, 11: 490–500.
- Sposito, G. 1975c. On the differential equation for the equilibrium moisture profile in swelling soils. *Soil Sci. Soc. Am. Proc.*, 39: 1053–1056.
- Sposito, G. and Jury, W. A. 1985. Inspectional analysis in the theory of water flow through unsaturated soil. *Soil Sci. Soc. Am. J.*, 49: 791–798.
- Sposito, G. 1996. *Scale Invariance and Scale Dependence in Hydrology*. Cambridge University Press, New York.
- Starr, J. L., de Roo, H. C., Frink C. R., and Parlange, J.-Y. 1978. Leaching characteristics of a layered field soil. *Soil Sci. Soc. Am. J.*, 42: 386–391.
- Swartzendruber, D. 1987. Rigorous derivation and interpretation of the Green and Ampt equation. In *Infiltration Development and Application*, Fok, Y.-S. (Ed.), Water Resources Research Center, University of Hawaii, Honolulu, Hawaii, U.S.A.
- Talsma, T. and Parlange, J.-Y. 1972. One-dimensional vertical infiltration. *Aust. J. Soil Res.*, 10: 143–150.
- Tietje, O. and Tapkenhinrichs, M. 1993. Evaluation of pedo-transfer functions. *Soil Sci. Soc. Am. J.*, 57: 1088–1095.
- Tillotson, P. M. and Nielsen, D. R. 1984. Scale factors in soil science. *Soil Sci. Soc. Am. J.*, 48: 953–959.
- Tomasella, J. and Hodnett, M. G. 1998. Estimating soil water retention characteristics from limited data in Brazilian Amazonia. *Soil Sci.*, 163: 190–202.
- Tomasella, J., Hodnett, M. G., and Rossato, L. 2000. Pedo-transfer functions for the estimation of soil water retention in Brazilian soils. *Soil Sci. Soc. Am. J.*, 64: 327–338.
- Topp, G. C. 1971. Soil–water hysteresis: the domain theory extended to pore interaction conditions. *Soil Sci. Soc. Am. Proc.*, 35: 219–225.
- Topp, G. C., Davis, J. L., and Annam, A. P. 1980. Electromagnetic determination of soil water content: measurement in coaxial transmission lines. *Water Resour. Res.*, 16: 574–582.
- Touma, J., Vachaud, G., and Parlange, J.-Y. 1984. Air and water flow in a sealed ponded vertical soil column. *Soil Sci.*, 137: 181–187.
- Vachaud, G. and Thony, J. L. 1971. Hysteresis during infiltration and redistribution in a soil column at different initial water contents. *Water Resour. Res.*, 7: 111–127.
- Vachaud, G., Vauclin, M., Khanji, D., and Wakil, N. 1973. Effects of air pressure on water flow in an unsaturated stratified vertical column of sand. *Water Resour. Res.*, 9: 160–173.
- Vachaud, G., Gaudet, J. P., and Kuraz, V. 1974. Air and water flow during ponded infiltration in a vertical bounded column of soil. *J. Hydrol.*, 22: 89–108.
- van Bavel, C. H. M. 1969. The three phase domain in hydrology. In *Water in the Unsaturated Zone*, Rijtema, P. E. and Wassink, H. (Eds.), UNESCO, Paris, France, Vol. 1, pp. 23–32.

- van Dam, J. C., Stricker, J. N. M., and Droogers, P. 1994. Inverse method to determine soil hydraulic functions from multistep outflow experiments. *Soil Sci. Soc. Am. J.*, 58: 647–652.
- van Genuchten, M. Th. 1980. A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. *Soil Sci. Soc. Am. J.*, 44: 892–898.
- van Genuchten, M. Th., Leij, F. J., and Yates, S. R. 1991. The RETC code for quantifying the hydraulic functions of unsaturated soils. *Report EPA/600/2-91/065*, U.S. Environmental Protection Agency, Ada, OK.
- Vauclin, M., Haverkamp, R., and Vachaud, G. 1984. Error analysis in estimating soil water content from neutron probe measurements II. Spatial standpoint. *Soil Sci.*, 137: 141–148.
- Vereecken, H., Maes, J., Feyen, J., and Darius, P. 1989. Estimating the soil moisture retention characteristic from texture, bulk density and carbon content. *Soil Sci.*, 148: 389–403.
- Vereecken, H., Diels, J., van Orshoven, J., Feyen, J., and Bouma, J. 1992. Functional evaluation of pedotransfer functions for the estimation of soil hydraulic properties. *Soil Sci. Soc. Am. J.*, 56: 1371–1378.
- Warrick, A. W., Mullen, G. J., and Nielsen, D. R. 1977. Scaling field-measured soil hydraulic properties using a similar media concept. *Water Resour. Res.*, 13: 355–362.
- Warrick, A. W. and Nielsen, D. R. 1980. Spatial variability of soil physical properties in the field. In *Applications of Soil Physics*, Hillel, D. (Ed.), Academic Press, New York, pp. 319–344.
- Warrick, A. W., Lomen, D. O., and Yates, S. R. 1985. A generalized solution to infiltration. *Soil Sci. Soc. Am. J.*, 49: 34–38.
- Whitmore, A. P. and Heinen, M. 1999. The effect of hysteresis on microbial activity in computer simulation models. *Soil Sci. Soc. Am. J.*, 63: 1101–1105.
- Wooding, R. A. 1968. Steady infiltration from a shallow circular pond. *Water Resour. Res.*, 4: 1259–1273.
- Woods, R. A., Sivapalan, M. and Duncan, M. 1995. Investigating the representative area concept: an approach based on field data. *Hydrol. Process.*, 9: 291–312.
- Wösten, J. H. M. and van Genuchten, M. Th. 1988. Using texture and other soil properties to predict the unsaturated soil hydraulic functions. *Soil Sci. Soc. Am. J.*, 52: 1762–1770.
- Yates, S. R., van Genuchten, M. Th., Warrick, A. W., and Leij, F. J. 1992. Analysis of measured, predicted, and estimated hydraulic conductivity using the RETC computer program. *Soil Sci. Soc. Am. J.*, 56: 347–354.
- Yong, R. N. and Warkentin, B. P. 1966. *Introduction to Soil Behavior*. Macmillan, New York.
- Youngs, E. G. 1968. An estimation of sorptivity for infiltration studies from moisture moment considerations. *Soil Sci.*, 106: 157–163.

Further Reading

1. Haverkamp et al. (2002) provides a useful prediction model for soil water hysteresis based on theory and geometric scaling.
2. Hillel (1980a, 1980b) provided a comprehensive review on the fundamentals and applications of soil physics. In spite of the fact that the books are probably difficult to get hold of, it is worthwhile trying as they are well-written and reader friendly.
3. Kutilek and Nielsen (1994) put together a comprehensive review on soil hydrology. The work is an excellent source of updated references.
4. Soil Science Society of America Book Series edited a volume: Methods of Soil Analysis, Part 4 (Eds. J. Dane and G. Clarke Topp) which provides a comprehensive review on methods of soil analysis.
5. Parlange et al. (1999) provided an excellent theoretical overview on infiltration and water movement in unsaturated soils.
6. Sposito (1996) has put together the latest developments about scaling and invariance in hydrology.

7

Infiltration

Jan W. Hopmans
University of California, Davis
Jean-Yves Parlange
Cornell University
Shmuel Assouline
Volcani Center, Bet Dagan

7.1	Introduction.....	7-1
7.2	Soil Properties and Processes of Infiltration	7-2
7.3	Infiltration Equations..... Philip Infiltration Equation (PH) • Parlange et al. (PA) Model • Swartzendruber (SW) Model • Empirical Infiltration Equations	7-7
7.4	Infiltration Measurement Methods.....	7-10
7.5	Spatial Variability Considerations	7-11
7.6	Infiltration into Sealed or Crusted Soils.....	7-12
7.7	Summary and Conclusions	7-14
	References	7-15
	Further Information	7-18

7.1 Introduction

Water entry into the soil by infiltration is among the most important of soil hydrological processes, as it controls the partitioning between runoff and soil water storage. Runoff water determines the surface water's quantity and quality, whereas infiltrated water determines the plant's available water, evapotranspiration, groundwater recharge, and groundwater quality. Also through exinfiltration, infiltrated water can affect the surface water's characteristics as in riparian zones. Yet, despite its relevance and our solid physical understanding of infiltration, we have generally great difficulty in predicting the infiltration in the field at larger spatial scales. Mostly, this is so because the infiltration rate is a time-varying quantity that is highly dependent on soil properties that are generally highly variable in space, both in the vertical and horizontal directions of the hydrologic basin. Moreover, infiltration rate and runoff are affected by vegetation cover, as it protects the soil's surface from the energy impacts of falling rain drops or intercepting rainfall, serving as a temporary water storage. The kinetic energy of rainfall causes soil degradation, leading to soil surface sealing and decreasing infiltration.

Historically, solutions to infiltration problems have been presented through analytical or empirical means. Analytical solutions provide values of infiltration rate or cumulative infiltration as a function of time, making simplified assumptions of soil depth variations of water content, before and during the infiltration. In contrast, powerful computers nowadays use numerical simulations of unsaturated water flow to solve for water content and water fluxes throughout the unsaturated soil domain in a single vertical direction or in multiple spatial dimensions, allowing complex initial and boundary conditions. However, although the multi-dimensional unsaturated water flow modeling is extremely useful for many vadose zone applications, it does not necessarily improve the soil surfaces infiltration rate prediction, in light of

the large uncertainty of the soil's physical properties and the initial and boundary conditions that control the infiltration. In contrast, empirical infiltration models serve primarily to fit model parameters to the measured infiltration, but have limited power as a predictive tool.

In this chapter, we will first review the basic processes of infiltration and introduce the relevant soil and infiltration parameters that control infiltration (Section 7.2). In doing so, we will use the notation introduced by Haverkamp et al. (this volume), and start with the so-called Green and Ampt equation for ponded infiltration. In Section 7.3 we introduce the most frequently used physically-based and empirical infiltration equations. We will present methods of measuring infiltration (Section 7.4), and demonstrate that specific field tests can be used to characterize the soil's physical properties in an inverse way. In Section 7.5, we review the spatial considerations of infiltration and highlight infiltration studies at the basin scale, whereas Section 7.6 presents the effects of the soil surface sealing on infiltration and the implications in terms of runoff estimates at the watershed scale. We will conclude with a summary in Section 7.7.

7.2 Soil Properties and Processes of Infiltration

The soil is a complex three-phase porous medium of mostly connected solid, liquid, and gaseous phases, with the spatial distribution and geometrical arrangement of each phase, and the partitioning of solutes between phases, controlled by physical, chemical, and biological processes. The vadose zone is bounded by the soil surface and merges with the groundwater in the capillary fringe. Water in the unsaturated soil matrix is attracted by capillary and adsorptive forces. Soils are subjected to fluctuations in water and chemical content by infiltration and leaching from rainfall or irrigation, water uptake by plant roots (transpiration), and evaporation from the soil surface.

As introduced by Haverkamp (this issue), we represent the soil's dry bulk density by ρ_b , and define the porosity ($\varepsilon L^3 L^{-3}$) by:

$$\varepsilon = 1 - \frac{\rho_b}{\rho_s} \quad (7.1)$$

with ρ_s being the soil's particle density. *Soil porosity* decreases as *bulk soil density* increases. The volumetric water content θ expresses the volume of water present per unit bulk soil as

$$\theta = \frac{w \rho_b}{\rho_w} \quad (7.2)$$

where w is defined as the mass water content and we take $\rho_w = 1000 \text{ kg m}^{-3}$. Alternatively, the soil water content can be described by the degree of saturation $S(-)$, or the equivalent depth of stored water $D_e(L)$,

$$S = \frac{\theta}{\varepsilon} \quad \text{and} \quad D_e = \theta D_{\text{soil}} \quad (7.3)$$

so that θ can also be defined by the equivalent depth of water per unit depth of bulk soil, D_{soil} . (L). The volumetric water content ranges between 0.0 (dry soil) to the saturated water content θ_s , which is equal to the porosity if the soil were completely saturated. The degree of saturation varies between 0.0 (completely dry) and 1.0 (all pores completely water-filled). When considering water flow, the porosity term is replaced by the saturated water content θ_s , and both terms in Equation 7.3 are corrected by subtracting the so-called residual water content θ_r , so that the effective saturation S_e , is defined as

$$S_e = \frac{\theta - \theta_r}{\theta_s} \quad (7.4)$$

The soil-water retention function determines the relation between the volume of water retained by the soil, expressed by θ , and the governing soil matric, or suction forces. These suction forces are typically expressed

by the soil water matric head (L, strictly negative) or soil suction (strictly positive). Also known as the soil-water release or soil-water characteristic function, this soil hydraulic property describes the increase of θ and the size of the water-filled pores with a decrease in matric potential, as occurs by infiltration. As the matric forces are controlled by pore size distribution, specific surface area, and type of physico-chemical interactions at the solid–liquid interfaces, the soil water retention curve is very soil-specific and very nonlinear. It provides an estimate of the soil's capacity to hold water after irrigation and free drainage (field capacity), minimum soil water content available to the plant (wilting point), and therefore water availability in the rootzone for plants. The soil water retention curves shows hysteresis, that is, the θ value is different for wetting (infiltration) and drying (drainage).

The relation between the soil's unsaturated hydraulic conductivity, K , and volumetric water content, θ , is the second essential fundamental soil hydraulic property needed to describe water movement in soils. It is also a function of the water and soil matrix properties, and determines water infiltration and drainage rates, and is strongly affected by water content and possibly by hysteresis. It is defined by the Darcy–Buckingham equation, which relates the soil-water flux density to the total driving force for flow, with K being the proportionality factor. Except for special circumstances, the total driving force for water flow in soils is determined by the sum of the matric and gravitational forces, expressed by the total water potential head gradient, $\Delta H/L$ (L L⁻¹), where ΔH denotes the change in total water potential head over the distance L , and $H = (h + z)$. For vertical flow, the application of Darcy's law yields the magnitude of water flux from the steady-state flow equation:

$$q = -K(\theta) \left(\frac{dh}{dz} + 1 \right) \quad (7.5)$$

where q is the Darcy water flux density (L³L⁻² T⁻¹), z is vertical position ($z > 0$, upwards, L). A soil system is usually defined by the bulk soil, without consideration of the size and geometry of the individual flow channels or pores. Therefore, the hydraulic conductivity (K) describes the ability of the bulk soil to transmit water, and is expressed by the volume of water flowing per unit area of bulk soil per unit time. Functional models for unsaturated hydraulic conductivity can be based on pore size distribution, pore geometry, and connectivity, and require the integration of soil water retention functions to obtain analytical expressions for the unsaturated hydraulic conductivity. The resulting expressions relate the relative hydraulic conductivity, K_r , defined as the ratio of the unsaturated hydraulic conductivity, K , and the saturated hydraulic conductivity, K_s , to the effective saturation, S_e , or matric potential head h , as in Equation 7.20b. We note that θ_s in the vadose zone is typically about 85% of the porosity, so that a saturated soil (e.g., as the result of ponded infiltration) is really a “saturated” soil due to entrapped air, with a saturated hydraulic conductivity that is significantly smaller than the true K_s .

For transient vertical water flow, with no root water uptake, the Darcy–Buckingham equation for unsaturated flow is combined with the mass balance equation to yield:

$$\frac{\partial \theta}{\partial t} = -\frac{\partial q}{\partial z} \quad (7.6)$$

which is defined as the Richards' equation after substitution of Equation 7.5 and written with h being the single dependent variable (Haverkamp, et al., this volume). Numerous studies have been published addressing different issues in the numerical modeling of unsaturated water flow in multiple spatial dimensions, using the Richards' equation. For the purpose of deriving analytical infiltration models, a linearized version of Equation 7.6 is often used. Specifically, for one-dimensional infiltration, the infiltration rate (L T⁻¹), $i(t)$, can be defined by Equation 7.5 at the soil surface (subscript $surf$), or

$$i(t) = -K(\theta) \left(\frac{\partial h}{\partial z} + 1 \right)_{surf} \quad (7.7a)$$

Cumulative infiltration $I(t)$, expressed as volume of water per unit soil surface area (L), is defined by

$$I(t) = \int_0^t i(t) dt \quad (7.7b)$$

Analytical solutions of infiltration generally assume that the wetted soil profile is homogeneous in texture with uniform initial water content. They also make distinction between ponded (positive h) and non-ponded soil surface (negative h) infiltration. The infiltration capacity of the soil is defined by $i_c(t)$, defined by the maximum rate at which a soil can absorb water for ponded soil-surface conditions. Its maximal value is at time zero, and decreases with time to its minimum value approaching the soil's saturated hydraulic conductivity, K_s , as the total water potential gradient decreases, and tends to unity, with the downward moving wetting front. As defined by Equation 7.7b, the soil's cumulative infiltration capacity, $I_c(t)$ is defined by the area under the capacity curve. It represents the maximum amount of water that the soil can absorb at any time. Typically at the onset of infiltration ($t = 0$), the rainfall rate $r(t)$, will be lower than $i_c(t)$, so that the infiltration rate is equal to the rainfall rate (i.e., $r(t) < i_c(t)$ for $h_{surf} < 0$). If at any point in time, the rainfall rate becomes larger than the infiltration capacity, ponding will occur ($h_{surf} > 0$), resulting in runoff. The time at which ponding occurs is defined as t_p . Thus, the actual infiltration rate will depend on the rainfall rate and its temporal changes. This makes prediction of infiltration and runoff much more difficult for realistic time-variable rainfall patterns.

Therefore infiltration rate prediction is often described as a function of the cumulative infiltration, I , or $i(I)$, independent of the time domain, and with $i(I)$ curves that are independent of rainfall rate (Skaggs, 1982). An example of such a time-invariant approach is the IDA or infiltrability-depth approximation (Smith et al., 2002). The main IDA assumption is that time periods between small rainfall events are sufficiently small so that soil water redistribution and evaporation between events does not affect the infiltration rate. IDA implies that the infiltration capacity at any given time depends only on the cumulative infiltration volume, regardless of the previous rainfall history. Following this approach, t_p is defined as the time during a storm event when I becomes equal to $I_c(t_p)$, or $R = \int_{t=0}^{t_p} r(t) dt = I_c(t_p)$, whereas $i(t) = r(t)$ for $t < t_p$. The time invariance of $i(I)$ holds true also when a layered/sealed soil profile is considered (Mualem and Assouline, 1989). However, it may not remain valid during the dynamic stage of seal formation (Assouline and Mualem, 2001).

For illustration purposes, we present an hypothetical storm event with time-varying $r(t)$ in Figure 7.1, in combination with an assumed soil-specific infiltration capacity curve, $i_c(t)$. At what time will ponding occur? It will not be at $t = 7$, when $r(t)$ exceeds i_c for the first time. To approximate t_p , we plot both I_c and R for the storm in Figure 7.2a, as a function of time and determine t_p as the time at which both curves intersect ($t_p = 13$, for $R = I_c = 110$), since at that time, the cumulative infiltration of the storm is identical to the soil's infiltration capacity. The final corresponding $i(I)$ for this soil and storm event is presented in Figure 7.2b, showing that the soil infiltration rate is equal to $r(t)$ until $I = R(t) = I_c(t_p) = 110$, after which the infiltration rate is soil-controlled and determined by $I_c(t)$. More accurate approximations to the time-invariant approach can be found in Sivapalan and Milly (1989) and Brutsaert (2005), using the time compression or time condensation approximation that more accurately estimate infiltration prior to surface ponding.

In addition to whether the soil is ponded or not, solutions of infiltration also distinguish between cases with and without gravity effects, as different analytical solutions apply. As Equation 7.7a shows, infiltration rate $i(t)$ is determined by both the soil water matric potential gradient, dh/dz , and gravity. However, at the early stage of infiltration into a relatively dry soil, infiltration rate is dominated by the matric potential gradient so that the gravity effects on infiltration can be ignored. Gravity becomes important in the later stages of infiltration, when the wetting front has moved down. For gravity-free drainage, a simple analytical solution can be found, after transforming Equation 7.6 into a diffusion type of equation by defining the diffusivity $D(\theta) = K(\theta)(dh/d\theta)$, and writing Equations 7.6 and Equation 7.7 in the θ -based

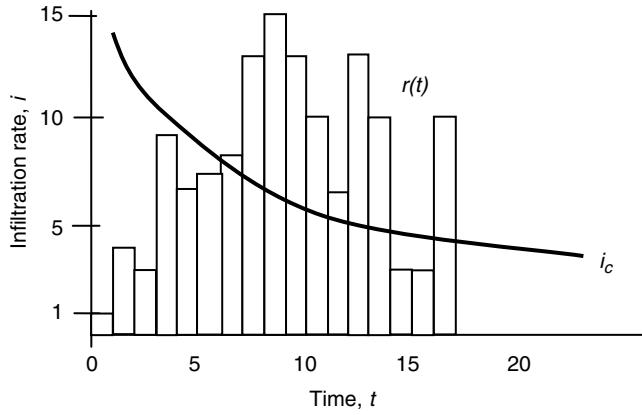


FIGURE 7.1 Hypothetical rainfall event, $r(t)$, and soil infiltration capacity, $i_c(t)$. The rainfall event starts at $t = 0$.

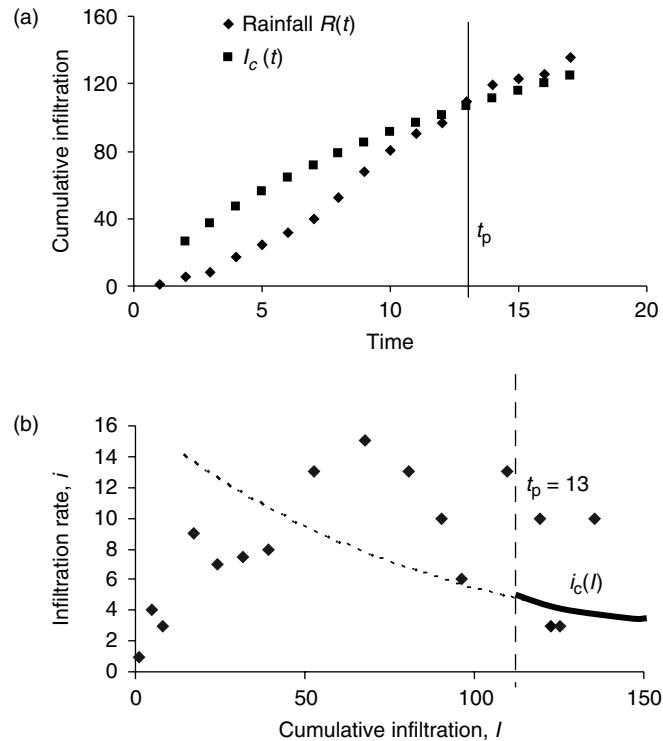


FIGURE 7.2 (a) Cumulative infiltration corresponding with infiltration capacity, $I_c(t)$, and cumulative rainfall, $R(t)$ and (b) actual infiltration rate vs I . Ponding starts only after $t_p = 13$, or $I = 110$.

form, so that:

$$i(t) = -D(\theta) \left(\frac{\partial \theta}{\partial z} \right)_{\text{surf}} \quad (7.8)$$

Using the Boltzmann transformation for a constant head boundary condition (Bruce and Klute, 1956), and defining the scaling variable $\varphi = z/t^{1/2}$, a combination of Equation 7.6 and Equation 7.8 results

in a unique solution of θ as a function of ϕ , from which the wetting profile can be computed for any time t (Kirkham and Powers, 1972). Using the terminology of Haverkamp et al. (this volume), θ_1 and θ_o define the surface water content during infiltration and uniform initial profile water content, respectively. Moreover, defining I as the volume of infiltrated water within the wetting soil profile, or

$$I = \int_{\theta_0}^{\theta_1} z d\theta = t^{1/2} \int_{\theta_0}^{\theta_1} \varphi d\theta \quad (7.9a)$$

results in the simple infiltration equation $I = St^{1/2}$ where the sorptivity $S(LT^{-1/2})$ is defined as:

$$S(\theta_1) = \int_{\theta_0}^{\theta_1} \varphi d\theta \quad (7.9b)$$

Equation 7.9a states that for “gravity-free” infiltration during the early times of vertical infiltration, and at all times for horizontal infiltration, I is a linear function of $t^{1/2}$, with S being defined as the slope of this line. Hence, for saturated soil conditions where $\theta_1 = \theta_s$, the infiltration capacity is computed from $i_c(t) = 1/2St^{-1/2}$. By the way, this also leads to $I_c = S^2/2i_c$.

A relatively simple analytical solution without and with gravity effects was suggested by Green and Ampt (GA, 1911) for a ponded soil surface, with $\theta_{surf} = \theta_1$. The assumptions are that the wetting front can be approximated as a step function (Figure 7.3) with a constant effective water potential, h_f , at the wetting front, a wetting zone hydraulic conductivity of $K(\theta_1) = K_1 = K_s$, and a constant soil water profile of $\Delta\theta = \theta_1 - \theta_o$. Using this so-called delta-function assumption of a $D(\theta)$ with a Dirac-delta function form, both solutions for horizontal and vertical infiltration can be relatively easily obtained (Jury et al., 1991; Haverkamp et al., this volume). Assuming that K_o at the initial water content, θ_0 , is negligible, the GA solution of vertical infiltration for ponded conditions is ($h = h_{surf} > 0$):

$$I = I_c = K_1 t + (h_{surf} - h_f) \Delta\theta \ln \left(1 + \frac{I}{(h_{surf} - h_f) \Delta\theta} \right) \quad (7.10)$$

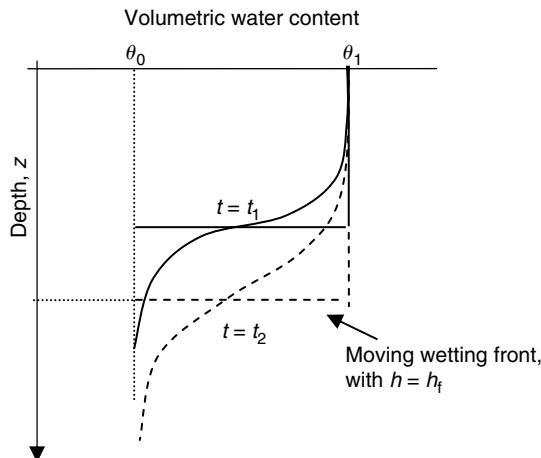


FIGURE 7.3 Water content distributions for an infiltrating soil profile under ponding at two times with $t_2 > t_1$. Step function represents Green and Ampt delta function infiltration model.

which can be solved iteratively for I . This simple, yet physically-based solution appears to work best for dry coarse-textured soils. A theoretical expression for the wetting front potential head, h_f , was defined by Mein and Farrel (1974), to yield that $h_f = \int_0^{h_o} K_r(h)dh$, where the relatively conductivity $K_r = K(h)/K_s$. The so-called S-form of the GA equation can be obtained by comparing the gravity-free solution of GA with the Boltzmann solution, to yield ($S_0^2 = -2K_1 h_f \Delta\theta$):

$$S_1^2(\theta_1) = 2K_1 \Delta\theta(h_{\text{surf}} - h_f) = S_0^2 + 2K_1 h_{\text{surf}} \Delta\theta \quad (7.11a)$$

so that

$$I = K_1 t + h_{\text{surf}} \Delta\theta + \frac{S_0^2}{2K_1} \ln \left(1 + \frac{I}{h_{\text{surf}} \Delta\theta + S_0^2 / 2K_1} \right) \quad (7.11b)$$

In reality, the wetting front is not a step function, but will consist of a time-dependent transition zone where water content changes from θ_1 to θ_0 (Figure 7.3). The shape of this transition zone will be a function of time and is controlled by soil type. The step function assumption is better for uniform coarse-textured soils that have a Dirac-like $D(\theta)$, for which there is a sharp decline in K with a decrease in water content near saturation. The wetting front is generally much more diffuse for finer-textured soils that have a wide pore-size distribution.

By now it must be clear that infiltration and its temporal changes are a function of many different soil's factors. In addition to rainfall intensity and duration and the soil physical factors such as soil water retention and hydraulic conductivity, infiltration is controlled by the initial water content, surface sealing and crusting, soil layering and the ionic composition of the infiltrated water (Kutilek and Nielsen, 1994; Assouline, 2004). For example, Vandervaere et al. (1998) applied the GA model to sealed soil profiles, by assuming that the wetting front potential decreases suddenly as it leaves the seal and enters the soil. This results in a discontinuous drop in the infiltration rate. Many relatively simple infiltration equations have been proposed and are successfully used in to characterize infiltration. This has been achieved despite the fact that these equations can be applied for homogeneous soils only, in theory.

7.3 Infiltration Equations

In addition to the solutions in Section 7.2, other physically-based analytical solutions have been presented. These can be potentially be used to predict infiltration from known soil hydraulic properties of homogeneous soils. However, in practice this is difficult as soil's physical characteristics near the soil surface are highly time-dependent because of the soil's structural changes, and because they are highly variable spatially. Alternatively, various empirical infiltration models have been proposed that are very useful for describing the measured infiltration data. A parameter sensitivity analysis of many of the presented infiltration models, analyzing the effects of measurement error, was given by Clausnitzer et al. (1998). This section presents the most frequently-used infiltration models in both categories.

7.3.1 Philip Infiltration Equation (PH)

Philip (1957a) presented an analytical infinite-series solution to the water-content based form of Richards' equation for the case of vertical infiltration. The so-called Fokker–Planck equation, or

$$\frac{\partial\theta}{\partial t} = \frac{\partial}{\partial z} \left(D \frac{\partial\theta}{\partial z} + K(\theta) \right) \quad (7.12)$$

for $h_{\text{surf}} = 0$ and $\theta_1 = \theta_s$ converges to the true solution for small and intermediate times, but fails for large times. In this case an alternative solution was presented (Philip, 1957b). With additional assumptions regarding the physical nature of soil water properties, Philip (1987) proposed joining solutions that are

applicable for all times. Philip (1957c) introduced a truncation of the small-time series solution that is a simple two-parameter model equation (PH model)

$$I_c = At + St^{1/2} \quad (7.13a)$$

“which should be accurate for all but very large t ”, and “suitable for applied hydrological studies.” The sorptivity S depends on several soil-physical properties, including initial water content θ_0 , and the hydraulic conductivity and soil-water retention functions. S is equal to the expression defined in Equation (7.9b). Because the coefficient A is not equal to K_1 , physical interpretation of Equation 7.13 fails for $t \rightarrow \infty$. Philip (1969) showed that A may take values between $0.38K_s$ and $0.66K_s$. However, for long times when gravity is dominant, A must be equal to K_s .

Differentiation of Equation 7.13 yields the infiltration rate, or

$$i_c = 1/2St^{-0.5} + A \quad (7.13b)$$

Using (Equation 7.13b) to express t as a function of i_c and substituting in Equation 7.13a yields $I(i)$, or

$$I = \frac{S^2(i - A/2)}{2(i - A)^2} \quad (7.13c)$$

For positive pressure heads, the correction of Equation (7.11a) to S can be applied. In many cases, values of S and A are obtained from curve fitting. We note that for gravity-free flow, the PH solution without the gravity term corresponds with the Boltzmann solution for horizontal flow in Equation 7.9.

7.3.2 Parlange et al. (PA) Model

Parlange et al. (1982) proposed the following universal model:

$$t = \frac{S^2}{2K_1^2(1-\delta)} \left[\frac{2K_1}{S^2} I - \ln \frac{\exp\left(\frac{2\delta K_1 I}{S^2}\right) + \delta - 1}{\delta} \right] \quad (7.14a)$$

assuming that K_0 is negligible small so that the ΔK in Parlange et al. (1982) is equal to K_1 . The value of the parameter δ can be chosen to approach various closed-form solutions. For example, Equation 7.14 reduces to the GA solution for δ equal to zero. Its value is a function of $K(\theta)$, and is defined by (Parlange et al., 1985) as

$$\delta = \frac{1}{\theta_s - \theta_0} \int_{\theta_0}^{\theta_s} \frac{K_s - K(\theta)}{K_s} d\theta \quad (7.14b)$$

An approximate value of $\delta = 0.85$ was suggested by Parlange et al. (1982) for a range of soil types. After taking the time derivative of I , the following $i(I)$ -relationship can be derived (Espinoza, 1999):

$$i = K_1 + \delta K_1 \left(1 - \exp\left(\frac{2I\delta K_1}{S^2}\right) \right)^{-1} \quad (7.14c)$$

Because this is based on integration of the water-content based form of Richards' equation, the theoretical scope of Equation 7.14 is limited to non-ponded conditions. A generalization of Equation 7.14 to include ponded conditions without affecting the value of S was introduced by Parlange et al. (1985). Haverkamp et al. (1990) presented a modification of their model to include upward water flow by capillary rise. The

resulting infiltration model contained six physical parameters, in addition to the interpolation parameter δ (Haverkamp et al., 1990). Both the PA and the Haverkamp model require an iterative procedure to predict $I(t)$. Barry et al. (1995) presented an explicit approximation to the Haverkamp et al. (1990) model, retaining all six physical parameters (BA model):

$$I = K_1 t + \frac{S^2 + 2K_1 h_{\text{surf}} \Delta\theta}{2\Delta K} \left[t^* + 1 - \gamma - \exp\left(\frac{-6(2t^*)^{0.5}}{6 + (2t^*)^{0.5}} - \frac{2t^*}{3}\right) \right. \\ \left. + \frac{\gamma}{1+t^*} \left\{ \exp\left(-\frac{2t^*}{3}\right) [1 - (1-\gamma)^8 t^{*2.5}] + (2\gamma + t^*) \ln\left(1 + \frac{t^*}{\gamma}\right) \right\} \right] \quad (7.15a)$$

where

$$t^* = \frac{2t(\Delta K)^2}{S^2 + 2K_1 h_{\text{surf}} \Delta\theta}, \quad \gamma = \frac{2K_1(h_{\text{surf}} + h_a)\Delta\theta}{S^2 + 2K_1 h_{\text{surf}} \Delta\theta} \quad (7.15b)$$

and h_a denotes the absolute value of the soil-water pressure head at which the air phase becomes discontinuous upon wetting. By defining

$$B_1 = (h_{\text{surf}} + h_a)\Delta\theta \quad \text{and} \quad B_2 = \frac{2}{S^2 + 2K_1 h_{\text{surf}} \Delta\theta} \quad (7.15c)$$

Equation 7.15a can be expressed by only four fitting parameters K_0 , K_1 , B_1 , and B_2 . The Clausnitzer et al. (1998) study concluded that both the PA and BA models described infiltration equally well, however, the BA model, while most advanced, was not as well suited to serve as a fitting model due to non-uniqueness problems caused by the larger number of fitting parameters.

7.3.3 Swartzendruber (SW) Model

Swartzendruber (1987) proposed an alternative series solution that is applicable and exact for all infiltration times, and also allows for surface ponding. Its starting point is similar to the GA approach, however, its derivation does not require a step function for the wetted soil profile. Its simplified form is a three-parameter infiltration equation (SW model)

$$I = K_1 t + \frac{S}{A_0} (1 - \exp(-A_0 t^{1/2})) \quad (7.16)$$

where A_0 is a fitting parameter of which its value depends on the surface water content, θ_1 . As $A_0 \rightarrow 0$, it reduces to a form of the Philip (1957b) model with K_1 as the coefficient of the linear term, and for which dI/dt approaches K_1 as $t \rightarrow \infty$. As for the GA model, the S-term can be corrected using Equation (7.11a) to account for ponded conditions.

7.3.4 Empirical Infiltration Equations

For most of these types of infiltration equations the fitting parameters do not have a physical meaning and are evaluated by fitting to experimental data only. However, in many cases, the specific form of the infiltration equation is physically intuitive. For example, the empirical infiltration equation by Horton (1940) is one of the most widely-used empirical infiltration equations. It considers infiltration as a natural “exhaustion process,” during which infiltration rate decreases exponentially with time from a finite initial value, $i_c|_{t=0} = (\alpha_1 + \alpha_2)$, to a final value, $\alpha_1 = K_1$. Accordingly, cumulative infiltration I (L) is predicted as a function of time t (HO model):

$$I = \alpha_1 t + \frac{\alpha_2}{\alpha_3} [1 - \exp(-\alpha_3 t)] \quad (7.17)$$

with the soil parameter $\alpha_3 > 0$, representing the decay of infiltration rate with time. In Equation 7.17, α_1 can be associated with the hydraulic conductivity (LT^{-1}) of the wetted soil portion, K_1 , for $t \rightarrow \infty$.

Another simple empirical infiltration equation is the Kostiakov (1932) model (KO):

$$i = at^{-b} \quad (7.18)$$

Clearly, this equation will not fit infiltration data at long times, as it predicts zero infiltration rate as $t \rightarrow \infty$. The value of a should be equal to the infiltration rate at $t = 1$, and $0 < b < 1$.

Mezencev (1948) proposed another infiltration model, and modified the KO model by including a linear term with a coefficient β_1 , so that $\beta_1 \rightarrow K_1$ for $t \rightarrow \infty$, provided $0 < \beta_3 < 1$ and $\beta_2 > 0$ (ME model)

$$I = \beta_1 t + \frac{\beta}{1 - \beta} t^{(1-\beta_3)} \quad (7.19)$$

Other models include the Soil Conservation Service (SCS) method and the Holtan solution (Kutilek and Nielsen, 1994; Espinoza, 1999). The only parameter needed for the SCS method (SCS, 1972) is the curve number (CN), which is available for most soil types in the United States, from which a soil water storage estimate is determined. The cumulative infiltration, I , is subsequently estimated from the measured precipitation and estimated runoff.

7.4 Infiltration Measurement Methods

Infiltration measurements can serve various purposes. In addition to characterizing infiltration, for example, to compare infiltration between different soil types, or to quantify macropore flow, infiltration is often measured to estimate the relevant soil hydraulic parameters from the fitting of the infiltration data to a specific infiltration model. This is generally known as inverse modeling. Infiltration is generally measured using one of three different methods: a sprinkler method, a ring infiltration method, or a permeameter method. The sprinkler method is mostly applied to determine the time of ponding for different water application rates, whereas the ring infiltrometer method is used when the infiltration capacity is needed. The permeameter method provides a way to measure infiltration across a small range of h -values ≤ 0 . A general review of all three methods was recently presented by Smettem and Smith in Smith et al. (2002), whereas a comparison of different infiltration devices using seven criteria was presented by Clothier (2001).

Rainfall sprinklers or rainfall simulators are also sprinkler infiltrometers, but they are typically used to study runoff and soil erosion (e.g., Morin et al., 1967). They mimic the rainfall characteristics (e.g., kinetic energy) of natural storms, specifically the rainfall rate, rainfall droplet size distribution, and drop velocity. Most of these devices measure infiltration by subtracting runoff from applied water. Using a range of water application rates, infiltration measurements can be used to determine the $i(I)$ curve for a specific soil type, with specific soil hydraulic properties such as K_s or S . Various design parameters for many developed rainfall simulators, specifically nozzle systems, were presented by Peterson and Bubenzier (1986). A portable and inexpensive simulator for infiltration measurements along hillslopes was developed by Battany and Grismer (2000). This low-pressure system used a hypodermic syringe needle system to form uniform droplets at rainfall intensities ranging from 20 to 90 mm/h.

Ring infiltrometers have historically been used to characterize soil infiltration by determining the infiltration capacity, i_c . A ring is carefully inserted in the soil so that water can be ponded over a known area. As a constant head is required, a constant water level is maintained by either manually adding water with a measuring stick in the infiltration to measure the depth of ponded water, by using a Mariotte tube, or by a valve connected to a float that closes at a predetermined water level. Measurements are usually continued until the infiltration rate is essentially constant. Water seepage around the infiltrometer is prevented by compaction of the soil around and outside of the infiltrometer. Multi-dimensional water flow under the ring is minimized by pushing the ring deeper into the soil, or by including an outer buffer

ring. In the latter case, the soil between the two concentric rings is ponded at the same depth as the inner ring, to minimize lateral flow directed radially outwards. The deviation from the assumed one-dimensionality depends on ring insertion depth, ring diameter, measurement time, and soil properties such as its hydraulic conductivity and the presence of the restricting soil layers. A sensitivity analysis on the diverging flow of infiltrometers was presented by Bouwer (1986) and Wu et al. (1997).

Permeameters are generally smaller than infiltrometers and allow easy control of the soil water pressure head at the soil surface. Generally, multi-dimensionality of flow must be taken into account, using Wooding's (1968) equation for steady flow (Q_∞ , L^3/T) from a shallow, circular surface pond of free water, or

$$Q_\infty = K_s \left(\pi r_0^2 + \frac{4r_0}{\alpha} \right) \quad (7.20a)$$

The first and second terms in parentheses denote the gravitational and capillary components of infiltration and α denotes the parameter in Gardner's unsaturated hydraulic conductivity function (1958)

$$K(h) = K_s \exp(\alpha h) \quad (7.20b)$$

In this model of the so-called Gardner soil, the macroscopic capillary length, λ_c , is equivalent to $1/\alpha$. The basic analysis for most permeameter methods relies on Wooding's solution. An extensive review of the use of permeameters was presented by Clothier (2001), including the tension infiltrometers and disk permeameters, by which the soil water pressure at water-entry is controlled by a bubble tower. Their use is relatively simple, and is based on analytical solutions of steady-state water flow. The permeameter method is economical in water use and portable. The soil hydraulic properties (S and K), in an inverse way, can be inferred from measurements using (7.1) both short and long-time observations, (7.2) disks with various radii, or (3) using multiple water pressure heads. Transient solutions of infiltration may be preferable, as it allows analysis of shorter infiltration times, so that the method is faster and likely will better satisfy the homogeneous soil assumption. Differences between one- and three-dimensional solutions for transient infiltration were analyzed by Haverkamp et al. (1994), Vandervaere et al. (2000), and Smith et al. (2002) from multi-dimensional numerical modeling analysis. These effects were reported to be small if gravity effects were included. Nowadays, permeameters are most often applied to estimate the soil's hydraulic characteristics in an inverse way, by fitting infiltration data to analytical solutions. The so-called mini-disk infiltrometer is available for simple and fast characterization of soil surface infiltration (<http://www.decagon.com/instruments/infilt.html>).

7.5 Spatial Variability Considerations

Although measurement scale is variable, all of the methods presented in Section 7.4 are small-scale measurements in the range of 0.2–1.0 m. This is relevant for irrigation purposes, especially for micro-irrigation applications. Yet, infiltration information is often needed for much larger spatial scales, at the pedon scale, hillslope, and watershed scale. Very little work has been done relating the infiltration process to measurement or support scale. Exceptions are the studies by Sisson and Wierenga (1981) and Haws et al. (2004), who measured steady-state infiltration at three spatial scales, ranging from 5 to 127 cm diameter infiltrometer rings. Their results showed that much of the larger-scale infiltration occurs through smaller-scale regions, and that the spatial variability of infiltration decreased as the measurement scale increased. Thus, in general we find that the process of infiltration might vary with spatial scale, and that larger spatial scales are required to estimate representative infiltration characteristics across a typical landscape.

Many field studies have dealt with the significant areal heterogeneity of soil hydraulic properties, and particularly that of the saturated hydraulic conductivity, K_s (Nielsen et al., 1973). The heterogeneity in K_s is recognized to have a major effect on unsaturated flow, and leading to significant variation in local infiltration. In general, accounting for areal heterogeneity leads to shorter ponding times and to a

more gradual decrease of the infiltration flux with time (Smith and Hebbert, 1979; Sivapalan and Wood, 1996). To characterize spatial variable infiltration rates, Sharma et al. (1980) measured infiltration with a double-ring infiltrometer at 26 sites in a 9.6 ha watershed. The infiltration data were fitted to the PH infiltration Equation 7.13a, and fitting parameters S and A were scaled to express their spatial variability and to describe the ensemble-average or composite infiltration curve of the watershed. A simpler but similar scaling technique for infiltration data was presented by Hopmans (1989), who measured transient infiltration at 50 sites along a 100 m transect. Data were fitted to both the PH and a modified KO model that includes an additional constant c as a second term in Equation 7.19. This paper showed that spatial variability of infiltration can be easily described by the probability density function of a single scaling parameter, to be used for applications in Monte Carlo simulation of watershed hydrology, as suggested for the first time by Peck et al. (1977). For application at the field scale, the so-called one point method was presented by Shepard et al. (1993) to estimate furrow-average infiltration parameters of PH Equation 7.13a, across a furrow-irrigated agricultural field. They used the volume-balance principle from advance time to the field end, inflow rate, and flow area measurements.

For modeling surface hydrology, by subtracting the infiltration rate, $i(t)$, from the rainfall rate, $r(t)$, it is possible to estimate spatial and temporal distributions of rainfall excess or runoff. The influence of spatially heterogeneity in rainfall and soil variability on runoff production was studied by Sivapalan and Wood (1986), from an analytical solution of infiltration and by making use of the IDA approximation. Statistical characteristics of ponding time and infiltration rate were presented for two cases, one with a spatially variable soil with a lognormal K_s distribution and uniform rainfall, and the other for a homogeneous soil with spatially variable rainfall. Among the various results, this study concluded the ensemble infiltration approach is biased for spatially variable soils. Their results also showed that the cumulative distribution of ponding times or proportion of ponded area is an excellent way of analyzing mean areal infiltration. Moreover, the spatial correction of infiltration rate is time-dependent and varies depending on the correlation lengths of rainfall and soil K_s . This study neglected the effects of surface water run-on, as caused by accumulated water upstream, running on to neighboring areas thereby contributing locally to infiltration. A quantitative analysis of soil variability effects on watershed hydraulic response that included surface water interactions such as run-on, was presented by Smith and Hebbert (1979), through analysis of the effects of deterministic changes of infiltration properties in the direction of surface water flow, using a kinematic watershed model. In a subsequent study by Woolhiser et al. (1996), it was clearly demonstrated that runoff hydrographs along a hillslope are significantly affected by spatial trends in the soil's saturated hydraulic conductivity. We expect to collect important new information can be collected by linking this interactive modeling approach with remote sensing and GIS tools. In addition to the soil's infiltration characteristics and its variability, surface runoff is controlled by spatial and temporal patterns in rainfall (Milly and Eagleson, 1988; Woolhiser and Goodrich, 1988). A detailed analysis and review of the control of spatially variable hydrologic properties on overland flow is presented by Govenderaju et al. (this issue).

Yet another concern regarding non-ideal infiltration, causing spatially variable infiltration at small spatial scales comes from the presence of water-repellent or hydrophobic soils. Since the 1980s much new research and findings have been presented, improving the understanding of the underlying physical processes and its relevance to soil water flow and water infiltration (DeBano, 2000; Wang et al., 2000).

7.6 Infiltration into Sealed or Crusted Soils

Seal formation is a complex phenomenon dominated by a wide variety of factors involving soil properties, rainfall characteristics, and local water flow conditions. Sealing involves the formation of a thin, dense layer of very low hydraulic conductivity at the soil surface. Soil surface sealing can have severe agricultural, hydrological, and environmental effects. It decreases the infiltration rate, thereby reducing plant-available water to the plant in the root zone, diminishing the natural recharge of aquifers, and increasing runoff and soil erosion. Two types of rainfall-induced soil seals can be identified (i) structural seals that are

directly related to rainfall through the impact of raindrops and sudden wetting, and (ii) depositional or sedimentary seals that are indirectly related to rainfall as it results from the settling of fine particles carried in suspension by the runoff in soil depressions. A recent review on concepts and modeling of rainfall-induced soil surface sealing was presented by Assouline (2004). The problem of the calculation of infiltration rate in crusted soils was first addressed by Hillel and Gardner (1969, 1970). They presumed that a sealed soil can be modeled as a uniform soil profile capped with a saturated thin layer of low permeability and they prescribed constant physical properties such as the saturated hydraulic conductivity. Their simplified solution was based on the Green and Ampt (1911) approach, and assumed constant water content (or suction) at the interface between the seal and the soil beneath. Variations and extensions of this basic approach included the simulation of infiltration with time-dependent seal hydraulic conductivity functions (Vandervaere et al., 1998).

Other studies such as those by Baumhardt et al. (1990) and Assouline and Mualem (1997) include the dynamic modeling of seal formation by assigning time-dependent physical properties to the seal that depend on soil and rainfall characteristics. The performance of the dynamic model of Assouline and Mualem (1997) is shown in Figure 7.4, where the predicted infiltration curves during seal formation under three different rainfall intensities, 20, 40, and 90 mm/h, are depicted. For the 90 mm/h rainfall case, the predicted infiltration rates can be compared to the one measured by Baumhardt (1985). The direct effect of the presence of the impeding seal layer at the soil surface is to reduce ponding time and infiltration rate during rainfall (compare the infiltration curves for the unsealed and the seal formation cases at the 40 mm/h rainfall intensity). As rainfall intensity increases, ponding time decreases, the infiltration rate curve decreases more rapidly, whereas the final infiltration rate increases. The latter can be caused by soil surface seal destruction after prolonged raindrop impact (Seginer and Morin, 1970) or reduction of seal thickness by erosion (Baumhardt, 1985). Other possible causes are the protection of the surface seal by the accumulated water layer on the soil surface (Schultz et al., 1985) as well as the reduction of the thickness of the saturated zone within the seal layer as the rainfall intensity increases (Mualem et al., 1993; Assouline and Mualem, 2001).

Two different approaches were followed in applying such a solution to flow processes in sealed soils or during seal formation. In one approach there is a gradual change in the soil water retention and hydraulic conductivity functions according to a prescribed exponential distribution of the seal bulk density such as done by Mualem and Assouline (1989). The other approach is to replace the nonuniform seal by a uniform equivalent layer, thereby generating a flow domain consisting of two uniform flow layers. The

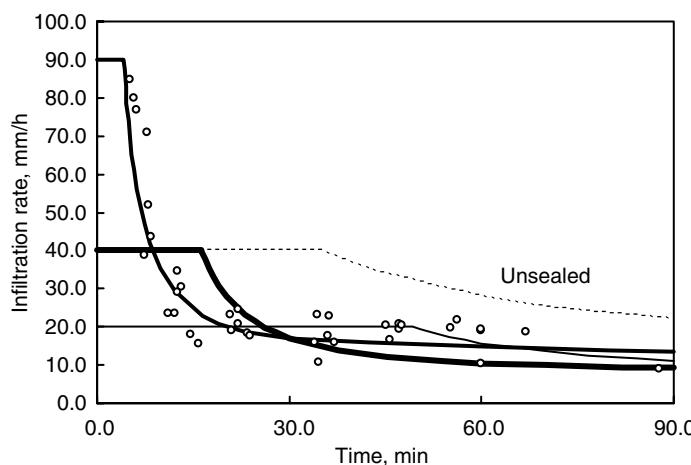


FIGURE 7.4 Infiltration curves predicted by the dynamic model of Assouline and Mualem (1997) for rainfall intensities of 20, 40 and 90 mm/h (solid lines), the corresponding infiltration curve for the unsealed soil case at 40 mm/h rainfall (dashed line); and the infiltration rates at 90 mm/h rainfall as measured by Baumhardt (1985).

effect of these two approaches was evaluated by Assouline and Mualem (2001). It appears that during seal formation, the approach adopted to represent the seal layer in the flow system has only a slight little effect on the infiltration curve. Still, a significant difference between the two approaches exists in terms of the soil water content distribution with depth, especially at the soil surface. However, in the case of a fully developed seal with constant properties, the approach adopted to represent the seal layer in the flow system has a much more significant effect on the infiltration curve. The use of an equivalent-uniform seal-layer causes a significant increase in the ponding time, assuming a uniform seal leads to overestimating the infiltration around ponding time. Therefore, approximating seals by equivalent uniform layers can induce errors in predicting the runoff hydrograph from bare fields or small bare watersheds where surface sealing predominates.

Under natural conditions, rainfall-infiltration relationships of bare soils at the field scale are mostly determined by the combined effect of soil seal formation and areal heterogeneity of soil hydraulic properties. The combined effect of these two factors on infiltration was presented by Assouline and Mualem (2002). Using the conceptual model of the non-uniform seal layer (Mualem and Assouline, 1989), this sensitivity study simulated the effect of areal variability of the initial soil properties on the seal properties. Subsequently, the dynamic model of seal formation (Assouline and Mualem, 1997) was applied to simulate the effect of the areal heterogeneity in the soil and the seal properties on the infiltration curve. The effect of areal heterogeneity on infiltration was estimated for a rainfall intensity of 40 mm/hr^{-1} and duration of 120 min. The infiltration curves for the unsealed case and during seal formation were simulated for a uniform field represented by mean values of the soil parameters and a heterogeneous one with all soil parameters being spatially distributed. The results are depicted in Figure 7.5, along with the corresponding measured infiltration rates in the packed soil columns (Baumhardt, 1985). For both the unsealed and the soil sealing case, the simulated infiltration curves corresponding to the uniform soil assumption fit the measured data, indicating that both the soil properties and the dynamic model represent the uniform conditions relatively well. For the unsealed field, areal heterogeneity of the soil properties has a considerable effect on infiltration. The ponding time is drastically reduced from 35 to 4 min, and the infiltration curve decreases more slowly. The maximal difference between the infiltration rates for the uniform and the heterogeneous fields is obtained at the time where ponding is estimated to occur for the uniform field case.

These trends are in agreement with the findings of Smith and Hebert (1979). In addition to the drastic reduction of the ponding time, areal heterogeneity in soil properties increases the steady infiltration rate reached after long rainfall. Similar trends are obtained when a seal develops at the soil surface, with some differences. The reduction of the ponding time when areal heterogeneity is accounted for is not as drastic as it is in the case of the unsealed soil. The interesting point here is that while the soil surface sealing significantly reduces the ponding time in the uniform field as expected, it does not affect the ponding time of the heterogeneous field. This is so as since part of the heterogeneous field has much lower hydraulic conductivity than that of the uniform field, and this generates the early runoff. This remains valid also when surface sealing is initiated.

Consequently, runoff will appear in the heterogeneous field about the same time, and is hardly affected by seal formation. Also, a higher quasi-steady infiltration rate is calculated in the case of the heterogeneous field during seal formation. Figure 7.5 shows that accounting for the areal heterogeneity in soil properties reduces the ponding time and increases the final steady infiltration rate, both for the unsealed case as well as under sealing conditions. Also, it seems that the relative effects of the areal heterogeneity on the ponding time and on the steady infiltration rate are both reduced when a seal develops at the soil surface. However, additional process occurring at the field scale, like run-on or depositional seal layers in surface depressions, can introduce some changes in these results.

7.7 Summary and Conclusions

Although important and seemingly simple, infiltration is a complicated process that is a function of many different soil properties and rainfall characteristics. In addition to rainfall intensity and duration and

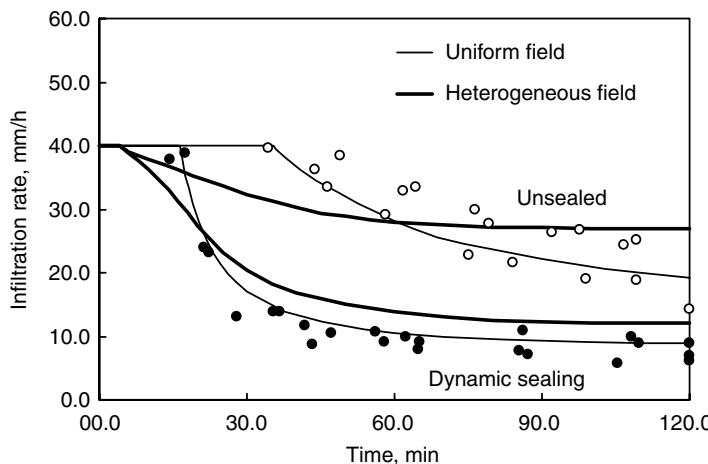


FIGURE 7.5 The infiltration curves simulated for a uniform (thin line) and a heterogeneous field (thick line) for the unsealed case and during seal formation, along with the corresponding measured infiltration rates in packed soil columns. From Baumhardt, R.L., 1985. The effect of rainstorm characteristics on soil sealing and filtration. Ph.D. Thesis, Mississippi State University, Mississippi State.

the soil physical factors such as soil water retention and hydraulic conductivity, infiltration is controlled by the initial water content, surface sealing and crusting, hydrophobicity, soil layering, and the ionic composition of the infiltrated water. Many relatively simple infiltration equations have been proposed and they are successfully used to characterize infiltration. Other physically-based analytical solutions have been presented that can potentially be used to predict infiltration. However, in practice this is difficult as the soils physical characteristics near the soil surface are time-dependent because of soil's structural changes and naturally high soil spatial variability. Alternatively, infiltration is often measured to estimate the relevant soil hydraulic parameters from the fitting of the infiltration data to a specific infiltration model by inverse modeling, such as by using permeameters.

Whereas most infiltration measurement techniques and infiltration models apply to relatively small spatial scales, infiltration information is often needed at the watershed and hill-slope scale. Yet, it has been shown that much of the larger scale infiltration occurs through smaller scale regions, for example because infiltration is largely controlled by spatial variations of the soil's saturated hydraulic conductivity. Infiltration and runoff predictions are even more complicated because of the surface soil's seal formation that is controlled by the impact of rain drops and type of vegetation cover. In general we find that the process of infiltration might vary with spatial scale, and that measurements at larger spatial scales are required to estimate representative infiltration characteristics across the hydrologic basin. We expect improved solutions to infiltration across scales from the field to the basin scale, using rapidly developing techniques such as remote sensing, GIS, and new measurement devices.

References

- Assouline, S., 2004. Rainfall-induced soil surface sealing: a critical review of observations, conceptual models and solutions. *Vadose Zone J.* 3: 570–591.
- Assouline, S. and Y. Mualem, 1997. Modeling the dynamics of seal formation and its effect on infiltration as related to soil and rainfall characteristics. *Water Resour. Res.* 33: 1527–1536.
- Assouline, S. and Y. Mualem, 2000. Modeling the dynamics of soil seal formation: Analysis of the effect of soil and rainfall properties, *Water Resources Research* 36: 2341–2349.
- Assouline, S. and Y. Mualem, 2001. Soil seal formation and its effect on infiltration: Uniform versus non-uniform seal approximation. *Water Resour. Res.* 37: 297–305.

- Assouline, S. and Y. Mualem, 2002. Infiltration during soil sealing: The effect of areal heterogeneity of soil hydraulic properties. *Water Resour. Res.* 38, doi:10.1029/2001WR001168.
- Barry D.A., J.-Y. Parlange, R. Haverkamp, and P.J. Ross, 1995. Infiltration under ponded conditions: 4. An explicit predictive infiltration formula. *Soil Sci.* 160: 8–17.
- Battany, M.C. and M.E. Grismer, 2000. Development of a portable field rainfall simulator for use in hillside vineyard runoff and erosion studies. *Hydraul. Process.* 14: 1119–1129.
- Baumhardt, R.L., 1985. The effect of rainstorm characteristics on soil sealing and infiltration. Ph.D. Thesis, Mississippi State University, Mississippi State.
- Baumhardt, R.L., M.J.M. Romkens, F.D. Whisler and J.-Y. Parlange, 1990. Modeling infiltration into a sealing soil. *Water Resour. Res.* 26: 2497–2505.
- Bouwer, H., 1986. Intake rate: Cylinder infiltrometer. In *Methods of Soil Analysis, Part 1*. Klute, A. (Ed.). Number 9 in the series Agronomy, American Society of Agronomy, Inc, Madison, WI, USA, pp.825–844.
- Bruce, R.R. and A. Klute, 1956. The measurement of soil moisture diffusivity. *Soil Sci. Soc. Am. Proc.* 20: 458–462.
- Brutsaert, W., 2005. Hydrology — An Introduction, Cambridge. University Press.
- Clausnitzer, V., J.W. Hopmans, and J.L. Starr, 1998. Parameter uncertainty analysis of common infiltration models. *Soil Sci. Soc. Am. J.* 62: 1477–1487.
- Clothier, B.E., 2001. Infiltration. In *Soil and Environmental Analysis, Physical Methods*, 2nd ed, Revised and Expanded. Smith, K.A. and Mullins, C.E. (Eds.). Marcel Dekker, Inc., New York, pp. 239–280.
- DeBano, L.F., 2000. Water repellency in soils: A historical Overview. *J. Hydrol.* 231: 4–32.
- Espinosa, R.D., 1999. Infiltration. In *The Handbook of Groundwater Engineering*. Delleur, J.W. (Ed.). CRC Press, Boca Raton, FL.
- Gardner, W.R., 1958. Some steady state solutions of unsaturated moisture flow equations with application to evaporation from a water table. *Soil Sci.* 85: 228–232.
- Govendaraju, R.S., C. Corradini, R. Morbidelli, and N. Nahar. Infiltration and run-on under spatially variable hydrologic properties. Chapter 8, this issue.
- Green, W.A. and G.A. Ampt, 1911. Studies on soils physics: 1. The flow of air and water through soils. *J. Agric. Sci. (Cambridge)* 4: 1–24.
- Haverkamp, R., Debionne, S., Viallet, R., Angulo-J.R., and de Condapp, D. 2006. Soil properties and moisture movement in the unsaturated zone. In *The Handbook of Groundwater Engineering*. Jacques W. Delleur (Ed.). CRC Press, Boca Raton, FL, 5-1 pp. 5–60.
- Haverkamp, R., J.-Y. Parlange, J.L. Starr, G. Schmitz, and C. Fuentes, 1990. Infiltration under ponded conditions: 3. A predictive equation based on physical parameters. *Soil Sci.* 149: 292–300.
- Haverkamp, R., P.J. Ross, K.R.J. Smettem, and J.-Y. Parlange, 1994. Three-dimensional analysis of infiltration from the disc infiltrometer. 2. Physically based infiltration equation. *Water Resour. Res.* 30: 2931–2935.
- Haws, N.W., C.W. Boast, Rao, P.S.C., Kladivko, E.J., and D.P. Franzmeier, 2004. Spatial variability and measurement scale of infiltration rate on an agricultural landscape. *Soil Sci. Soc. Am. J.* 68: 1818–1826.
- Hillel, D. 2004. *Introduction to Environmental Soil Physics*. Elsevier Academic Press, Amsterdam.
- Hillel, D. and W.R. Gardner, 1969. Steady infiltration into crust topped profiles. *Soil Sci.* 108: 137–142.
- Hillel, D. and W.R. Gardner, 1970. Transient infiltration into crust topped profiles. *Soil Sci.* 109: 69–76.
- Hopmans, J.W., 1989. Stochastic description of field-measured infiltration data. *Trans. ASAE* 32: 1987–1993.
- Horton, R.E., 1940. An approach towards a physical interpretation of infiltration capacity. *Soil Sci. Soc. Am. Proc.* 5: 399–417.
- Jury, W.A., W.R. Gardner, and W.H. Gardner, 1991. *Soil Physics*. John Wiley & Sons, Inc., New York.
- Kirkham, D. and W.L. Powers, 1972. *Advanced Soil Physics*. Wiley Interscience, John Wiley & Sons, Inc., New York. pp. 256–267.

- Kostiakov, A.N., 1932. On the dynamics of the coefficient of water percolation in soils and on the necessity of studying it from a dynamic point of view for purposes of amelioration. *Trans. Sixth Comm. Int. Soc. Soil Sci. A*: 17–21.
- Kutilek, M. and D.R. Nielsen, 1994. *Soil Hydrology*. GeoEcology textbook, Catena Verlag, 38162 Cremlingen-Destedt, Germany.
- Mein, R. and D.A. Farrel, 1974. Determination of wetting front suction in the Green–Ampt equation. *Soil Sci. Soc. Am. Proc.* 38: 872–876.
- Mezencev, V.J., 1948. Theory of formation of the surface runoff [Russian]. *Meteorologiya i Gidrologiya* 3: 33–40.
- Milly, P.C.D. and Eagleson, 1988. Effect of storm scale on runoff volume. *Water Resources Research* 24: 620–624.
- Morin, J., D. Goldberg, and I. Seginer, 1967. A rainfall simulator with a rotating disc. *Trans. ASAE* 10: 74–77.
- Mualem, Y. and S. Assouline, 1989. Modeling soil seal as a non-uniform layer. *Water Resour. Res.* 25: 2101–2108.
- Mualem, Y., S. Assouline, and D. Eltahan, 1993. Effect of rainfall induced soil seals on soil water regime: Wetting processes. *Water Resour. Res.* 29: 1651–1659.
- Nielsen, D.R., J.B. Biggar and K.T. Ehr, 1973. Spatial variability of field measured soil water properties, *Hilgardia* 42: 215–260.
- Parlange, J.-Y., I. Lisle, R.D. Braddock, and R.E. Smith, 1982. The three-parameter infiltration equation. *Soil Sci.* 133: 337–341.
- Parlange, J.-Y., R. Haverkamp, and J. Touma, 1985. Infiltration under ponded conditions: 1. Optimal analytical solution and comparison with experimental observations. *Soil Sci.* 139: 305–311.
- Peck, A.J., R.J. Luxmoore, and J.L. Stolzy, 1977. Effects of spatial variability of soil hydraulic properties in water budget modeling *Water Resour. Res.* 13: 348–354.
- Peterson, A.E. and G.D. Bubenzier, 1986. Intake rate: Sprinkler infiltrometer. In *Methods of Soil Analysis, Part 1*. Klute, A. (Ed.). Number 9 in the series Agronomy, American Society of Agronomy, Inc, Madison, WI, USA, pp. 845–870.
- Philip, J.R., 1957a. The theory of infiltration: 1. The infiltration equation and its solution. *Soil Sci.* 83: 345–357.
- Philip, J.R., 1957b. The theory of infiltration: 2. The profile at infinity. *Soil Sci.* 83: 435–448.
- Philip, J.R., 1957c. The theory of infiltration: 4. Sorptivity and algebraic infiltration equations. *Soil Sci.* 84: 257–264.
- Philip, J.R., 1969. Theory of infiltration. In *Advances in Hydroscience*. (Ed.), Vol. 5. Chow, V.T. Academic Press, New York, pp. 215–296.
- Philip, J.R., 1987. The infiltration joining problem. *Water Resour. Res.* 12: 2239–2245.
- Schultz, J.P., A.R. Jarrett, and J.R. Hoover, 1985. Detachment and splash of a cohesive soil by rainfall. *Trans. ASAE* 28: 1878–1884.
- Seginer, I. and J. Morin, 1970. A model of surface crusting in bare soils. *Water Resources Research* 6: 629–633.
- Sharma, M.L., G.A. Gander, and C.G. Hunt, 1980. Spatial variability of infiltration in a watershed. *J. Hydrol.* 45: 101–122.
- Shepard, J.S., W.W. Wallender, and J.W. Hopmans, 1993. One-point method for estimating furrow infiltration. *Trans. ASAE* 36: 395–404.
- Sisson, J.B. and P.J. Wierenga, 1981. Spatial variability of steady-state infiltration rates as a stochastic process. *Soil Sci. Soc. Am. J.* 45: 699–704.
- Sivapalan, M. and P.C.D. Milly, 1989. On the relationship between the time condensation approximation and the flux-concentration relation. *J. Hydrol.* 105: 357–367.
- Sivapalan, M. and E.F. Wood, 1996. Spatial heterogeneity and scale in the infiltration response of catchments. In *Scale Problems in Hydrology*. Gupta, V.K., Rodriguez-Iturbe, I., and Wood, E.F. (Eds.). Reidel, Hingham, MA. pp. 81–106.

- Skaggs, R.W., 1982. Infiltration. In *Hydrologic Modeling of Small Watersheds*. Haan, C.T., Johnson, H.P., and Brakensiek D.L. (Eds.) ASAE Monograph Nr. 5, ASAE, 2950 Niles Road, P.O. Box 410, St. Joseph, MI, 49085, pp. 121–166.
- Smith, R.E. and R.H.B. Hebbert, 1979. A Monte-Carlo analysis of the hydrologic effects of spatial variability of infiltration. *Water Resour. Res.* 15: 419–429.
- Smith, R.E., K.R.J. Smettem, P. Broadbridge, and D.A. Woolhiser, 2002. Infiltration theory for hydrologic applications. *Water Resources Monograph 15*, American Geophysical Union, Washington, DC.
- Soil Conservation Service-USDA, 1972. Estimation of direct runoff from storm rainfall, *National Engineering Handbook, Section 4—Hydrology*, pp. 10.1–10.24.
- Soil Science Society of America, 1996. Glossary of Soil Science Terms. Soil Science Society of America, 677 South Segoe Road, Madison, WI 53711.
- Stephens, D.B., 1996. *Vadose Zone Hydrology*. CRC Press Inc., Boca Raton, FL.
- Swartzendruber, D., 1987. A quasi-solution of Richards' equation for the downward infiltration of water into soil. *Water Resour. Res.* 23: 809–817.
- Vandervaere, J.-P., M. Vauclin, R. Haverkamp, C. Peugeot, J.-L. Thony, and M. Gilfedder, 1998. A simple model of infiltration into crusted soils, *Soil Sci.* 163: 9–21.
- Vandervaere, J.-P., M. Vauclin, and D.E. Elrick, 2000. Transient flow from tension infiltrometers: I. The two-parameter equation. *Soil Sci. Soc. Am. J.* 64: 1263–1272.
- Wang, Z., Q.J. Wu, L. Wu, C.J. Ritsema, L.W. Dekker, and J. Feyen, 2000. Effects of soil water repellency on infiltration rate and flow instability. *J. Hydrol.* 231: 265–276.
- Warrick, A.W., 2002. *Soil Physics Companion*. CRC Press, Boca Raton, FL.
- Woolhiser, D.A., and D.C. Goodrich, 1988. Effect of storm rainfall intensity patterns on surface runoff. *J. Hydrol.* 102: 335–354.
- Woolhiser, D.A., R.E. Smith, and J.-V. Giraldez, 1996. Effects of spatial variability of saturated hydraulic conductivity on Hortonian overland flow. *Water Resour. Res.* 32: 671–678.
- Wu, L., L. Pan, M.J. Robertson, and P.J. Shouse, 1997. Numerical evaluation of ring-infiltrometers under various soil conditions. *Soil Sci.* 162: 771–777.

Further Information

Although this chapter deals specifically with infiltration, the basic physical concepts are treated in the following texts of unsaturated water flow:

- Hillel (2004) is a generally-applicable textbook on soil physics, and includes comprehensive chapters on water flow in field soils, including infiltration and redistribution.
- Jury et al. (1991) provides additional material related to infiltration, including a treatment of various infiltration equations.
- Kutilek and Nielsen (1994) present an in-depth and extensive review of the various infiltration models and associated assumptions.
- Soil Science Society of America (1996) is a Glossary of Soil Science Terms, defining soil parameters and properties used in this chapter.
- Stephens (1996) introduces the concept of vadose zone hydrology, blending soil physics with hydrogeology.
- Warrick (2002) addresses the most significant topics of contemporary soil physics, written by leading scientists in the respective topics.

8

Infiltration and Run-On under Spatially Variable Hydrologic Properties

R.S. Govindaraju and
N. Nahar
Purdue University

C. Corradini
R. Morbidelli
University of Perugia, Italy

8.1	Introduction.....	8-1
8.2	Characterization of Spatial Variability	8-3
	Spatial Variability of Saturated Conductivity • Spatial Variability of Rainfall • Generation of Random Fields	
8.3	Numerical Simulation Results	8-5
	Local Scale Model • Soil Properties and Selected Cases • Expected Spatial Average Infiltration Rates • Discussion of Numerical Results • When Is Run-On Important?	
8.4	Theoretical Results.....	8-9
	Spatial Averages for Cases of Negligible Run-On • Empirical Averages for Cases of Appreciable Run almost On	
8.5	Conclusions and Recommendations	8-12
	Glossary.....	8-12
	References	8-13
	Further Information	8-15

8.1 Introduction

The processes governing water movement in natural fields are influenced by different factors such as rainfall intensity and duration, infiltration rates, catchment area, slope, and soil cover. Among these, a key role is played by infiltration, the process of water entering the soil (see Chapter 7). Hillslope hydrologic processes are strongly influenced by the patterns of infiltration rate through space and time (Freeze, 1980). The Hortonian mode of overland flow (Horton, 1933; Rubin and Steinhardt, 1963) is generated at a point on a hillslope when the rainfall intensity exceeds the infiltration rate, and surface ponding is initiated. A second mechanism, as described by Dunne (1978), requires surface saturation to occur from a rising water table over a shallow impeding layer; ponding and overland flow occur when no further soil moisture storage is available. The Horton mechanism is more common on upslope areas whereas the Dunne mechanism is generally observed in near-channel wetlands. While several other streamflow

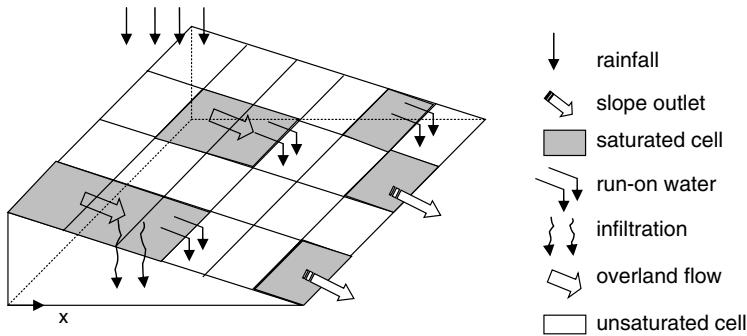


FIGURE 8.1 Schematic representation of the flow mechanisms that take place on a hillslope with the inclusion of run-on. (Adapted from Corradini et al., 1998, *J. Hydrol.*, 204:52–67.)

generation mechanisms exist (see Chow et al., 1988; Dingman, 2002), the focus of this chapter is on Hortonian runoff generation.

As mentioned in Chapter 6, natural watersheds exhibit spatial heterogeneity in topography, surface roughness, vegetation, and soil infiltration characteristics. Field measurements of soil hydraulic properties (Nielsen et al., 1973; Loague and Gander, 1990) and infiltration characteristics (Grah et al., 1983; Govindaraju et al., 1996) have shown that these properties exhibit a high degree of spatial variability, and several studies have demonstrated that this spatial variability has a significant effect on the generation of Hortonian runoff (Smith and Hebbert, 1979; Woolhiser and Goodrich, 1988; Corradini et al., 1998). Smith and Goodrich (2000) observed that ponding does not occur instantaneously over an entire field as conceptualized in many models because of spatial heterogeneity.

When ponded water is allowed to move in response to slope as overland flow, an important process called “run-on” manifests itself. The run-on process can be defined as the infiltration of surface water that, as it moves downslope, encounters areas where moisture deficit has not yet been satisfied. The infiltration rate here is larger than the runoff generated from upslope cells. In such cases, water available for infiltration includes rainfall and water supply from upslope areas. As seen in the schematic in Figure 8.1, run-on occurs when surface runoff from upstream cells infiltrates in pervious downstream cells. When the spatial variation of infiltration properties is taken into account, the run-on effect can be very significant in some cases. The goal of this chapter is to identify conditions when run-on is or is not significant, and to describe field-scale infiltration for both cases in a simple manner.

The importance of run-on has been documented in a few studies, but has largely been ignored in physically-based models. Run-on was first incorporated in an analysis of infiltration and Hortonian surface runoff by Smith and Hebbert (1979). They employed Monte Carlo simulations to analyze the effects of random distribution of soil infiltration properties on the hydrologic performance of catchments areas. The time for the initiation of hillslope hydrograph as well as runoff volume were strongly influenced by the inclusion of run-on even with deterministic variation in the soil saturated hydraulic conductivity, K_s . Saghafian and Julien (1995) carried out an extensive study of Hortonian overland flow with horizontal variability of saturated hydraulic conductivity K_s . They incorporated the run-on process utilizing two-dimensional finite-difference solutions of the diffusive wave approximation for surface flow with interactive infiltration described by the Green-Ampt equation. Again, hillslope hydrographs were markedly different when run-on effect was included from those when run-on was ignored. The relative placement of high hydraulic conductivity regions in the field was found to have a dramatic effect on hillslope hydrographs. Woolhiser et al. (1996) showed that runoff hydrographs are strongly affected by trends in hydraulic conductivity, particularly for small runoff events. They also demonstrated that run-on has major impacts on hydrograph peaks, runoff volumes, and time to peak. Corradini et al. (1998) observed the effects of run-on on Hortonian overland flow generated at the hillslope scale under conditions of horizontal heterogeneity of K_s . They concluded that the run-on process acted to reduce the sensitivity of the average overland flow hydrograph to the level of spatial correlation of K_s , which has found to be very

low, and that there exists a range of storm characteristics for which a deterministic formulation of effective K_s is acceptable. These conclusions were further confirmed by Nahar et al. (2004), where an extensive set of Monte Carlo simulations was used to examine infiltration and Hortonian runoff hydrographs both with and without run-on. They showed that the degree of spatial correlation in K_s also influenced the variance of field scale infiltration rate and overland flow, and this variance was higher for no run-on.

Even though several models describe infiltration into vertically-homogeneous soils “at a point” fairly accurately (see Chapter 7), upscaling of these models to the field scale has been complicated even in the absence of run-on. Most soils exhibit a high degree of spatial variability in their infiltration properties, particularly in the saturated hydraulic conductivity of the soil (Nielsen et al., 1973; Russo and Bresler, 1982; and others). The infiltration process, as described by these models, is highly nonlinear, so that “effective parameters” cannot be defined easily. The problem is aggravated further if macropore flow is present (Stillman et al., 2006). There appears to be a general consensus that the saturated hydraulic conductivity exhibits the maximum variability among the soil hydraulic properties that influence infiltration, and that this variability can be represented by a spatially-correlated lognormal random field for most practical purposes (Dagan and Bresler, 1983).

Researchers have followed a two-pronged approach to address the problem of describing field scale infiltration with or without run-on. The first approach has primarily relied on Monte-Carlo simulation techniques (Sharma and Seely, 1979; Smith and Hebbert, 1979; Maller and Sharma, 1981; Woolhiser and Goodrich, 1988), and are described in Chapter 14. Despite requiring significant computer effort, Monte-Carlo simulations are sometimes the only recourse when the governing equations are such as to render the mathematical analysis intractable as is the case when run-on is included. Such simulations can often be used to parameterize simple expressions for practical use (Smith and Goodrich, 2000; Corradini et al., 2002; Morbidelli et al., 2006).

Alternatively, many researchers have opted for providing simplified (analytical or semi-analytical) solutions. Maller and Sharma (1981, 1984) obtained distributions of time to surface ponding, infiltration rates, and cumulative infiltration volumes based on a known distribution of saturated hydraulic conductivity values. Hawkins and Cundy (1987) demonstrated a relationship between long term field scale infiltration and rainfall for spatially variable soils. Sivapalan and Wood (1986) used the two-term Philip equation to find the probabilistic distribution of ponding times, and subsequently presented approximate expressions for areal mean and variance of local infiltration rates. When macropore flow is important in determining infiltrated water, then simplified approaches may be utilized (Arabi et al., 2006; Stillman et al., 2006). Govindaraju et al. (2001) developed semi-analytical solutions to address the problem of field-scale infiltration under correlated hydraulic conductivity fields, and for time dependant rainfall. Corradini et al. (2002) presented an empirical analysis for field-scale infiltration and overland flow including the effect of run-on. Nahar et al. (2004) examined a wide range of cases of variability in K^* ($= K_s/r$, with r being rainfall intensity assumed uniform in space) and provided a framework for developing and solving averaged-equations for infiltration and overland flow when run-on can be neglected. Most of these studies utilized Monte-Carlo simulations to support their theoretical findings. However, none of these studies established definitive criteria as to when run-on could be neglected. In recent years, researchers (Morbidelli et al., 2006) have expressed interest in the problem of representing the effects of coupled spatial heterogeneity of both r and K_s on hydrological response.

8.2 Characterization of Spatial Variability

The generation of Horton runoff depends on both rainfall and soil infiltration properties, and the run-on process is influenced by spatial variation in these quantities. These are described briefly.

8.2.1 Spatial Variability of Saturated Conductivity

The saturated hydraulic conductivity, K_s , is treated as a spatial random field, while other soil parameters such as residual water content, θ_r , saturated water content, θ_s , and capillary head gradient, ψ (a parameter

in the Green-Ampt infiltration model), are assumed constant. The justification for this assumption is that K_s changes over several orders of magnitude for a single field, while the other parameters vary over a much narrower range (Dagan and Bresler, 1983). Previous statistical analysis (Russo and Bresler, 1982) indicated that the impact of the variability of other parameters is indeed limited compared to the influence of variability exhibited by K_s . The probability distribution of K_s has long been treated as a correlated log-normal random field based on several experimental studies (see e.g., Law, 1944; Bulnes, 1946; Warren and Price, 1961; Willardson and Hurst, 1965; McMillan, 1966; Rogowski, 1972; Nielsen et al., 1973; Baker and Bouma, 1976; Sharma et al., 1980, 1987; Burden and Selim, 1989; Govindaraju et al., 1996).

In line with previous studies, univariate probability density function of K_s is given by:

$$f_{K_s}(k) = \frac{1}{\sqrt{2\pi}\sigma_Y k} \exp\left[-\frac{(\ln k - \mu_Y)^2}{2\sigma_Y^2}\right] \quad (8.1)$$

where $Y = \ln(K_s)$ with mean μ_Y and variance σ_Y^2 . The isotropic autocovariance function of Y is assumed as

$$\Gamma_Y(\delta) = E[Y(x_1)Y(x_2)] = \sigma_Y^2 \exp\left(-\frac{\delta}{L_c}\right), \quad \delta = |x_1 - x_2| \quad (8.2)$$

where $E[\cdot]$ stands for the expectation operation, x_1 and x_2 denote vectorial spatial locations, δ denotes separation distance, and L_c determines the correlation length.

8.2.2 Spatial Variability of Rainfall

In addition to the heterogeneity of soil hydraulic properties, experimental studies show that rainfall also exhibits spatial variability at a larger scale than K_s (Goodrich et al., 1995; Krajewski et al., 2003). There is interaction between the effects of the variability of r and K_s on the areal-average infiltration and generation of surface runoff that needs to be accounted for in models. Studies which considered a coupled spatial variability of r and K_s were carried out by Wood et al. (1986), Castelli (1996), and Govindaraju et al. (2006). Govindaraju et al. (2006) and Morbidelli et al. (2006) argue that a uniform distribution of rainfall between the limits of r_{\min} and $r_{\min} + R$ allows for a reasonable representation of spatial variability of rainfall rate.

8.2.3 Generation of Random Fields

Several algorithms are available for generation of correlated random fields (see Chapter 14). Among these, the turning bands method (Mantoglou, 1987; Tompson et al., 1989) has been popular for the generation of random fields of saturated hydraulic conductivity for use in Monte-Carlo simulations. The turning bands method is a multidimensional random number generator designed to simulate spatially correlated random fields (Mejia and Rodriguez-Iturbe, 1974; Journel and Hujibregts, 1978; Delhomme, 1979; Smith and Freeze, 1979; Smith and Schwartz, 1981; Clifton and Neuman, 1982; Brooker, 1985; Miller and Borgman, 1985; Frankel and Clayton, 1986; Ababou et al., 1988). The method performs simulations along uni-dimensional lines instead of synthesizing the multidimensional field directly and the method is more efficient than, for instance, the LU decomposition algorithm (Gotway, 1994; Gotway and Rutherford, 1994). Tompson et al. (1989) provided new recommendations for accurate and efficient implementation of the three-dimensional algorithm.

Following Dagan and Bresler (1983), two-dimensional random fields over the soil surface (see Figure 8.1) of K_s and r may be also generated through the theory of similar media. Specific realizations of K_s are obtained by assuming a log-normal probability density function (pdf) and fixing the mean value $\langle K_s \rangle$ and the coefficient of variation $cv(K_s)$. Since previous studies (Corradini et al., 1998; Govindaraju et al., 2001) have shown that spatial correlation of K_s has a minor influence on average overland flow hydrograph and expected areal-average infiltration, K_s values may be assumed to be spatially independent

for studying the ensemble mean behavior. Because of the one-dimensional flow paths along the x -direction and in the absence of a spatial correlation structure for K_s and r along the y -direction, a one-dimensional generation would suffice in principle. However, a two-dimensional representation has been adopted in several studies (Govindaraju et al., 2006; Morbidelli et al., 2006) because it contributes to reduction in sampling statistical errors.

Similarly, by fixing the mean value of rainfall rate, $\langle r \rangle$, and the coefficient of variation, $cv(r)$, two-dimensional realizations of r corresponding to a uniform pdf may be carried out. Govindaraju et al. (2006) conducted a comparative analysis of the behavior of the expected areal-average infiltration $\langle I \rangle$ by adopting a uniform and a lognormal pdf for r and found that they produced substantially similar results, but the uniform pdf assumption leads to a simpler mathematical representation of $\langle I \rangle$.

8.3 Numerical Simulation Results

The role of run-on has been primarily evaluated numerically through simulations for lack of detailed experimentation. To clarify the role of run-on, a one-dimensional representation is often adopted. As shown in Figure 8.1, a typical field is represented by a plane of slope S_0 along the flow direction. For a rainfall event that generates Hortonian overland flow, the local infiltration is governed both by the rainfall and surface runoff generated from upstream areas. The soil is assumed to be vertically homogeneous, implying that subsurface horizontal flow components are neglected in comparison to vertical ones. This assumption is satisfied when the correlation scale in the horizontal plane is much larger than the vertical length scale associated with the head gradient (Dagan and Bresler, 1983; Chen et al., 1994a, 1994b).

8.3.1 Local Scale Model

The saturated hydraulic conductivity over the soil surface is represented by a stochastically homogeneous random field with a log-normal pdf characterized by the mean value $\langle K_s \rangle$ and the coefficient of variation $cv(K_s)$. Given a rainfall rate randomly varying in space with uniform pdf, with mean value $\langle r \rangle$ and coefficient of variation $cv(r)$, the local generation of surface runoff will be governed by infiltration of both rainfall and surface water from saturated areas running downslope over a pervious soil. For a given realization of K_s and r , local infiltration rate, q_0 , is formally expressed as:

$$q_0 = r + v_0 \quad \text{over unsaturated surface} \quad (8.3)$$

$$q_0 = f \quad \text{over saturated surface} \quad (8.4)$$

where v_0 is the discharge per unit surface generated upstream, with $v_0 = 0$ until the adjacent upstream area is unsaturated; f is the local infiltration capacity at time t computed by the Green–Ampt equation (see Chapter 7 for details):

$$f = K_s + \frac{K_s(\theta_s - \theta_i)\psi}{F}, \quad t > t_p \quad (8.5)$$

where F is the cumulative infiltration; θ_i and θ_s are the surface soil water content at $t = 0$ and saturation, respectively; t_p is the time to ponding, determined by the relation:

$$\int_0^{t_p} (r + v_0) dt = \frac{K_s(\theta_s - \theta_i)\psi}{r(t_p) - K_s} \quad (8.6)$$

where ψ is the capillary head gradient or suction head.

Surface runoff is routed over the plane by the kinematic wave approximation with flow resistance expressed by the Manning law (Singh, 1996):

$$\frac{\partial h}{\partial t} + \frac{S_0^{1/2}}{n} \frac{\partial h^m}{\partial x} = r - q_0 \quad (8.7)$$

with boundary and initial conditions given by:

$$h(0, t) = 0 \quad (8.8)$$

$$h(x, 0) = 0 \quad (8.9)$$

where x is the spatial coordinate increasing downslope, h is the flow depth, n is the Manning roughness coefficient and $m = 5/3$. Specifically, Equation 8.7 has to be combined with Equation 8.3 and Equation 8.4 in order to obtain, at each time, the values of h over saturated locations and v_0 over the unsaturated computational cells where Equation 8.7 becomes:

$$v_0 = - \frac{S_0^{1/2}}{n} \frac{\partial h^m}{\partial x} \quad (8.10)$$

which gives $v_0 > 0$ when $h > 0$ in the adjacent upstream cell.

8.3.2 Soil Properties and Selected Cases

The results in this chapter are mostly based on the work of Nahar et al. (2004), Govindaraju et al. (2006), and Morbidelli et al. (2006) who conducted an extensive set of numerical simulations over a wide range of cases by considering a 50 m long and 20 m wide plane with uniform slope $S_0 = 4\%$. Two soils representative of a clay loam soil, Soil B, and a sandy loam soil, Soil SL, were evaluated in detail by these authors (see Table 8.1). Conclusions are based on Monte-Carlo simulations of the coupled numerical solutions of the infiltration and the kinematic wave equations. These authors also carried out Monte-Carlo simulations for $v_0 = 0$ in order to quantify the relevance of the run-on process. These computations were performed using a Manning roughness coefficient $n = 0.15$ (metric units) and adopting a 0.5 m long and 2 m wide grid to meet stability requirements of the numerical procedure which are strictly governed by the representation of the run-on process.

It should be noted that the representation of the run-on process requires the coupled numerical solution of the infiltration and kinematic wave equations and incurs a significant computational burden.

TABLE 8.1 Hydraulic Data Used for the Study of Soils

Property	Soil B	Soil SL
Saturated water content, θ_s	0.3325	0.4120
Residual water content, θ_r	0.1225	0.0410
Initial water content, θ_i	0.1659	0.0613
Suction head, ψ [mm]	1116	360
Mean saturated hydraulic conductivity, $\langle K_s \rangle$ [mm h ⁻¹]	0.75	20.0
Coefficient of variation, $cv(K_s)$	0.3–1.0	0.3–1.0

Source: Adapted from Morbidelli et al., 2006, *Hydrol. Proc.*, John Wiley & Sons. Reproduced with permission.

8.3.3 Expected Spatial Average Infiltration Rates

With spatial variability in K_s and r , the areal-average infiltration rate, I , over the plane of area A is therefore partly controlled by the rainfall rate and run-on (unsaturated area A_1) and partly controlled by the soil (saturated area A_2):

$$I = \frac{1}{A} \iint_A q_0(x, y) dA = \frac{1}{A} \left[\iint_{A_1(t)} \left(r - \frac{S_0^{1/2}}{n} \frac{\partial h^m}{\partial x} \right) dA_1 + \iint_{A_2(t)} f dA_2 \right] \quad (8.11)$$

For each realization, the hillslope hydrograph $Q(t)$ is obtained through the spatial average of the product of velocity and flow depth at the hillslope outlet. Thus, using a variety of K_s and r realizations, the expected areal-average infiltration and hillslope response, $\langle Q(t) \rangle$, can be estimated by averaging I and Q over the ensemble of realizations.

8.3.4 Discussion of Numerical Results

Numerical simulations performed for the study soils revealed that the spatial variability of K_s and r greatly influence the ensemble-averaged infiltration rate and the production of surface runoff, the latter in terms of both the peak discharge and the water volume, especially for storms of short duration and low intensity. Figure 8.2 shows representative simulations of the these effects on outflow hydrographs, for a fixed value of $cv(K_s)$ combined with a wide range of $cv(r)$ values, applied to Soil B. As can be seen from this figure,

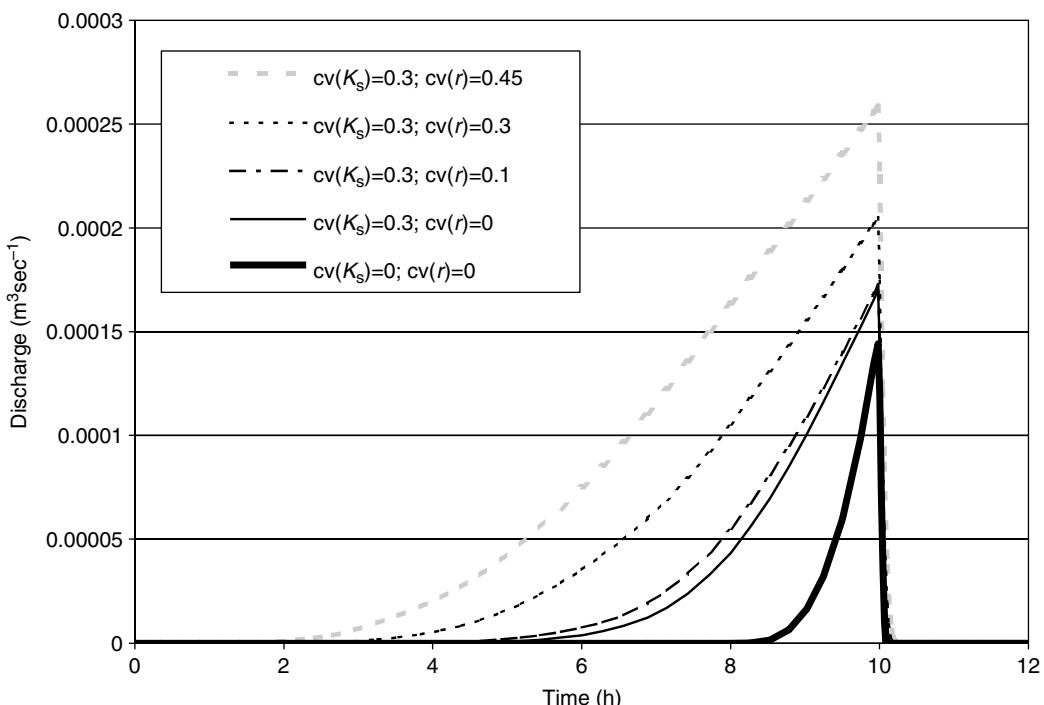


FIGURE 8.2 Surface outflow hydrograph for different combinations of the coefficients of variation of saturated hydraulic conductivity, $cv(K_s)$, and rainfall rate, $cv(r)$. Soil B with average value of rainfall rate $\langle r \rangle = 4.5 \text{ mm/h}$, and with rainfall duration 10 h. (Adapted from Morbidelli et al., 2006, *Hydrol. Proc.*, John Wiley & Sons. Reproduced with permission.)

the coefficient of variation of r plays a role very similar to that of K_s , and it is only when $\text{cv}(r) \ll \text{cv}(K_s)$ that the assumption of K_s as random variable and r as a uniform quantity appears reasonable.

Furthermore, when a larger rainfall rate with the same duration was used over each soil, the variation in $\langle Q \rangle$ with the level of heterogeneity in r and K_s was substantially reduced. However, when the rainfall duration decreased with increasing $\langle r \rangle$ the effects of spatial heterogeneity were found to be significant. Computations carried out over Soil B choosing appropriately the storm duration, t_d , revealed that the effects of spatial heterogeneity of r and K_s were in some way dependent on the rainfall rate. An overall analysis indicated that the role of spatial variability of r and K_s becomes less prominent for heavy storms and particularly when rainfall duration is much greater than the time to ponding t_{pa} obtained using $\langle K_s \rangle$ and $\langle r \rangle$. This implies that for small and moderate flood events in real watersheds the spatial variability of r and K_s need to be explicitly accounted for in models in order to avoid misleading results.

8.3.5 When Is Run-On Important?

To quantify the importance of run-on under various degrees of spatial variability, the Nash and Sutcliffe (1970) efficiency criterion can be used. The efficiency criterion E_f is expressed in this context as follows

$$E_f = \left[\sum_{i=1}^n (I_i - \bar{I})^2 - \sum_{i=1}^n (\hat{I}_i - I_i)^2 \right] \cdot \left[\sum_{i=1}^n (I_i - \bar{I})^2 \right]^{-1} \quad (8.12)$$

where I_i is the expected field-scale infiltration rate, \bar{I} is the mean value of field-scale quantity over n time steps considering run-on, and \hat{I}_i is the expected field-scale infiltration rate without run-on. According to this criterion, if the field-scale infiltration remains unchanged with or without run-on, then E_f attains its maximum value 1.0. As suggested by James and Burges (1982), an E_f value greater than 0.97 would imply that run-on may be ignored without incurring serious errors.

Figure 8.3 shows a contour map of E_f values over a range $0 \leq \mu_k^*, \sigma_k^* < 2$, to cover a wide range of field conditions. Here μ_k^* and σ_k^* are the mean and standard deviation, respectively, of $K^* = K_s/r$. Based

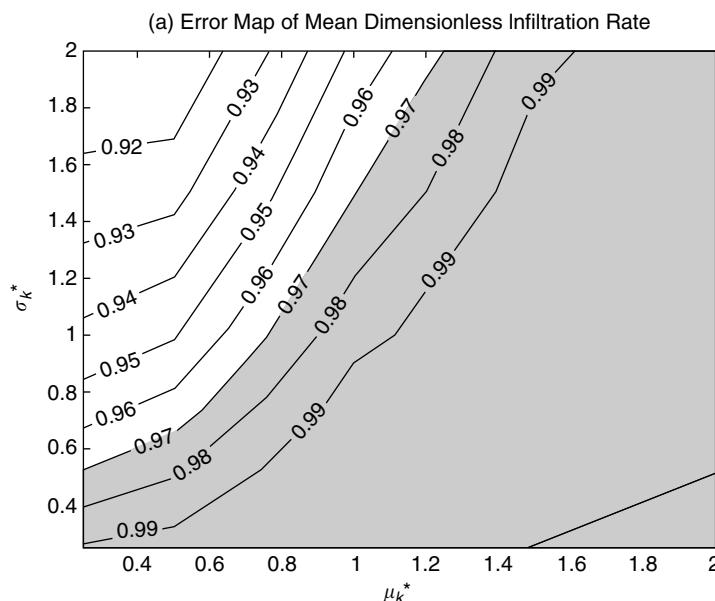


FIGURE 8.3 Error map of expected areal-average infiltration rate for various combinations of mean, μ_k^* , and standard deviation, σ_k^* , of the ratio between saturated hydraulic conductivity and rainfall rate.

on the threshold value of $E_f > 0.97$, the shaded regions in this figure shows when run-on process may be ignored. It is clear that if $\mu_k^* \gg \sigma_k^*$, then the nature of spatial variability is weak implying that run-on process will not be significant. The non-shaded region shows where influence of run-on process on the mean dimensionless infiltration rate is substantial and cannot be ignored. However, this figure is only valid for a spatially and temporally constant rainfall rate.

8.4 Theoretical Results

8.4.1 Spatial Averages for Cases of Negligible Run-On

By choosing the cumulative infiltration F in Equation 8.5 as the independent variable, that is, as a surrogate for time t , the average infiltration rate at a given F over the whole field can be expressed as:

$$I(F) = \frac{1}{N} \left\{ \sum_{i=1}^N r_i [1 - H(F - F_{pi})] + \sum_{i=1}^N f_i H(F - F_{pi}) \right\} \quad (8.13)$$

where the subscript i denotes the i th cell of Figure 8.1 out of a total of N cells, and the time to ponding and cumulative infiltration at ponding for the i th cell and for r invariant with time are given as:

$$t_{pi} = \frac{K_{si} \psi(\theta_s - \theta_i)}{r_i(r_i - K_{si})} \quad \text{and} \quad F_{pi} = t_{pi} r_i \quad (8.14)$$

The first and the second terms on the right-hand side of Equation 8.13 are the contributions from cells where ponding has not occurred and ponded cells, respectively, and $H(\cdot)$ is the Heaviside step function. Starting from Equation 8.13, as $N \rightarrow \infty$, the expected areal-average infiltration from infinite realizations can be written as:

$$\langle I \rangle = E[I(F)] = \int_0^\infty \int_{K_c}^\infty r f_r(r) f_{K_s}(K) dr dK + \int_0^\infty \int_0^{K_c} \left(1 + \frac{\psi \Delta \theta}{F} \right) K f_r(r) f_{K_s}(K) dr dK \quad (8.15)$$

where $f_r(r)$ is the pdf of r , $f_{K_s}(K)$ is the pdf of K_s , and

$$K_c = \frac{Fr}{\psi(\theta_s - \theta_i) + F} \quad (8.16)$$

Equation 8.15 was simplified by Govindaraju et al. (2006), with $\langle I \rangle$ approximated by $\langle I_n \rangle$ as

$$\begin{aligned} \langle I_n \rangle = E[I(F)] &= \frac{1}{2RF_c^2} \{ G_{K_s}[(r_{min} + R)F_c, 2] - G_{K_s}[r_{min}F_c, 2] \} \\ &\quad - \frac{r_{min}^2}{2R} \{ G_{K_s}[(r_{min} + R)F_c, 0] - G_{K_s}[r_{min}F_c, 0] \} \\ &\quad + \left(r_{min} + \frac{R}{2} \right) \{ 1 - G_{K_s}[(r_{min} + R)F_c, 0] \} \\ &\quad + \frac{1}{F_c} \left(\frac{r_{min} + R}{R} \right) \{ G_{K_s}[(r_{min} + R)F_c, 1] - G_{K_s}[(r_{min}F_c), 1] \} \\ &\quad - \frac{1}{F_c} \left(\frac{1}{RF_c} \right) \{ G_{K_s}[(r_{min} + R)F_c, 2] - G_{K_s}[(r_{min}F_c), 2] \} \\ &\quad + \frac{1}{F_c} G_{K_s}[r_{min}F_c, 1] \end{aligned} \quad (8.17)$$

where the function G_{K_s} is the truncated moment of the log-normal pdf of K_s defined as:

$$\begin{aligned} G_{K_s}(K_1, \xi) &= \int_0^{K_1} K^\xi f_K(K) dK \\ &= \exp\left(\xi \mu_Y + \frac{\sigma_Y^2 \xi^2}{2}\right) \left[1 - \frac{1}{2} \operatorname{erfc}\left(\frac{\ln K_1 - \mu_Y}{\sqrt{2}\sigma_Y} - \frac{\sigma_Y \xi}{\sqrt{2}}\right) \right] \end{aligned} \quad (8.18)$$

Here K_1 and ξ represent the first and the second argument of the G_{K_s} function in Equation 8.17 and $\operatorname{erfc}(\cdot)$ is the complimentary error function (see appendix in Chapter 3). Thus Equation 8.17 allows estimation of the expected areal-average infiltration rate as a function of cumulative infiltration. However, many hydrological applications require the estimate of $\langle I \rangle$ as a function of time t and therefore it is required to relate t to F . We start with the relation between these two quantities at the i th cell:

$$t_i = \frac{F_i}{r_i}, \quad \text{when } F_i < F_{pi} \text{ or } t_i < t_{pi} \quad (8.19a)$$

$$t_i = \frac{1}{K_{si}} \left[F_i + \psi \Delta \theta \ln \left(\frac{\psi \Delta \theta}{\psi \Delta \theta + F_i} \right) \right] + \psi \Delta \theta \sum_{j=1}^{\infty} \frac{K_{si}^j}{(j+1)r^{j+1}}, \quad \text{when } F_i > F_{pi} \text{ or } t_i > t_{pi} \quad (8.19b)$$

Equation 8.19b has been obtained by the classical relation of Green-Ampt for post-ponding conditions (Chow et al., 1988), modified by a series expansion of a log-term and simplification along the lines of Govindaraju et al. (2001). Since F takes on different values in each cell of Figure 8.1, an average time for the whole field may be computed as in Morbidelli et al. (2006):

$$\begin{aligned} E[t; r = \langle r \rangle] &= \frac{F}{\langle r \rangle} \{1 - G_{K_s}[\langle r \rangle F_c, 0]\} + \left[F + \psi \Delta \theta \ln \left(\frac{\psi \Delta \theta}{\psi \Delta \theta + F} \right) \right] \{G_{K_s}[\langle r \rangle F_c, -1]\} \\ &\quad + \psi \Delta \theta \sum_{j=1}^{\infty} \frac{1}{(j+1)\langle r \rangle^{j+1}} \{G_{K_s}[\langle r \rangle F_c, j]\} \end{aligned} \quad (8.20)$$

Using Equation 8.17 and Equation 8.20, one can find the expected areal-average infiltration rate and the corresponding expected time for a given cumulative infiltration depth. The series in the last term of Equation 8.20 converges very rapidly, and the first five terms are sufficient for most practical purposes.

8.4.2 Empirical Averages for Cases of Appreciable Run almost On

The inclusion of run-on process makes the derivation of analytical solutions almost impossible for field scale infiltration. After examining numerous results of field-scale infiltration, with and without run-on, Morbidelli et al. (2006) proposed the following empirical representation with parameters a , b , and c :

$$\langle I(t) \rangle \cong \langle I_n(t) \rangle + \langle r \rangle a t^b \exp(-c t^*) \quad (8.21)$$

where $\langle I_n(t) \rangle$ is obtained by ignoring run-on effect as given in Equation 8.17 and Equation 8.20, and $t^* = t/t_{pa}$, with t_{pa} defined as the ponding time obtained from average values of K_s and r in Equation 8.14. The quantities a , b , and c of Equation 8.21 were found to be well represented by the following relations:

$$a = 2.8[\operatorname{cv}(r) + \operatorname{cv}(K_s)]^{0.36} \quad (8.22)$$

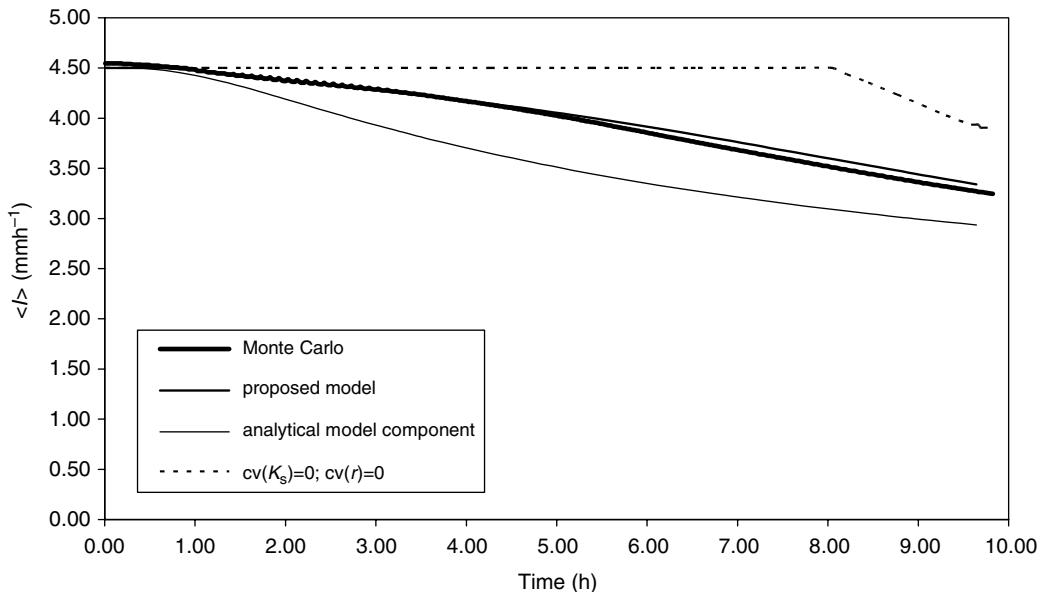


FIGURE 8.4 Comparison of expected field-scale infiltration rate, $\langle I \rangle$, derived by Monte Carlo simulations and obtained by the proposed model. Areal-average rainfall rate, $\langle r \rangle = 4.5 \text{ mm h}^{-1}$, invariant with time over Soil B of areal average saturated hydraulic conductivity $\langle K_s \rangle = 0.75 \text{ mm h}^{-1}$. The coefficients of variation were $cv(r) = 0.45$ and $cv(K_s) = 1.0$. (Adapted from Morbidelli et al., 2006, *Hydrol. Proc.*, John Wiley & Sons. Reproduced with permission.)

$$b = 5.35 - 6.32[cv(r)cv(K_s)] \quad (8.23)$$

$$c = 2.7 + 0.3 \left[\frac{\langle r \rangle / \langle K_s \rangle}{cv(r)cv(K_s)} \right]^{0.3} \quad (8.24)$$

Equations 8.22–8.24 require that both K_s and r be expressed in units of millimeters per hour. Morbidelli et al. (2006) note that much better accuracy is attained by the inclusion of the run-on component in Equation 8.21, thereby illustrating the importance of this component for representing field-scale infiltration. The model of Equation 8.21 was found to be very effective in simulating the expected cumulative infiltration over a wide range of spatial variabilities in K_s and r , with errors less than 5 and 2% over Soil B and SL, respectively. Figure 8.4 shows one example of expected field-scale infiltration rate behavior obtained from Monte-Carlo simulations vs. the models presented here. This figure is representative of cases where (i) the role of spatial variability of r and K_s on the production of effective rainfall is rather pronounced, as can be deduced by observing the difference between the curves from Monte Carlo simulations and from the use of averaged properties, and (ii) the run-on process produces a prominent effect on $\langle I \rangle$. The latter feature is illustrated by the difference between the Monte Carlo results and the analytical component model of Equation 8.17 and Equation 8.20 as first proposed in Govindaraju et al. (2006).

As can be seen here, the proposed model of Equation 8.21 shows a great improvement over Equation 8.17 and Equation 8.20 that ignore run-on. Based on numerous simulation experiments, Morbidelli et al. (2006) concluded that the proposed model Equation 8.21 is well structured because the two components interact appropriately under a wide range of conditions ranging from negligible run-on to those cases where run-on is very significant. For the sake of simplicity, only the case of rainfall rate invariant with time has been discussed here. This is usually not the case for real events. However, by following along the

lines developed by Govindaraju et al. (2001), it is possible to apply the semi-analytical model component for time-varying r as well.

8.5 Conclusions and Recommendations

Based on studies dealing with run-on effect, the following conclusions can be made. The influence of run-on on field-scale infiltration and runoff can be quite substantial in heterogeneous hillslopes, as has been demonstrated in numerous studies by using Monte Carlo simulations. These studies demonstrate that heterogeneity of K_s and r can have a substantial influence on the expected areal-average infiltration and on the surface hydrograph at the slope outlet. Existing infiltration models do not account appropriately for these effects when the run-on process cannot be neglected. The model described in this chapter overcomes this problem by providing accurate estimates of infiltration for the spatial variability in these two quantities. The use of the model does require knowledge of the first two truncated moments of the pdfs of K_s and r . Given the nature of K_s and r , Figure 8.3 may be used to first determine if run-on is important. When run-on is thought to be not important, then equations Equation 8.17 and Equation 8.20 may be used to describe field-scale infiltration. When run-on is taken into account, Equation 8.20 to 8.24 may be used. With the inclusion of run-on, expected field-scale infiltration increases, which in turn, has a significant influence on the hillslope hydrograph.

Glossary

Capillary head gradient: An estimate of the constant suction head that causes vertical water movement in the Green-Ampt model.

Correlation length: A measure of the extent of similarity in statistical properties in a heterogeneous quantity.

Diffusive wave model: An approximation for representing surface water movement where the momentum equation is simplified by equating the friction slope to the bed slope and a diffusive term.

Effective properties: Averaged values of spatially heterogeneous quantities that can be used in a model to describe average behavior at larger scale than the scale of heterogeneity.

Green-Ampt equation: A simplified physical model to represent local infiltration.

Infiltration: The process of water entering the soil through the soil surface.

Kinematic wave model: An approximation for representing surface water movement where the momentum equation is simplified by equating the friction slope to the bed slope.

Monte-Carlo simulations: A method of accounting for uncertainty by first conceptualizing the uncertain quantities to be particular realizations of a random process, and then generating numerous realizations of similar statistical characteristics to assess average behavior.

Overland flow: Movement of water over the land surface usually conceptualized as a thin sheet flow.

Ponding: The collection of the excess rain water on the soil surface.

Random field: A probabilistic model representation for describing quantities that vary in space in an uncertain fashion.

Residual water content: The water content below which water ceases to exist as a continuous phase in the soil.

Runoff: The accounting of water that moves on the soil surface.

Run-On: The process of allowing water from upstream areas to infiltrate into downstream regions where the moisture deficit has not been satisfied.

Saturated hydraulic conductivity: A measure of the ease with which water moves through a porous formation when all the pores are filled with water.

Saturated water content: The water content when all the pores of a porous medium are filled with water.

Storm duration: The time from the start to end of a rainfall event.

Time to ponding: The time it takes for rain water to start collecting on the surface from the start of a rainfall event.

Turning bands method: A technique used for generating random fields.

Upscaling: A method to represent heterogeneous quantities by their averaged or effective properties to be used in a constitutive equation.

Water content: The ratio of the volume of water to bulk volume in a soil.

References

- Ababou, R., McLaughlin, D.B., and Gelhar, L.W., 1988, Three Dimensional Flow in Random Porous Media. *Tech. Rep. 318, R. M. Parsons Lab.*, Dep. of Civ. Eng., Mass. Inst. of Technol., Cambridge, MA.
- Arabi, M., Stillman, J.S., and Govindaraju, R.S., 2006, A process-based transfer function approach to model tile drain hydrographs. In Press, *Hydrol. Process.* 13 pages.
- Baker, F.G. and Bouma, J., 1976, Variability of hydraulic conductivity in two subsurface horizons of two silt loam soils. *Soil Sci. Soc. Am. J.*, 40: 219–222.
- Brooker, P., 1985, Two-dimensional simulation by turning bands. *Math. Geol.*, 17: 81–90.
- Bulnes, A.C., 1946, An application of statistical methods to core analysis data of dolomitic limestone. *Trans. AIME*, 165: 223–240.
- Burden, D.S., and Selim, H.M., 1989, Correlation of spatially variable soil water retention for a surface soil. *Soil Sci.*, 148: 436–447.
- Castelli, F., 1996, A simplified stochastic model for infiltration into a heterogeneous soil forced by random precipitation. *Adv. Water Res.*, 19: 133–144.
- Chen, Z.Q., Govindaraju, R.S., and Kavvas, M.L., 1994a, Spatial averaging of unsaturated flow equations under infiltration conditions over areally heterogeneous fields. I. Development of models. *Water Resour. Res.*, 30: 523–533.
- Chen, Z.Q., Govindaraju, R.S., and Kavvas, M.L., 1994b, Spatial averaging of unsaturated flow equations under infiltration conditions over areally heterogeneous fields. II. Numerical simulations. *Water Resour. Res.*, 30: 535–548.
- Chow, V.T., Maidment, D.R., and Mays, L.W., 1988, *Applied Hydrology*, McGraw-Hill, New York.
- Clifton, P.M. and Neuman, S.P., 1982, Effects of kriging and inverse modeling on conditional simulation of the Avra Aquifer in Southern Arizona. *Water Resour. Res.*, 18: 1215–1234.
- Corradini, C., Govindaraju, R.S., and Morbidelli, R., 2002, Simplified modeling of areal average infiltration at hillslope scale. *Hydrol. Proc.*, 16: 1757–1770.
- Corradini, C., Morbidelli, R., and Melone, F., 1998, On the interaction between infiltration and Hortonian runoff. *J. Hydrol.*, 204: 52–67.
- Dagan, G. and Bresler, E., 1983, Unsaturated flow in spatially variable fields, 1, Derivation of models of infiltration and redistribution. *Water Resour. Res.*, 19: 413–420.
- Delhomme, J.P., 1979, Spatial variability and uncertainty in groundwater flow parameters: A geostatistical approach. *Water Resour. Res.*, 15: 269–280.
- Dingman, S.L., 2002, *Physical Hydrology*, Prentice Hall, Upper Saddle River, NJ.
- Dunne, T., 1978, Field studies of hillslope flow processes., In Kirkby, M.J., ed., *Hillslope Hydrology*, Wiley-Interscience, New York, pp. 227–293..
- Frankel, A. and Clayton, R., 1986, Finite difference simulations of seismic scattering: Implications for the propagation of short-period seismic waves in the crust and models of crustal heterogeneity. *J. Geophys. Res.*, 91: 6465–6489.
- Freeze, R.A., 1980, A stochastic-conceptual analysis of rainfall runoff processes on a hillslope. *Water Resour. Res.*, 16: 391–408.
- Goodrich D.C., Faurès, J.-M., Woolhiser, D.A., Lane, L.J., and Sorooshian, S., 1995, Measurement and analysis of small-scale convective storm rainfall variability, *J. Hydrol.*, 173: 283–308.

- Gotway, C.A., 1994, The use of conditional simulation in nuclear-waste-site performance assessments. *Technometrics*, 36: 129–144.
- Gotway, C.A. and Rutherford, B.M., 1994, Stochastic simulation for imaging spatial uncertainty, comparisons and evaluation of available algorithms, In Armstrong, M. and Dowd, P.A., eds., *Geostatistical Simulations: Proc. Geostatistical Simulation Workshop*, Fontainebleau, France, 27–28 May, 1993, Kluwer Acad. Publ. Co., Dordrecht, Netherlands, pp. 1–21.
- Govindaraju, R.S., Koelliker, J.K., Banks, M.K., and Schwab, A.P., 1996, Comparison of spatial variability of infiltration properties at two sites in the Konza Prairie. *J. Hydrol. Eng. ASCE*, 1: 131–138.
- Govindaraju, R.S., Morbidelli, R., and Corradini, C., 2001, Areal infiltration modeling over soils with spatially correlated hydraulic conductivities. *J. Hydrol. Eng. ASCE*, 6: 150–158.
- Govindaraju, R.S., Corradini, C., and Morbidelli, R., 2006, A semi-analytical model of expected areal-average infiltration under spatial heterogeneity of rainfall and soil saturated hydraulic conductivity. *J. Hydrol.*, 316: 184–194.
- Grah, O.J., Hawkins, R.H., and Cundy, T.W., 1983, Distribution of infiltration on a small watershed. *Proceedings of ASCE Irrigation and Drainage Speciality Conference on Advances in Irrigation and Drainage: Surviving External Pressures*, Eds. J. Barrelli, V.R. Hasfurther, and R.D. Burman, Jackson, Wyoming, pp. 44–54.
- Hawkins, R.H. and Cundy, T.W., 1987, Steady-state analysis of infiltration and overland flow for spatially-varied hillslopes. *Water Resour. Res.*, 32: 251–256.
- Horton, R.E., 1933, The role of infiltration in the hydrological cycle. *Trans. Am. Geophys. Union*, 12: 189–202.
- James, L.D. and Burges, S.J., 1982, Selection calibration and testing of hydrologic models, In Haan, C.T., Johnson, H.P., and Brakensiek, D.L., eds., *Hydrologic Modeling of Small Watersheds*, American Society of Agricultural Engineers, St. Joseph, MI, pp. 437–472.
- Journel, A.B. and Huijbregts, Ch.J., 1978, *Mining Geostatistics*. Academic, San Diego, CA.
- Krajewski, W.F., Ciach, G.J., and Habib, E., 2003, An analysis of small-scale rainfall variability in different climatic regimes. *Hydrol. Sci. J.*, 48: 151–162.
- Law, J., 1944, A statistical approach to the interstitial heterogeneity of sand reservoirs. *Trans. AIME*, 155: 202–222.
- Loague, K. and Gander, G.A., 1990, R-5 revisited 1. Spatial variability of infiltration on a small rangeland catchment. *Water Resour. Res.*, 26: 957–971.
- Maller, R.A. and Sharma, M.L., 1981, An analysis of areal infiltration considering spatial variability. *J. Hydrol.*, 52: 25–37.
- Maller, R.A. and Sharma, M.L., 1984, Aspects of rainfall excess from spatially varying hydrologic parameters. *J. Hydrol.*, 67: 115–127.
- Mantoglou, A., 1987, Digital simulation of multivariate two- and three dimensional stochastic processes with a spectral turning bands method. *Math. Geol.*, 19: 129–149.
- McMillan, W.D., 1966, Theoretical analysis of groundwater basin operations. *Water Resour. Center Contrib.*, 114: 167, University of California, Berkeley.
- Meija, J. and Rodriguez-Iturbe, I., 1974, On the synthesis of random fields from the spectrum: An application of the generation of hydrologic spatial processes. *Water Resour. Res.*, 10: 705–711.
- Miller, S. and Borgman, L., 1985, Spectral-type simulation of spatially correlated fracture set properties. *Math. Geol.*, 17: 41–52.
- Morbidelli, R., Corradini, C., and Govindaraju, R.S., 2006, A field-scale infiltration model accounting for spatial heterogeneity of rainfall and soil hydraulic properties. *Hydrol. Proc.*, 20: 1465–1481.
- Nahar, N., Govindaraju, R.S., Corradini, C., and Morbidelli, R., 2004, Role of run-on for describing field-scale infiltration and overland flow over spatially variable soils. *J. Hydrol.*, 286: 36–51.
- Nash, J.E. and Sutcliffe, J.V., 1970, River flow forecasting through conceptual models, 1, A discussion of principles. *J. Hydrol.*, 10: 282–290.
- Nielsen, D.R., Biggar, J.W., and Erh, K.T., 1973, Spatial variability of field measured soil-water properties. *Hilgardia*, 42: 215–259.

- Rogowski, A.S., 1972, Watershed physics: Soil variability criteria. *Water Resour. Res.*, 8: 1015–1023.
- Rubin, J. and Steinhhardt, R., 1963, Soil water relations during rain infiltration, 1. Theory. *Proc. Soil. Sci. Soc. Am.*, 27: 246–251.
- Russo, D. and Bresler, E., 1982, A univariate versus a multivariate parameter distribution in a stochastic-conceptual analysis of unsaturated flow. *Water Resour. Res.*, 18:483–488.
- Saghafian, B. and Julien, P.Y., 1995, Similarity in catchment response. 1. Stationary rainstorms. *Water Resour. Res.*, 31: 1533–1541.
- Sharma, M.L., Barron, R.J.W., and Fernie, M.S., 1987, Areal distribution of infiltration parameters and some soil physical properties in lateritic catchments. *J. Hydrol.*, 94: 109–127.
- Sharma, M.L., Gander, G.A., and Hunt C.G., 1980, Spatial variability of infiltration in a watershed. *J. Hydrol.*, 45: 101–122.
- Sharma, M.L. and Seely, E., 1979, Spatial variability and its effect on infiltration. *Proc. Hydrol. Water Resour. Symp.*, Institute of Engineering, Perth, Australia, pp. 69–73.
- Singh, V.P., 1996, *Kinematic Wave Modeling in Water Resources: Surface Water Hydrology*, John Wiley, New York, pp. 453–489.
- Sivapalan, M. and Wood, E.F., 1986, Spatial heterogeneity and scale in the infiltration response of catchments. In Reidel, D. ed., *Scale Problems in Hydrology*, Norwell, MA, pp. 81–106.
- Smith, L. and Freeze, R.A., 1979, Stochastic analysis of steady state groundwater flow in a bounded domain, 2. Two-dimensional simulations. *Water Resour. Res.*, 15: 1543–1559.
- Smith, L. and Schwartz, F.W., 1981, Mass transport, 2. Analysis of uncertainty in prediction. *Water Resour. Res.*, 17: 351–369.
- Smith, R.E. and Goodrich, D.C., 2000, Model for rainfall excess patterns on randomly heterogeneous area. *J. Hydrol. Eng. ASCE*, 5: 355–362.
- Smith, R.E. and Hebbert, R.H.B., 1979, A Monte Carlo analysis of the hydrologic effects of spatial variability of infiltration. *Water Resour. Res.*, 15: 419–429.
- Stillman, J.S., Haws, N.W., Govindaraju, R.S., and Rao, P.S.C., 2006, A model for transient flow to a subsurface tile drain under macropore-dominated flow conditions. *Journal of Hydrology*, 317: 49–62.
- Tompson, A.F.B., Ababou, R., and Gelhar, L.W., 1989, Implementation of the Three-dimensional turning bands random field generator. *Water Resour. Res.*, 25: 2227–2243.
- Warren, J.E. and Price, H.S., 1961, Flow in heterogeneous porous media. *Soc. Pet. Eng. J.*, 1: 153–169.
- Willardson, L.S. and Hurst, R.L., 1965, Sample size estimates in permeability studies. *J. Irrig. Drain. Div. Am. Soc. Civil Eng.*, 91(1RI): 1–9.
- Wood, E.F., Sivapalan, M., and Beven, K., 1986, Scale effects in infiltration and runoff production. *Proc. of the Symposium on Conjunctive Water Use*, IAHS Publication No. 156, Budapest.
- Woolhiser, D.A. and Goodrich, D.C., 1988, Effect of storm rainfall intensity patterns on surface runoff. *J. Hydrol.*, 102: 335–354.
- Woolhiser, D.A., Smith, R.E., and Giraldez, J.V., 1996, Effects of spatial variability of saturated hydraulic conductivity on Hortonian overland flow. *Water Resour. Res.*, 32: 671–678.

Further Information

The focus of this chapter has been on clarifying the role of run-on process in field-scale infiltration. Further, only natural variability of rainfall and infiltration properties have been accounted for by representing them as random fields. Despite its importance, there is not much information available on run-on, and how it influences field-scale behavior of water and associated contaminants. This topic has primarily been dealt with as a specialized research problem that appears in journal articles. The question of how man-made land use changes alter field- and watershed-scale infiltration has not been investigated in any detail. At the time of writing of this chapter, an ASCE task committee is being formulated to study and summarize the progress in some of these topics. The report from the task committee should contain more detailed information, but is still two to three years into the future.

9

Modeling the Movement of Water and Solute through Preferential Flow Paths and Fractures

Rony Wallach
The Hebrew University of Jerusalem
Tammo S. Steenhuis
Jean-Yves Parlange
Cornell University

9.1	Introduction.....	9-1
9.2	Dual-Porosity Models for a Sharp Change at the Fracture–Matrix Interface	9-3
	The Zeroth-Order Approximation • The First-Order Approximation	
9.3	A Dual-Porosity Model for Gradual Change in Matrix Structure and Concentration at the Fracture–Matrix Interface	9-9
	The Boundary Layer Effect on the Breakthrough Curve at the Fracture Outlet • The Transition Layer Effect on the Concentration Distribution in the Matrix • Modeling the Diffusion and Back-Diffusion — The Transition Layer Concept	
9.4	Dual-Porosity Models with a Distribution Layer at the Soil Surface.....	9-22
9.5	Multi-Pore Group Models.....	9-26
	Symbols	9-28
	Subscripts	9-29
	References	9-29
	Further Information	9-34

9.1 Introduction

Recent experimental research in soil structure and nonideal chemical transport is forcing scientists to refine their analysis of the movement and fate of contaminants in the soil and consider processes that have become important when small concentrations of highly toxic chemicals must be accurately accounted for to protect groundwater reserves from contamination (Stagnitti et al., 1995). The mobility of the

chemicals and contaminants in well-structured soils is affected by the continuity as well as the size of the pores. Networks of interconnected, highly conductive pathways that result from biological and geological activities, such as subsurface erosion, faults and fractures, shrink-swell cracks, animal burrows, wormholes, and decaying roots may be responsible for transmitting moisture, solutes, and contaminants to groundwater at faster velocities than those predicted by the theory based on the convective-dispersive equation (CDE) in which the water velocity is represented by the local average values (Bear, 1972; Sposito et al., 1986). In most cases, the volume of water within the preferential paths is much lower than the volume of the stagnant fluid, and only a small part of the solution (but very significant when the solute is toxic at low concentrations) may be moving within the well-defined preferential paths ahead of the main flow. For example, Smettem and Collis-George (1985) found that a single continuous macropore can conduct more water than a surrounding soil sample 10 cm in diameter. In structured soils, the convective-dispersive model gives unsatisfactory results under field conditions, and the interchange of fluids between the matrix and the fracture or large pore must be modeled explicitly (Steenhuis et al., 1990; Stagnitti et al., 1995; Parlange et al., 1996, 1998).

Contrary to widely held belief, even in relatively uniform sandy or water-repellent soil profiles, water does not necessarily take an average flow path and flow can take place through fingers caused by dynamic instabilities of the wetting front as was first documented by Hill and Parlange (1972). Further research has led to an understanding of many of the mechanisms involved in this phenomenon (Raats, 1973; Philip, 1975; Parlange and Hill, 1976; Hillel and Baker, 1988; Glass et al., 1989a, 1989b, 1989c, 1991; Kluitenberg and Horton, 1990; Selker et al., 1992a, 1992b, 1993; Liu et al., 1993, 1994a, 1994b; Sidle et al., 1998). The phenomena are not only restricted to the laboratory but can also be observed under field conditions. Fingered flow in a field was first noted in the Connecticut valley (Starr et al., 1978) and later by Glass et al. (1989a) on Long Island. Hendrickx and Dekker (1991), using dye-tracing techniques, found that only 10 to 20% of the vadose zone was actually involved in the transport of the dye for a water-repellent soil in Ouddorp (Netherlands). Also, Ritsema and Dekker (1993a, 1993b) and Dekker and Ritsema (1994) found that preferential flow was more the rule than the exception during the late summer and early fall in the Netherlands. Finally, Rice et al. (1991) in the midwestern U.S. found that solutes and herbicides travel at velocities 1.6 to 2.5 times faster than traditional water balance models of piston flow in sandy soils with little or no structure.

Finger properties can be derived from soil physical theory. Analysis of the wetting front instability suggests that the finger diameter, d , is given by

$$d = \frac{\kappa S_F^2}{K_F(\theta_F - \theta_i)} \left[\frac{1}{1 - \frac{Q}{K_F}} \right] \quad (9.1)$$

where θ_i is the water content ahead of the finger and θ_F , K_F , and S_F are the water content, conductivity, and sorptivity at the wettest spot in the finger, respectively, usually close to the tip. Q is the imposed flux at the surface ($Q \geq K_F$) and κ is a constant, which is equal to π in two dimensions (Parlange and Hill, 1976) or 4.8 in three dimensions (Glass et al., 1991). Hillel and Baker (1988) and Baker and Hillel (1990) suggest that at the wettest spot in the tip of the finger, the soil-water potential is at its water entry value. Indeed this seems to be a general result when the soil is initially dry. However, in subsequent infiltration events (using the finger paths created earlier), hysteresis is becoming a dominant factor and the finger becomes dryer (Liu et al., 1994b). The form of Equation 9.1 is physically quite intuitive. $S_F^2/(\theta_F - \theta_i)$ represents the sorptive properties of the soil, and the larger this term is, the wider the finger. K_F , on the other hand, represents the gravity effect, which drives the instability. The larger it is, the thinner the finger. Finally, increasing the imposed flux results in wider fingers. However, the main effect of the imposed flux is on the density of the fingers. The larger the Q , the closer the fingers are spaced. Hence, if Q approaches K_F , the whole space is necessary to carry the water entering the soil. At that point all the fingers merge so that $d \rightarrow \infty$.

Another mechanism for initiation of fingers in field soils is interbedded coarse layers in fine sand. Coarse layers concentrate the water from a large area into a finger-like structure. Kung (1990a, 1990b)

coined the term *funnel flow* for this phenomenon. He found that water and solutes flowed through less than 10% of the total soil matrix between depths of 3.0 and 3.6 m and less than 1% between 5.6 and 6 m. Funnel flow was also observed in Delaware (Boll et al., 1997) and in Massachusetts along the Connecticut River where the water flowed over coarse layers under low flux conditions and then broke through at higher fluxes (Steenhuis et al., 1998).

In contradiction to the fingers or funnel flow — formed by the dynamic instability in sandy soils, and where the shape, size, and distribution of the fingers depend on the imposed flux — the preferential flow paths in structured loamy or clayey soils and in fractured rocks are intrinsically defined as part of the soil and rock profile, respectively, with shape, size, and distribution independent of the fluids involved. In this chapter, we will concentrate on the structural defined flow paths. One of the main difficulties in modeling preferential flow in the “structural” large-pore system is the interaction of the preferentially moving chemicals with the surrounding matrix. To predict the solute distribution in structured soils, one ideally should know the exact spatial distribution, shape, size, and connectiveness of the preferential flow paths. This information is only available for a few soil cores (Heijmans et al., 1996), and not for field soils. Consequently, simplified models that overcome this lack of information are being used to interpret the vast amount of solute concentration data from field and laboratory experiments (Flühler et al., 1996).

In this chapter, we will give a survey of the various mathematical approaches for predicting solute breakthrough curves (BTC) in structured soils and fractured rocks with an emphasis on the effect of the interaction between the moving and the stagnant fluids in the surrounding matrix. Preferential flow in sandy and layered soils has been presented elsewhere (Glass and Nichol, 1996; Ritsema et al., 1996). Dual-porosity models with and without a distribution layer are discussed first, followed by multi-pore group models.

9.2 Dual-Porosity Models for a Sharp Change at the Fracture–Matrix Interface

Flow through cracks and fissures was initially studied in relation to the exploitation of nonhomogeneous groundwater and petroleum reservoirs. The naturally fractured reservoir can be taken as a collection of porous rock blocks, which are usually called the matrix, and an interconnected system of fracture planes. Most of the fluid resides in the matrix, where it moves very slowly, but the small amount of fluid rapidly flowing in the fractures can have a profound effect on the overall solute flow. A common approach to modeling this system mathematically is the dual-porosity concept, first presented by Barenblatt et al. (1960) and Warren and Root (1963). It is based on the assumption that naturally fractured reservoirs behave as two porous structures rather than one. The difference between the dual-porosity media and the usual porous media is that the “solid part” is in itself permeable and both matrix and fracture flow are defined at each point of the matrix. Thus, fluid flow is a combination of general macroscopic motions and takes place within the matrix blocks and also around them. The dual-porosity models have also been used to describe flow and transport in structured porous media where less permeable soil aggregates are analogous to the matrix (Passioura, 1971).

An intensive use in the dual-porosity concept has been made to model the non-ideal BTCs measured at the outlet of soil columns in the laboratory and soil cores from the field. Non-ideal BTCs are asymmetric and characterized by early breakthrough and an extensive tail that indicates that some part of the fluid is moving ahead of a front calculated by the CDE (Biggar and Nielsen, 1962; Krupp and Elrick, 1969; Starr and Parlange, 1977; van Genuchten and Wierenga, 1977; Rao et al., 1980; Nkedi-Kizza et al., 1983; Seyfried and Rao, 1987). The inter-aggregate porosity of aggregated soil and the dead-end pores in structured soils are considered as the low conductivity matrix, while the inter-aggregate porosity consists of the pathways where preferential flow takes place. The advection within the regions with the relatively low hydraulic conductivity is usually assumed zero, therefore, these domains act as sink/source components and the dual-porosity, dual velocity model becomes a mobile–immobile model.

The dissolved chemical transport in a single porosity model is usually described by the CDE (Biggar and Nielsen, 1967)

$$\frac{\partial(\theta_m c_m)}{\partial t} = \frac{\partial}{\partial z} \left[D_e \frac{\partial c_m}{\partial z} \right] - \frac{\partial(J_w c_m)}{\partial z} \quad (9.2)$$

where the subscript m stands for mobile, z is the direction of flow, and c is the solute concentration. D_e is the effective dispersion coefficient, combining the influence of diffusion in the dissolved phase with the long-time dispersion representation of the effect of small-scale convection on the mixing of solute about the center of motion (Jury et al., 1991). J_w is the water flux in the mobile paths. When water flow is steady and there is no chemical exchange with immobile porosity, Equation 9.2 reduces to

$$\frac{\partial c_m}{\partial t} = D \frac{\partial^2 c_m}{\partial z^2} - v \frac{\partial c_m}{\partial z} \quad (9.3)$$

where $D = D_e/\theta_m$ is the effective diffusion-dispersion coefficient and $v = J_w/\theta_m$ is the pore water velocity. Analytical solutions of Equation 9.3 for many initial and boundary conditions can be found in van Genuchten and Alves (1982).

In the mobile-immobile model, exchange takes place between the two regions and the chemical flux to/from the immobile porosity appears as a sink term in Equation 9.2. For the convenience of mathematical presentation, the equations which follow will be written for saturated mobile and immobile porosities, and Equation 9.3 becomes

$$\frac{\partial c_m}{\partial t} + \frac{\theta_{im}}{\theta_m} \frac{\partial \bar{c}_{im}}{\partial t} = D \frac{\partial^2 c_m}{\partial z^2} - v \frac{\partial c_m}{\partial z} \quad (9.4)$$

where the subscript im stands for immobile, $\theta_{im} = \theta - \theta_m$ is the immobile water content, and \bar{c}_{im} is the uniformly distributed solute concentration in the immobile porosity. The dissolved chemical exchange between the mobile and immobile porosities can be represented by two conceptually different mechanisms: a local equilibrium and a rate-limited exchange. The choice between the two is dictated by how fast the rate of exchange is compared with the transport rate in the mobile porosity. The exchange has been usually modeled by a diffusion equation based on Fick's law or by employing a first-order mass transfer approximation which, for particular cases, can be derived directly from a diffusion-based model (Rao et al., 1980; Parker and Valocchi, 1986; van Genuchten and Dalton, 1986). The rate-limited first-order exchange between the mobile and immobile porosities may be written as

$$\theta_{im} \frac{\partial \bar{c}_{im}}{\partial t} = \alpha [c_m - \bar{c}_{im}] \quad (9.5)$$

where α is the rate coefficient describing the mass transfer process between the mobile and immobile regions. A brief review of mass transfer models and their characteristics is presented in Brusseau and Rao (1989) and Brusseau et al. (1994).

Two time scales are involved in the mobile-immobile problem: the convective-dispersive transport in the mobile pores and the rate-limited dissolved chemical exchange between the two regions. For cases where the microscopic processes are rapid enough in relation to the fluids flow rate in the mobile pores, the local equilibrium assumption (LEA) was suggested to replace the rate-limited exchange (Equation 9.5) (Rubin, 1983). This replacement provides conceptual and mathematical simplifications. Models based on LEA have proven to be useful in a variety of solute transport applications, but there is an increasing body of experimental evidence indicating that the LEA may not be applicable under many conditions of interest in the study of solute transport (Jennings and Kirkner, 1984; Valocchi, 1985; Bahr and Rubin, 1987). These studies aimed to express conditions that are based on the flow and transport parameters under which the LEA is valid and provide small deviations from the rate-limited model. The identification

of these parameter values was based on comparing concentration profiles, which were calculated by the LEA model and solutions to the rate-limited model, obtained by different mathematical methods, for different scenarios and conditions. Wallach (1998) took advantage of the different time scales for the two rate-limited processes and defined a set of dimensionless variables that were substituted into the mass balance equations (Equation 9.4 and Equation 9.5). For cases when the characteristic time scale for the lateral solute exchange is faster than the convective transport time scale (as used to define the conditions for LEA), the ratio between the two time scales forms a small dimensionless parameter that multiplies the time derivative in the dimensionless version of Equation 9.5. The method of matched asymptotic expansions can then be used to form a uniform solution for this singular perturbations problem. The leading-order approximation, obtained when this parameter approaches zero, yields the solution obtained by the LEA and is valid only for the longer times ("outer" solution). A transition period exists near $t = 0$ within which the initial concentration equilibrates. The LEA validity, as well as the length of the transition period during which the initial mobile and immobile concentrations reach local equilibrium depends on how small the ratio between the two time scales is.

A closer look at the process of chemical transfer from the moving solute within the crack into the stagnant matrix brings forth the conclusion that the distribution of solute concentration within the moving solute varies only slightly across the crack, perpendicular to the flow, except at the interface with the matrix. The solute transferred across the interface is doing so by diffusion (Tang et al., 1981; Neretnieks et al., 1982; Neretnieks, 1983; Moreno and Rasmussen, 1986; Fujikawa and Fukui, 1990; Maloszewski and Zuber, 1990; and others). A mass balance equation is then written for each medium, and the mass continuity is preserved via the boundary condition at the interface between these two media. The mass balance equation for the preferential path (Tang et al., 1981) is

$$\frac{\partial c_m}{\partial t} - D_m \frac{\partial^2 c_m}{\partial z^2} + v \frac{\partial c_m}{\partial z} = -q \quad (9.6)$$

where D_m is the effective dispersion coefficient for the crack, and q ($\text{ML}^{-3}\text{T}^{-1}$) is the solute flux (concentration per unit time) from the crack to the matrix. The lateral transport in the matrix is described by

$$\frac{\partial c_{im}}{\partial t} = D_{im} \frac{\partial^2 c_{im}}{\partial x^2} \quad (9.7)$$

where x is the length in the matrix perpendicular to the flow direction in the crack, z , and D_{im} is the effective diffusion coefficient for the matrix. The boundary condition at the interface between the crack and the matrix ($x = 0$) is that local equilibrium is valid

$$c_m(z, t) = c_{im}(x = 0, z, t) \quad (9.8)$$

Solutions of this model reveal that the matrix diffusion can be viewed as a beneficial safety mechanism in soil contamination problems involving a concentrated source in fractured media. The dissolved chemicals are not spreading instantaneously throughout the matrix as assumed in the mobile-immobile models. The degree of spreading depends strongly on the difference in the time scales characterizing the dispersive-convective transport in the fracture and the transport by diffusion in the matrix. For example, in a system with large matrix porosity and low fracture fluid velocity the contamination of the matrix soil will be much less than in a system with low matrix porosity and high fracture fluid velocity.

Piquemal (1993) pointed out that it is erroneous, in general, to assume that the concentrations in the crack and the matrix are in local equilibrium. Developing the equations by the volume-averaging technique, he concluded that it applies only when the mean concentration of the flowing fluid is equal to the concentration in the immobile fluid at the boundary separating them. In many models where local equilibrium between the mobile and immobile regions is assumed, the solute flow is described by a

one-dimensional convective-dispersive model which *a priori* ignores the lateral concentration gradient in the mobile region. Therefore, the assumption is self-consistent in these models.

In the mobile-immobile (MIM) model, the preferential-flow paths are not physically determined but have much larger volume than the immobile volume and are assumed to be uniformly distributed throughout the soil profile. On the contrary, in the case of preferential path models the preferential paths are defined and the immobile volume is much larger than the volume of the preferential paths. The second difference is the dissolved chemical distribution in the matrix. Assuming a uniform concentration distribution with the lateral direction within the matrix (Equation 9.5 for the MIM model) rather than a diffusion front that propagates laterally as a function of $-t^{1/2}$ is equivalent to the assumption that concentration changes taking place at any point along the interface between the mobile and immobile domains equilibrate instantaneously with the whole stagnant solution in the horizontal direction. Since the chemical transfer between the mobile and immobile domains is controlled by the concentration difference (Equation 9.5), this assumption has a major effect on the values of α and D estimated by fitting the MIM model to measured BTCs. The fitting-routine of Parker and van Genuchten (1984) is an often-used tool to estimate the MIM model parameters from the measured BTCs. For field soils where dye tracers have shown that only few pores carry most of the solutes, the assumption of instantaneous equilibrium throughout the entire stagnant pore group is poor. The solute cannot move through the large lateral distances between the preferential paths by diffusion in the time frame of the transport in the preferential paths. Application of the MIM model to BTCs measured for soils with well-defined preferential paths provided poor predictions (Anderson and Bouma, 1977).

An important question for modeling flow and transport in preferential paths is whether the flow is laminar or turbulent. It has a great effect on the concentration distribution within the preferential path, especially in the direction perpendicular to the flow. Water flow in soil pores has traditionally been assumed laminar (following Poiseuille's law), but Chen and Wagenet (1992) argue that the flow is turbulent for macropores larger than 0.2 mm in diameter. Ela et al. (1991) measured rates of water flow into individual pores (size not given) ranging from 0 to 2550 ml/sec, and Wang et al. (1991, 1994) have measured inflow rates of 1 to 17 ml/sec into individual worm or ant holes (1.5 to 3.6 mm diameter). Reynolds numbers calculated for these open-channel pores and rates are out of range of laminar flow and into the transitional flow range (Chow, 1964). Turbulent flow theory was developed for large conduits, and laminar flow theory was developed for capillary-size pores. Large macropores are in between these sizes and may be only partially filled (Beven and Germann, 1981); thus, one may argue that there is no reason to assume one or the other type of flow.

Wallach and Steenhuis (1998) proposed a model for such field soils with well-defined preferential paths. The model follows the main assumptions made for the MIM model, that is, the velocity of the mobile fluid in the well-defined preferential paths is constant and the dissolved chemical exchange between the saturated stagnant pore group and the mobile pore group is a rate-limiting process. For the sake of simplicity, the dissolved chemical transport in the crack is modeled by the kinematic-wave approach rather than by the convective-dispersive equation, and analytic solutions are obtained. The model takes advantage of the characteristic time scales for chemical transfer through the boundary layer along the interface between the mobile and immobile domains being faster than the lateral transport of the chemicals within the stagnant porosity by diffusion. This assumption is expressed mathematically as multiplying the immobile concentration in the mass balance equations by a small dimensionless parameter, $\varepsilon' \ll 1$. The mass balance equation for the preferential paths becomes (Wallach and Steenhuis, 1998)

$$\frac{\partial c_m}{\partial t} + v \frac{\partial c_m}{\partial z} = -k_m [c_m - \varepsilon' c_{im}] \quad (9.9)$$

where $k_m = k/\theta_m$. The boundary condition at the inlet to the preferential path is

$$c_m(0, t) = c_0(t) \quad (9.10)$$

Postulating that the concentration within the matrix is not laterally uniform and that it varies significantly only within a short lateral distance from the interface during the travel time of the mobile fluid to the outlet, the solute concentration is assumed to be uniformly distributed only within a thin matrix layer, θ_{im} (defined as the pore volume in this layer per sample volume), located at the interface between the two pore groups. The mass balance equation for this matrix layer is

$$\frac{\partial c_{im}}{\partial t} = k_{im}[c_m - \varepsilon' c_{im}] \quad (9.11)$$

where $k_{im} = k/\theta_{im}$.

Equation 9.9 and Equation 9.11 form a regular perturbations problem whose solution can be expressed by the series

$$c_m(z, t) = \sum_{i=0}^n (\varepsilon')^i A_i = A_0 + \varepsilon' A_1 + (\varepsilon') A_2 + \dots \quad (9.12)$$

$$c_{im}(z, t) = \sum_{i=0}^n (\varepsilon')^i B_i = B_0 + \varepsilon' B_1 + (\varepsilon') B_2 + \dots \quad (9.13)$$

Substituting Equations 9.12 and Equation 9.13 into Equations 9.9 to 9.11 and equating terms that are multiplied by equal powers of ε' , sequences of equations for the different orders of approximations are obtained.

9.2.1 The Zeroth-Order Approximation

This approximation, known also as the leading-order approximation, is obtained by equating terms multiplied by $(\varepsilon')^0$. The mass balance equations for the preferential path and matrix solutes are

$$\frac{\partial A_0}{\partial t} + \nu \frac{\partial A_0}{\partial z} = -\frac{k}{\theta_m} A_0 \quad (9.14)$$

$$\frac{\partial B_0}{\partial t} = \frac{k}{\theta_m} A_0 \quad (9.15)$$

The right-hand side of Equation 9.14 expresses the chemical transfer from the preferential paths to the matrix when the back transfer due to the accumulation of chemicals in the matrix is ignored. As a result, Equation 9.14 and Equation 9.15 include one dependent variable each and can be solved independently, as opposed to Equation 9.4 and Equation 9.7, which contain two dependent variables and, therefore, should be solved simultaneously. The higher-order approximations in the following will account for the dissolved chemical flux from the matrix.

This problem is solved by applying the Laplace transform with respect to time. The transformed mobile concentration is represented as

$$a_i = \int_0^\infty A_i e^{-st} dt; \quad b_i = \int_0^\infty B_i e^{-st} dt \quad (9.16)$$

where s is the transformation variable. Application of Equation 9.16 to Equation 9.14 and Equation 9.15 leads to a pair of ordinary differential equations with a boundary condition in the Laplace domain.

The equation for the preferential path with zero initial concentration is

$$sa_0 + \nu \frac{da_0}{dz} = -\frac{k}{\theta_m} a_0 \quad (9.17)$$

and its boundary condition at the soil surface in the Laplace domain is

$$\bar{c}_0(s) = \frac{c_i}{s} \quad (9.18)$$

The solution of Equation 9.17 with Equation 9.18 in the Laplace domain is

$$a_0(z, s) = \frac{c_i}{s} \exp \left[-\frac{kz}{\theta_m \nu} - \frac{sz}{\nu} \right] \quad (9.19)$$

The solution for A_0 in the time domain is obtained through the application of the inverse Laplace transform to Equation 9.19

$$A_0(z, t) = c_i \exp \left[-\frac{kz}{\theta_m \nu} \right] \cdot u \left(t - \frac{z}{\nu} \right) \quad (9.20)$$

here $u(t)$ is the unit step function. The leading solution for the mobile domain is of an advancing sharp front with a simultaneous transfer of dissolved chemicals to the active matrix layer. The chemical transfer depends only on the mobile concentration and its concentration at any point z as a function of the travel time from the inlet to this point.

Applying the Laplace transform Equation 9.16 to Equation 9.15 with initial zero immobile concentration and substituting $a_0(z, s)$ from Equation 9.19 gives

$$b_0(z, s) = \frac{k}{\theta_{im}} \frac{c_i}{s^2} \exp \left[-\frac{kz}{\theta_m \nu} - \frac{sz}{\nu} \right] \quad (9.21)$$

whose inverse Laplace transform gives the zeroth-order approximation of the chemical distribution in the matrix. As the local concentration in the matrix depends on the local chemical concentration in the preferential path, it also varies with time and depth

$$B_0(z, t) = \frac{c_i k}{\theta_2} \exp \left[-\frac{kz}{\theta_m \nu} \right] \left\{ t - \frac{z}{\nu} \right\} \cdot u \left(t - \frac{z}{\nu} \right) \quad (9.22)$$

9.2.2 The First-Order Approximation

The first-order approximation is obtained by equating terms multiplied by ϵ' when Equation 9.12 and Equation 9.13 are substituted into Equation 9.9 to Equation 9.11.

The mass balance equations for the preferential path and matrix solutions are

$$\frac{\partial A_1}{\partial t} + \nu \frac{\partial A_1}{\partial z} = -\frac{k}{\theta_m} (A_1 - B_0) \quad (9.23a)$$

$$\frac{\partial B_1}{\partial t} = \frac{k}{\theta_{im}} (A_1 - B_0) \quad (9.23b)$$

Applying the Laplace transform Equation 9.16 to Equation 9.23 gives

$$a_1(z, s) = \frac{k^2}{\theta_m \theta_{im}} \frac{c_i}{s^2} \exp \left[-\frac{kz}{\theta_m v} - \frac{sz}{v} \right] \quad (9.24)$$

whose inverse Laplace transform is

$$A_1(z, t) = \frac{c_i k^2}{\theta_m \theta_{im}} \exp \left[-\frac{kz}{\theta_m v} \right] \frac{z}{v} \left[t - \frac{z}{v} \right] \cdot u \left(t - \frac{z}{v} \right) \quad (9.25)$$

Higher-order approximations can be similarly obtained and solved, but the first two terms in the asymptotic expansion provide a sufficiently good approximation. The breakthrough concentration from a soil column of length L is obtained by substituting Equation 9.20 and Equation 9.25 into Equation 9.12 after replacing z by L , as

$$c_1(L, t) = c_i \exp \left[-\frac{kL}{\theta_m v} \right] \left\{ 1 + \varepsilon' \frac{c_i k^2}{\theta_m \theta_{im}} \frac{L}{v} \left[t - \frac{L}{v} \right] \right\} \cdot u \left(t - \frac{L}{v} \right) \quad (9.26)$$

The shape of the BTC predicted by Equation 9.26 is similar to BTCs measured for initially drained undisturbed soil materials with well-defined preferential paths (Anderson and Bouma, 1977; White et al., 1984, 1986; Kluitenberg and Horton, 1990). These BTCs are characterized by a breakthrough at very early times followed by a sharp concentration increase afterward when the chloride concentration increases rapidly from 0 to a higher value after which the concentration increases at a lower rate until the initial concentration is reached. Note that the sharp increase in concentration immediately after the initial breakthrough and sharp break later on is due to the elimination of the diffusion-dispersion term from the mass balance equation for the preferential paths.

Equation 9.26 was fitted to BTCs that were measured by Anderson and Bouma (1977) at the bottom of saturated columns filled with structured soils that were initially drained. The measured BTCs are taken from figure 3 in Anderson and Bouma (1977) and are shown together with the predicted BTCs in Figure 9.1. After determining the velocity v by dividing the column length by the breakthrough time, k is estimated by the concentration where the BTC changes its shape from the vertical increase to a moderate increase, and ε is determined from the slope at which the BTC reaches the initial concentration. The estimated k is close to the values calculated independently for this BTC by Skopp et al. (1981). Equation 9.26 fits well the measured BTCs (Figure 9.1).

9.3 A Dual-Porosity Model for Gradual Change in Matrix Structure and Concentration at the Fracture–Matrix Interface

The widely used assumption in the diffusion models in the previous section was that concentration over crack width is constant and equal to the matrix concentration at the interface between the two domains (Equation 9.10). In the following we focus on the interface between the fracture and the matrix and postulate that there is a boundary layer at the flowing and stagnant fluids (fracture and matrix) and that the matrix porosity is not uniform in the lateral direction (perpendicular to the fracture axis). The sharp porosity decrease at the vicinity of the fracture–matrix interface is denoted herein as “transition layer.” In the following the effect of the boundary and transition layers at the matrix–fracture interface on the solute exchange dynamics between the fracture and matrix is demonstrated by mathematical models and measured data.

The matrix–fracture system is shown on the right-hand side of Figure 9.2. The “matrix” is composed from a thin transition layer at the interface between the matrix and the fracture and the “real” elsewhere. The transition layer has diffusion properties different from those of the rest of the matrix. A close look at

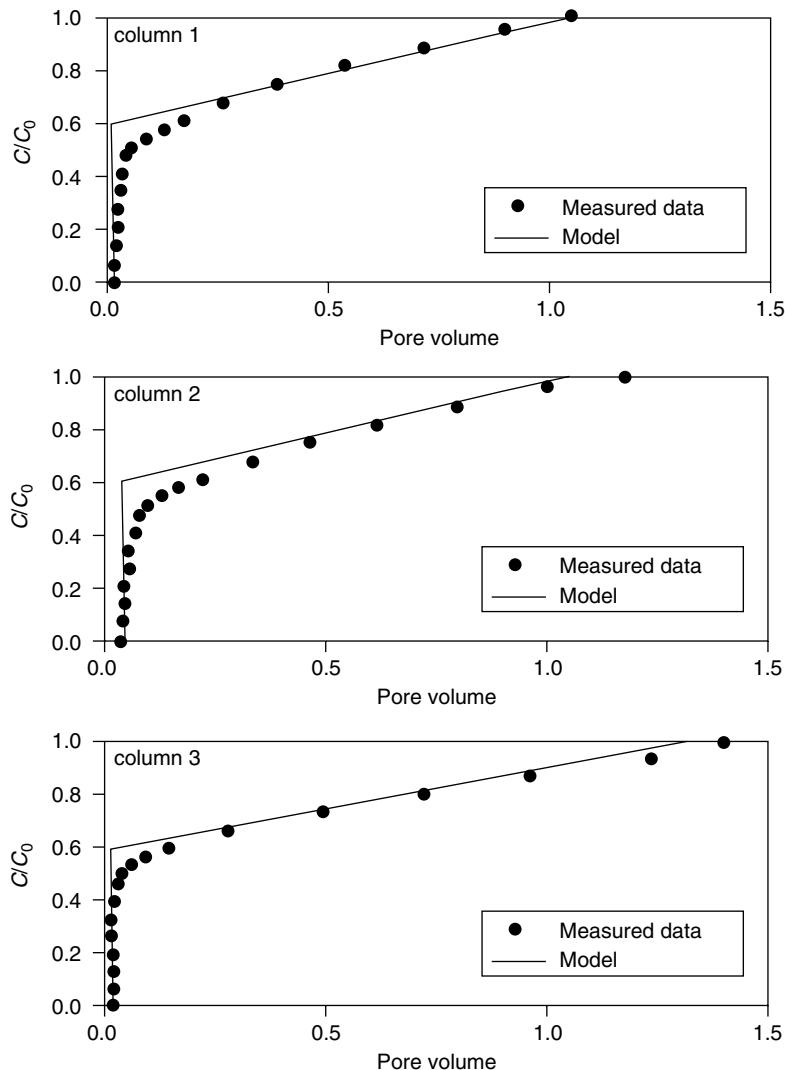


FIGURE 9.1 Measured (Data from Anderson, J. L. and Bouma, J. 1977. *Soil Sci. Soc. Am. J.* 41: 413–418.) and predicted BTCs at the outlet of three initially drained columns.

the fracture–matrix interface on the left-hand side of Figure 9.2 reveals the concentration variation along the fracture, boundary and transition layers, and matrix continuum.

The following mathematical model describes the convection in the fracture and lateral diffusion through the boundary and transition layers into the matrix. We assume that fluid flow through the crack is steady, one-dimensional, and isothermal. The concentration in the fracture is assumed to be laterally uniform beyond the boundary layer. Following the model presented by Wallach and Parlange (1998, 2000), the dissolved chemical transport in the crack is modeled by the kinematic-wave approach and the chemical exchange between the fracture and the matrix through the stagnant boundary layer is rate-limited. The kinematic-wave is a good approximation for the convection–dispersion equation for such system and enables an evaluation of an analytic solution.

$$\frac{\partial c_f}{\partial t} + \nu \frac{\partial c_f}{\partial z} = -k[c_f - c_1(-l, t)] \quad (9.27)$$

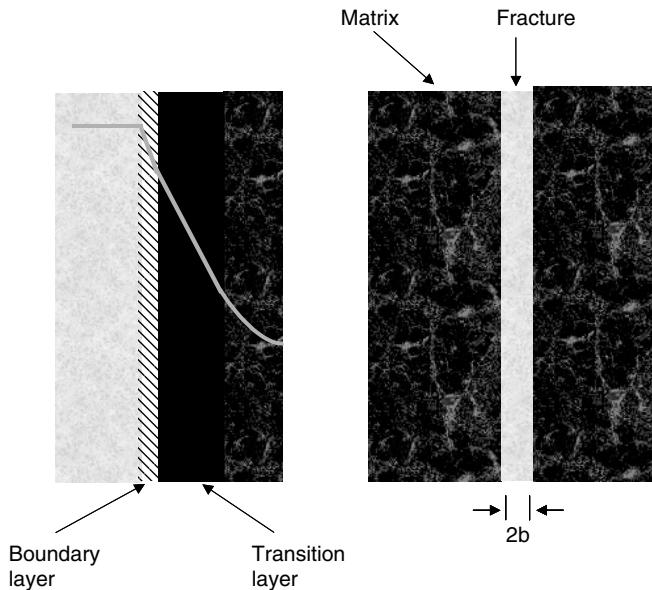


FIGURE 9.2 The fracture–matrix system together with the boundary and transition layers and an illustration of the lateral concentration variation.

where c_f is the dissolved chemical concentrations in the fracture; z is distance in the vertical direction; t is time; and $k [T^{-1}]$ is the mass-transfer coefficient per unit crack width.

The “matrix” is a composite medium. It is composed from a thin transition layer at the interface between the fracture and the matrix and the “real” elsewhere. The transition layer has diffusion properties different from those of the rest of the matrix. For simplicity it is assumed that the matrix is semi-infinite, $-l < x < \infty$, where x is the distance into the matrix perpendicular to the flow direction (Figure 9.2). The diffusion coefficient in the transition layer, $-l < x < 0$ is D_1 and the concentration is defined as c_1 there, while the corresponding quantities in $x > 0$ are D_2 and c_2 .

The differential equation for the transition layer and matrix are

$$\frac{\partial c_1}{\partial t} = D_1 \frac{\partial^2 c_1}{\partial x^2} \quad -l < x < 0, \quad t > 0 \quad (9.28)$$

$$\frac{\partial c_2}{\partial t} = D_2 \frac{\partial^2 c_2}{\partial x^2} \quad x > 0, \quad t > 0 \quad (9.29)$$

The boundary conditions for Equation 9.28 at the interface between the fracture and the transition layer ($x = -l$) is

$$\frac{\theta D_1}{b} \frac{\partial c_1(-l, t)}{\partial x} = k[c_f - c_1(-l, t)] \quad (9.30)$$

where $2b$ is the fracture/crack width and θ is the matrix porosity. Note that the matrix is assumed infinite: see Wallach and Parlange (1998) to remove this constraint if desired.

The boundary layer introduces an additional characteristic time scale to the equilibrium model that postulates an equal concentration over crack width and the matrix interface, namely, rate-limited mass-transfer between the crack and the matrix. Under certain circumstances, this additional time scale plays a major role in determining the concentration distribution along the crack and at its outlet.

The mass-transfer coefficient that describes the chemical transfer driven by a concentration gradient across the hydrodynamic boundary layer is generally described in a simplistic way by the film theory. The concentration is assumed to vary linearly within the boundary layer from its value at the matrix interface and the laterally uniform concentration in the flowing fluid (Figure 9.2). The concentration gradient throughout the boundary layer is not known because the thickness of the concentration boundary layer cannot be measured. The mass-transfer coefficient depends upon relative velocity near the interface and the thickness of the film. If bulk flow in the crack is reasonably steady, then the thickness of the film is essentially constant. The mass-transfer coefficient can be indirectly determined by way of experimental conditions, commonly done by fitting transport models to measured BTCs. Alternatively, a relationship can be obtained directly between the rate of mass-transfer and hydrodynamic conditions at the boundary. For example,

$$b \frac{kL}{D} = 0.626 \left(\frac{Lv\rho}{\mu} \right)^{1/2} \left(\frac{\mu}{\rho D} \right)^{1/3} \quad (9.31)$$

where ρ is the fluid density and μ is its dynamic viscosity, L is the fracture length. Equation 9.31 is written in this form to emphasize its relation to the dimensionless numbers $Sh = kL/D$ (Sherwood number), $Re = Lv\rho/\mu$ (Reynolds number), and $Sc = \mu/\rho D$ (Schmidt number). Different mass-transfer correlations for different fluids, geometries and values of Reynold number are given in Weber and DiGiano (1996). Equation 9.31 can be used in cases where the fracture shape is regular and well-defined with a priori known b , and L .

The boundary condition for Equation 9.29 at the interface between the transition layer and matrix assumes that there is no contact resistance at the surface of separation, $x = 0$.

$$c_1 = c_2 \quad x = 0 \quad (9.32)$$

$$D_1 \frac{\partial c_1}{\partial x} = D_2 \frac{\partial c_2}{\partial x} \quad x = 0 \quad (9.33)$$

An analytic solution of the entire system of differential equations (including the fracture) and associated boundary conditions is complicated. Before an effort to find a solution is allocated it is essential to find out whether the boundary and transition layers exist and what is their effect on the transport dynamics. In the following, intermediate stages of the entire effort are presented. The experimental and mathematical efforts to reach a complete understanding of the transport processes at fracture–matrix interface and their effect on the entire solute transport in the fracture and matrix during contamination and remediation continues.

9.3.1 The Boundary Layer Effect on the Breakthrough Curve at the Fracture Outlet

Wallach and Parlange (2000) presented an analytical solution for the case where the matrix is uniform matrix. The system consists of a fracture a uniform matrix and a boundary layer at their interface (Figure 9.2). The solution was developed by using the Laplace transform. The model was verified by comparing its simulated output to measured BTCs of two experimental studies (Kluitenberg and Horton, 1990; Sidle et al., 1998).

In their study, Kluitenberg and Horton used undisturbed cores of glacial-till-derived soil to displace CaSO_4 with CaCl_2 during saturated, steady fluid flow. The column length was 33 cm and the diameter 18 cm. Following the displacement study, a dye solution was applied at the column inlet and the column was

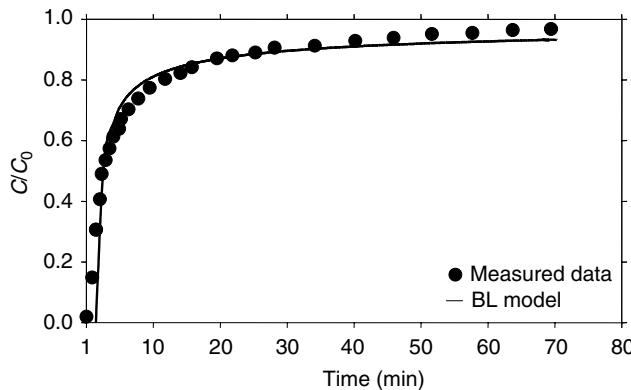


FIGURE 9.3 Comparison between predicted and measured (Kluitenberg and Horton, 1990) BTCs.

sectioned horizontally in 5-cm increments to allow visual characterization of the stained and unstained soil voids. The dye was transmitted through the lower half of the column in one continuous 2-mm diameter channel (column B, table 2 in Kluitenberg and Horton, 1990). As the fracture configuration in column B fit the model assumption of a single-fracture displacement with simultaneous matrix exchange better than the other columns, its BTC was used for the model validation. The predicted BTC fit the measured BTC well: both are shown in Figure 9.3. By using the kinematic wave equation for fracture transport (Equation 9.27), the predicted BTC increases more sharply than the curved, measured BTC increase soon after initial breakthrough. Note that the dispersion in the predicted BTC is formed solely by the solute exchange between the fracture and matrix.

The average pore water velocity, 1.1 cm min^{-1} , determined by Kluitenberg and Horton by the solute flux through the entire soil porosity, is much smaller than the velocity in the preferential flow path determined by the measured first breakthrough time. The flow velocity in the continuous fracture along the entire column length as estimated directly from the measured BTC is $v = 23 \text{ cm min}^{-1}$. This velocity was determined by neglecting hydrodynamic dispersion in the column and assuming a larger first breakthrough time than was actually measured. The model-estimated parameters obtained by best fit between the simulated and measured BTCs were $k = 4.5 \text{ min}^{-1}$, and $b^2/D\theta^2 = 1.4 \text{ min}^{-1}$. Note that the equivalent fracture aperture, b , depends on the matrix diffusion coefficient and vice versa (both unknown). This model was also verified by its successful comparison to natural gradient tracer test of Sidle et al. (1998) which simulated one-dimensional flow in a large ($4 \text{ m} \times 4.8 \text{ m}$ surface area) isolated block of fractured till in Denmark (Wallach and Parlange, 2000).

Recognizing the existence of a boundary layer (BL) at the fracture–matrix interface, rather than assuming an equilibration of concentrations (LE), affects the shapes of BTCs at the fracture outlet. A qualitative analysis of this effect was presented by Wallach and Parlange (1998). The analysis divided the role of the boundary layer resistance on fracture–matrix solute exchange into two stages. The first is soon after displacement initiation, when the lateral chemical flux into the matrix is mostly controlled by boundary layer resistance. Subsequently, the rate of matrix diffusion predominantly controls this flux and the BL model can be replaced by the LE model which involves easy mathematical manipulations and no induction of major errors. The extension of the first stage, during which the boundary layer resistance controls solute exchange, depends on the value of the dimensionless parameter $\varepsilon = kb^2/\theta^2 D$.

The outcome of this quantitative analysis is shown in Figure 9.4 where simulated BTCs at the fracture’s outlet were calculated by the BL and LE models for similar parameters.

The two simulated BTCs deviate initially and come closer as time progresses, as was qualitatively discussed in Wallach and Parlange (1998). Since chemical transport in the crack by both BL and LE models is expressed by the kinematic-wave equation, the first chemical appearance at a certain point along

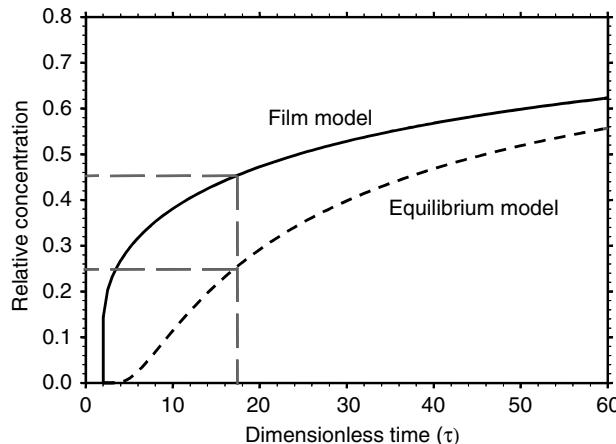


FIGURE 9.4 Comparison between BTCs calculated by the boundary-layer (—) and local-equilibrium (---) models for similar parameter values.

the fracture occurs at the same instant. The BL model predicts lower solute influx into the matrix than does the LE model, owing to the lower predicted concentration at the fracture–matrix interface. The varying concentration-differences between the interface and fracture lead to higher predicted concentrations at the fracture outlet by the BL model. The difference between the concentrations at the fracture–matrix predicted by the two models is initially high and decreases with time as more solute enters the matrix. As time progresses, the differences between the fracture–matrix interface and fracture concentrations converge to a finite value that provides a solute flux equal to the flux predicted by the LE model, the latter being solely controlled by the matrix concentration gradient at the interface.

The parameter ε has a significant effect on the deviation between BTCs predicted by the BL and LE models and its extent during the first stage. This parameter is similar to the Biot number, Bi , used by Crittenden et al. (1986) and others to estimate the role of different rate-limited processes on the adsorption of chemical onto soil aggregates. Larger ε and Bi values denote a higher effect of the mass-transfer coefficient (fast time scale) than matrix diffusion on solute flux to the matrix, and that concentration through the boundary layer becomes rapidly uniform. This can be obtained by relatively large k or small D values. In contrast, high diffusion coefficients in the matrix (smaller ε values) cause rapid removal of solutes from the fracture–matrix interface into the bulk matrix and extend the time during which the mass-transfer coefficient controls the flux. Although not rate-limited parameters, crack width, b , and matrix porosity, θ , have a significant effect on ε since both are raised to the second power.

9.3.2 The Transition Layer Effect on the Concentration Distribution in the Matrix

The verification of models in which diffusion was assumed to control the solute transport in the matrix was widely based on its effect on the breakthrough curves shape at the fracture outlet. A direct evidence of the concentration distribution within the matrix should include concentration measurement at high spatial resolution within the matrix. This can be efficiently and accurately done by non-invasive methods, such as CT, MRI, and others.

Polak et al. (2003a, 2003b) monitored the concentration changes inside a chalk matrix due to NaI diffusion from a vertical and horizontal fracture, respectively, and the back-diffusion process from the contaminated matrix into the fracture using a medical-based X-ray CT scanner. The sample used in the experiment was a 20-cm long, 5-cm diameter chalk core retrieved from a corehole at a depth of ~ 18.3 m. The study area is located in the northern Negev desert of Israel and is underlain by fractured Eocene chalk formations with a thickness of 285 m.

The characteristics of the mass transfer between the fracture and the surrounding matrix were obtained by the concentration distribution within the matrix. The time-dependence of the diffusion and back-diffusion processes was monitored by consecutive scans. The diffusion and back-diffusion experiments were conducted on an artificially fractured chalk core. The confined core was placed in the scanner and the concentration distribution within the core was monitored. The discussion to follow is related to the vertical fracture experiment (Polak et al., 2003a).

Following core saturation, the tracer solution (5% by weight of NaI) was injected into the fracture using a peristaltic pump at a constant concentration and rate of $2.5 \text{ cm}^3 \text{ min}^{-1}$. The volume of the fracture void was about 10 cm^3 , hence it was replaced every 4 min. This injection process lasted for six days, during which the core was scanned eight times (hereafter denoted as sequences 1 through 8). The scanner produces two-dimensional slices through the sample with a thickness of 4 mm and an in-plane pixel resolution of about 0.25 mm. The X-ray energy level used in the experiment was 130 kV at 105 mA and the acquisition time for each scan was 4 sec.

After six days of tracer-solution injection, distilled water was injected using the same peristaltic pump at the same flow rate. The water injection generated a reverse concentration gradient of the tracer from the matrix to the fracture. The flow rate ensured complete removal of the tracer from the fracture without accumulation. This water injection lasted for 11 days, during which the core was scanned 10 times to track tracer migration from the matrix to the fracture.

The tracer's lateral distribution in the matrix for the three selected images during the eight scan sequences is shown (a1, b1, c1) in CT numbers and in Figure 9.5 (a2, b2, c2) in terms of relative concentration, C/C_0 . Figure 9.5a refers to image 38, where both sides of the fracture resembled each other. This resulted in a similar temporal diffusion pattern as evidenced by the identical colored bands in the eight scanning sequences (a1) as well as by the four relative concentration graphs (a2). The effect of the sealed boundaries at the core's top and bottom began dominating tracer diffusion from the third day (fifth scan) through tracer migration from the core edges down and up toward the matrix center. The distribution of C/C_0 across the chalk matrix (a2) formed two distinct zones. The first zone was characterized by a sharp concentration decrease over a thin distance adjacent to the fracture–matrix interface, hereafter denoted transition zone. In the second zone, adjacent to the first one and further into the matrix, the concentration decrease followed a diffusion-type pattern.

Image 12 (Figure 9.5b1) contained a natural fracture, first observed after the core's fracturing, that ran across the upper right face of the artificial fracture. The natural fracture appeared impermeable to water and the saturation of the right face of the natural fracture lagged behind its counterpart on the left. Saturation was finally accomplished by water advecting from the top of the artificial fracture. A similar pattern was observed during tracer injection (b1, second scan on) and was reflected in the sharp relative concentration decrease in the naturally fractured zone (b2, at 3.7 cm along the x-axis). The white circle located at the right edge of image 30 (Figure 9.5c1) demonstrates the effect of a broken core on the diffusion process. The lack of color indicates lack of tracer accumulation, as also evidenced by the zero relative concentration (c2) in this zone. In fact, this hole at the core edge acts like a sink.

The results of the back-diffusion phase measured by Polak et al. (2003a) are exemplified through the same three images used to examine the diffusion phase. Figure 9.6 presents the tracer concentration decrease (in CT numbers) in the matrix of these three images during six of the 10 scan sequences (a1, b1, and c1), as well as the calculated relative concentration (C/C_0) during 4 of these scans (a2, b2, and c2). Note that the first picture and graph for each image represent the last sequence of the diffusion experiment.

Tracer displacement by the water injected into the fracture was documented in image 38 (where both sides of the fracture were almost identical) as a white line in the center of the image (Figure 9.6a1). The relative concentration in the fracture dropped to almost zero as soon as back-diffusion started (Figure 9.6a2, scan sequence 1). However, tracer concentrations away from the fracture did not change much with respect to their distribution during the end of the diffusion phase. Note that while tracer movement from the matrix into the fracture was occurring near the fracture as a result of the reversed concentration gradient, tracer further away from the fracture continued diffusing toward the core edge (scan sequence 3, Figure 9.6a2). By the end of the back-diffusion experiment, tracer concentration inside the core matrix had

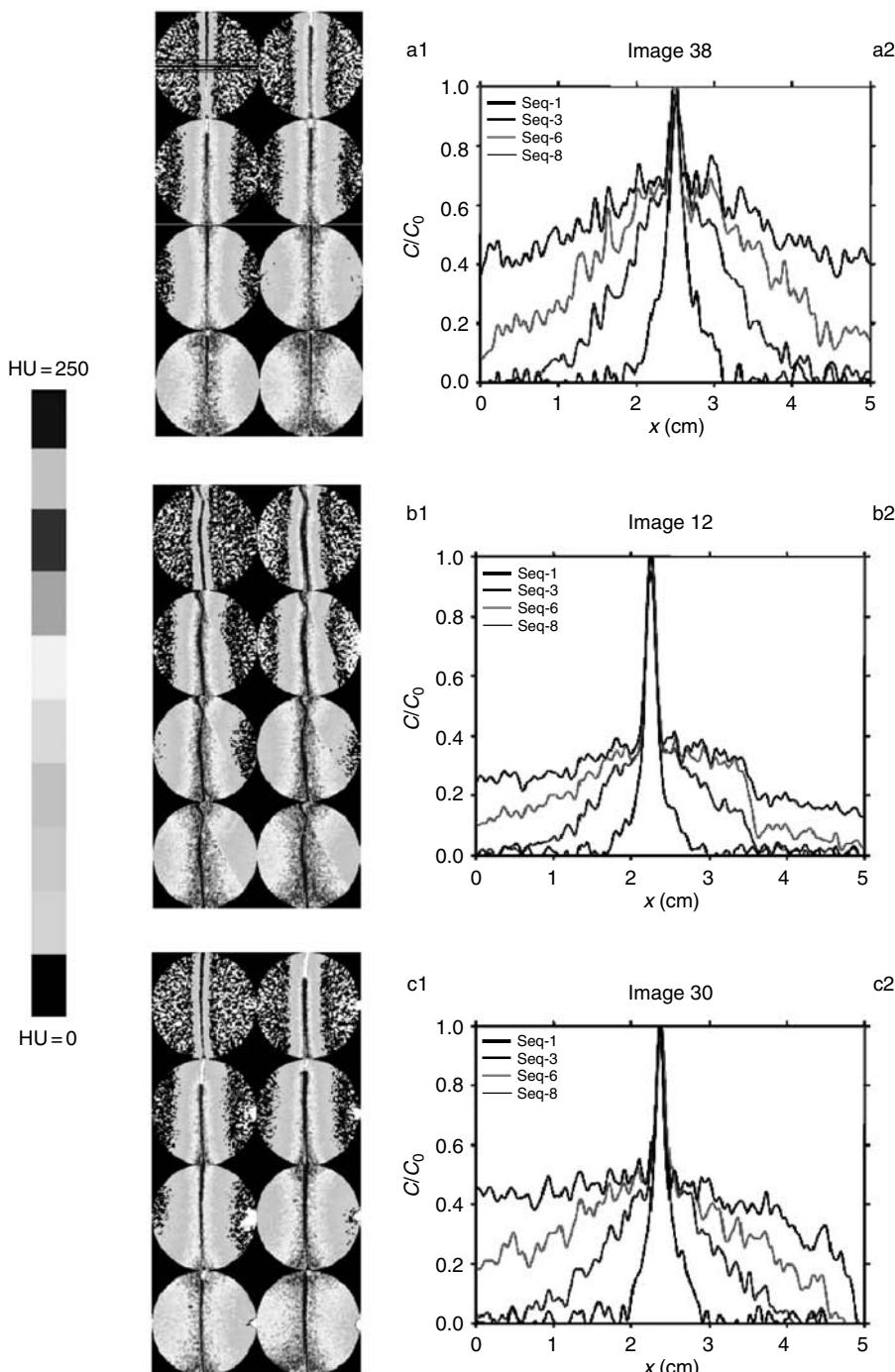


FIGURE 9.5 Tracer distribution during the diffusion phase in the scanned core, presented as both concentrations (in CT numbers) and relative concentrations during scan sequences 1, 3, 6, and 8 in images 38, 12, and 30 (a1,2; b1,2; and c1,2, respectively). The CT numbers refer to the attenuation coefficient, converted to an internationally standardized scale, known as Hounsfield units (HU). The scale is linear, and HU units for air and water are defined as -1000 and 0, respectively. High HU numbers correspond to high-density materials. (From Polak et al., 2003. Chemical diffusion between a fracture and the surrounding matrix: Measurements by computed tomography and modeling. *Water Resour. Res.* 39(4), 1106, doi 10.1029/2001WR000813. With permission)

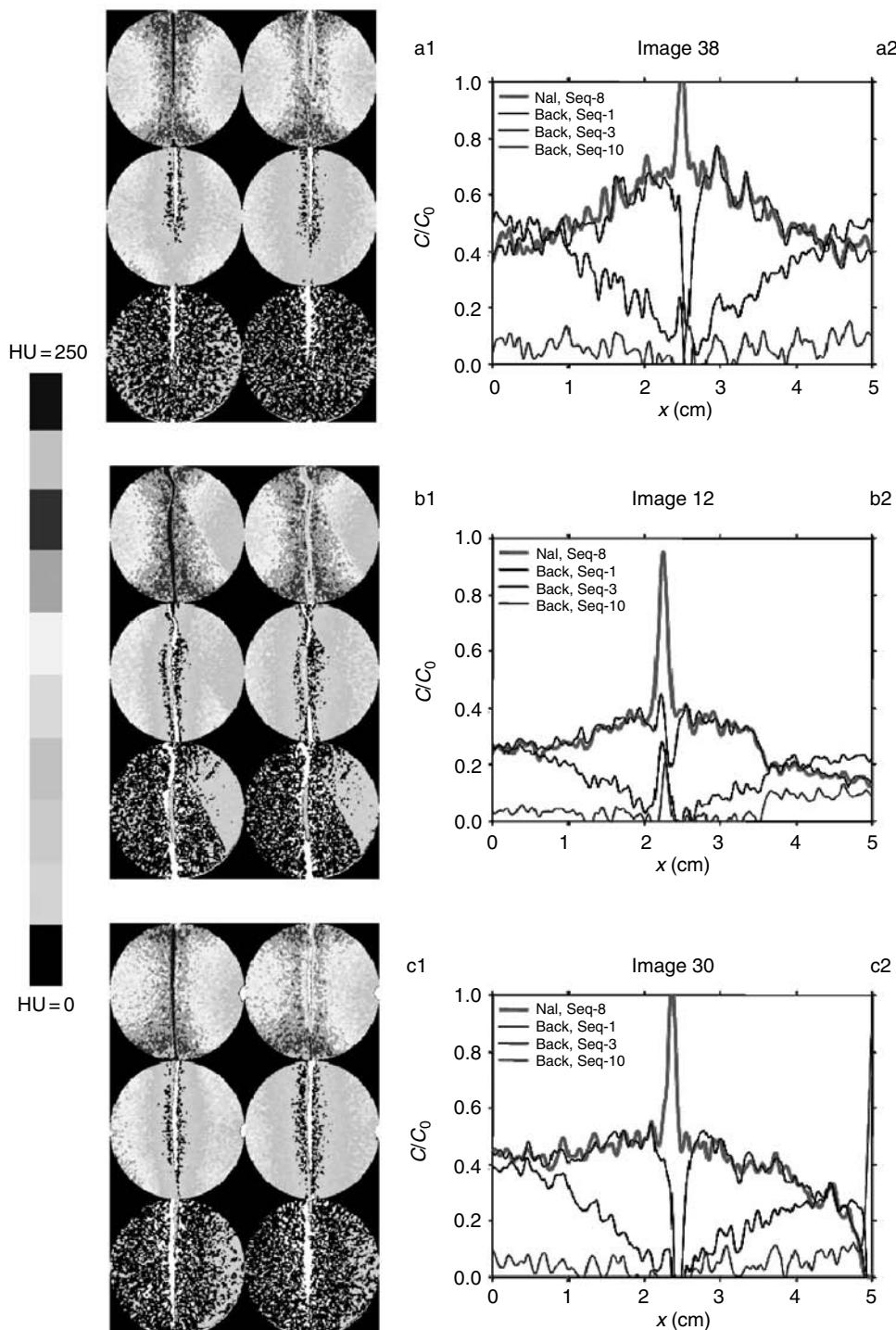


FIGURE 9.6 Tracer distribution during the back-diffusion phase in the scanned core, presented as both concentrations (in CT numbers) and relative concentrations during scan sequences 8 (last for the diffusion phase) and 1, 3, and 10 in images 38, 12, and 30 (a1,2; b1,2; and c1,2, respectively). (From Polak et al., 2003. Chemical diffusion between a fracture and the surrounding matrix: Measurements by computed tomography and modeling. *Water Resour. Res.* 39(4), 1106, doi 10.1029/2001WR000813. With permission)

dropped to almost zero (blue color in scan sequence 10, Figure 9.6a1 and 9.6a2, respectively), although the relative concentration at the core's edges was still higher than that near the fracture. Similar to observations during the diffusion experiment, concentration patterns during back-diffusion in this image were alike on both faces of the fracture. The bottom part of the fracture appears closed (Figure 9.6a1), probably as a result of the confining pressure applied to the core.

The natural impermeable fracture intercepted by image 12 slowed the tracer's back-diffusion (Figure 9.6b). Whereas the left face of the core exhibited a decreasing concentration pattern similar to those observed for image 38, tracer relative concentration simultaneously increased in the right face (Figure 9.6b1, 9.6b2). The reduction in back-diffusion rate was also visible at the end of the experiment (scan sequence 10) where the left side of the artificial fracture was almost tracer-free, while in the right side the relative concentration was much higher (0.02 vs. 0.1, respectively). By the end of the experiment, the tracer had not been completely displaced from the artificial fracture void (Figure 9.6b1 and scan sequence 10 in Figure 9.6b2).

In image 30, the hole in the core's edge that acted like a sink during the tracer diffusion phase turned out to be a tracer source, as evidenced by the spot on the right side of the core (Figure 9.6c1) and the increasing relative concentration (scan sequence 10, Figure 9.6c2).

9.3.3 Modeling the Diffusion and Back-Diffusion — The Transition Layer Concept

Attempts to fit the models of Grisak and Pickens (1980, 1981) and Sudicky and Frind (1982) to the measured concentration distributions failed as they could not simultaneously predict the sharp concentration decrease near the fracture–matrix interface and the concentration decrease further into the matrix (Figure 9.7). These commonly used models for chemical transport along a fracture surrounded by porous matrix assume a continuous concentration variation at the fracture–matrix interface. The simulated curve (using the Sudicky and Frind's model) shown in Figure 9.7 was calculated with a diffusion coefficient of $4 \times 10^{-6} \text{ cm}^2 \text{ sec}^{-1}$ and provides the best fit between the measured and predicted data. This value of the diffusion coefficient used for this simulation is within the range of independently measured values in this chalk (Polak et al., 2002).

The failure of the aforementioned models to simulate the sharp concentration decrease near the fracture–matrix interface possibly indicates that the assumption of equal fracture and matrix

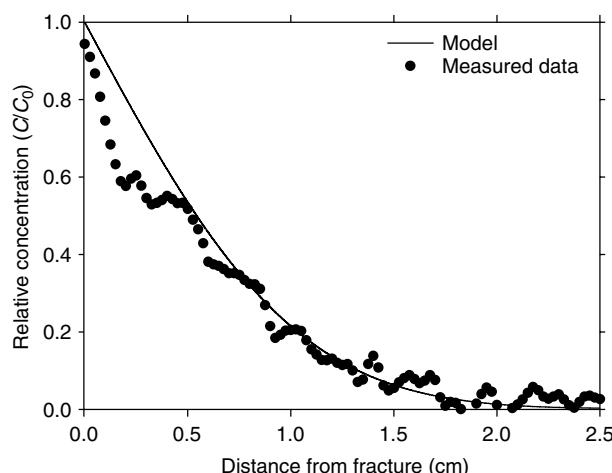


FIGURE 9.7 Measured (scanned) NaI concentrations along the right side of image 38 and its predicted concentrations using the model of Sudicky and Frind (1982). (From Polak et al., 2003. Chemical diffusion between a fracture and the surrounding matrix: Measurements by computed tomography and modeling. *Water Resour. Res.* 39(4), 1106, doi 10.1029/2001WR000813. With permission)

concentrations at their interface should be replaced by a different approach in that zone. The sharp concentration decrease near the matrix–fracture interface indicates the presence of a transition layer, the diffusion coefficient of which is significantly higher than that of the bulk matrix, but lower than that in the fracture. This transition layer consists of mini-fissures and small fractures that developed when the fracture was formed (Labuz et al., 1987; Huang and Kim, 1993). These mini-fissures and small fractures have a higher porosity, which leads to a higher diffusion coefficient.

As a first approximation, it is assumed that the transition-layer concept (Equation 9.32 and Equation 9.33) for this experiment will be simplified and conceptually treated as a boundary-layer (Equation 9.31), in which the concentration varies linearly between its value at the fracture and its value at the interface between the transition layer and the bulk matrix. This approximation is related to the fact that the fracture concentration in this experiment was kept artificially constant. The concept of a thin layer having no storage but with a fixed resistance to diffusion is identical to the concept of well-bore skin in groundwater hydrology and petroleum engineering. The simplified approach for the transition layer to predict the measured concentration distribution within the matrix (Figure 9.5 and Figure 9.6) was used in Polak et al. (2003a). This approach postulates that the transition layer is thin and that the concentration varies linearly within this layer (similar to the assumption used previously by Wallach and Parlange [2000] for the boundary layer). The concentration distribution within the matrix $c_2(x, t)$ for the finite core matrix dimension was obtained by applying the numerical inversion to equation (12) in Wallach and Parlange (2000) and equation (19) in Polak et al. (2003a). The Stehfest algorithm (Stehfest, 1970) for inverse Laplace transform was used for this purpose.

The analysis of the measured data and the model predictions presented from hereon are relevant to either side of the fracture (left or right) of each image. The origin of the x -axis in the following figures is at the interface between the transition layer and the bulk matrix. The transition layer width was determined from the measured lateral concentration distribution within the matrix in the region where the concentration sharply decreased. As it turns out, this layer contained two to three CT pixels (0.5 to 0.75 mm). Figure 9.8a presents the measured concentrations and fitted curves for image 38 at three of the eight scan times run during the diffusion phase. The fitted values for the three time steps were $D = 4 \times 10^{-6} \text{ cm}^2 \text{ sec}^{-1}$ and $\varepsilon = 0.18$. Note that each curve in Figure 9.8a represents the average of five cross-sections in the middle of the sample. The fitted diffusion coefficient, D , was kept constant for the entire core whereas the value of ε was changed for the different images in order to obtain the best fit to the measured data. As D is constant ($4 \times 10^{-6} \text{ cm}^2 \text{ sec}^{-1}$), $b = 0.05 \text{ cm}$ (half the fracture width as estimated in prior experiments), and L and θ are calculated from the CT measurements, the fitted ε values actually provide the fitted values of the mass-transfer coefficient, k (Equation 9.8). The measured and fitted values for images 38, 42, and 47 are presented in Table 3 in Polak et al. (2003a), where it can be seen that the variations among the fitted k values are small.

The initial tracer concentration inside the rock matrix at the beginning of the back-diffusion phase in Polak et al. (2003a) was identical to that at the end of the diffusion phase. As the results were calculated in the Laplace domain and the inversion back to the time domain was performed numerically, the superposition principle was applied to replace the initial condition into a boundary conditions. The diffusion model was run for each image from the beginning of the diffusion phase, using fitting parameters identical to those used during the diffusion phase for each image. The model was then run again from the beginning of the back-diffusion phase without changing the parameters' values. The concentration distribution for the back-diffusion phase was obtained by subtracting the latter from the former. Consequently, the initial condition of the back-diffusion phase was modified into two new boundary conditions (the concentration values at different times).

Figure 9.9a–c presents the measured (CT scanned) and predicted relative tracer concentrations for images 38, 42, and 47 at the beginning of the back-diffusion phase (scan sequence 1) and during two more of the 10 scans performed. The model successfully predicted the concentration decrease near the transition layer–matrix boundary from $c/c_0 = 0.6$ in scan sequence 1 down to 0.1 in scan sequences 3 and 5. It also successfully simulated the before discussed relative concentration increase at the core's edge (sealed boundary) for sequence 3.

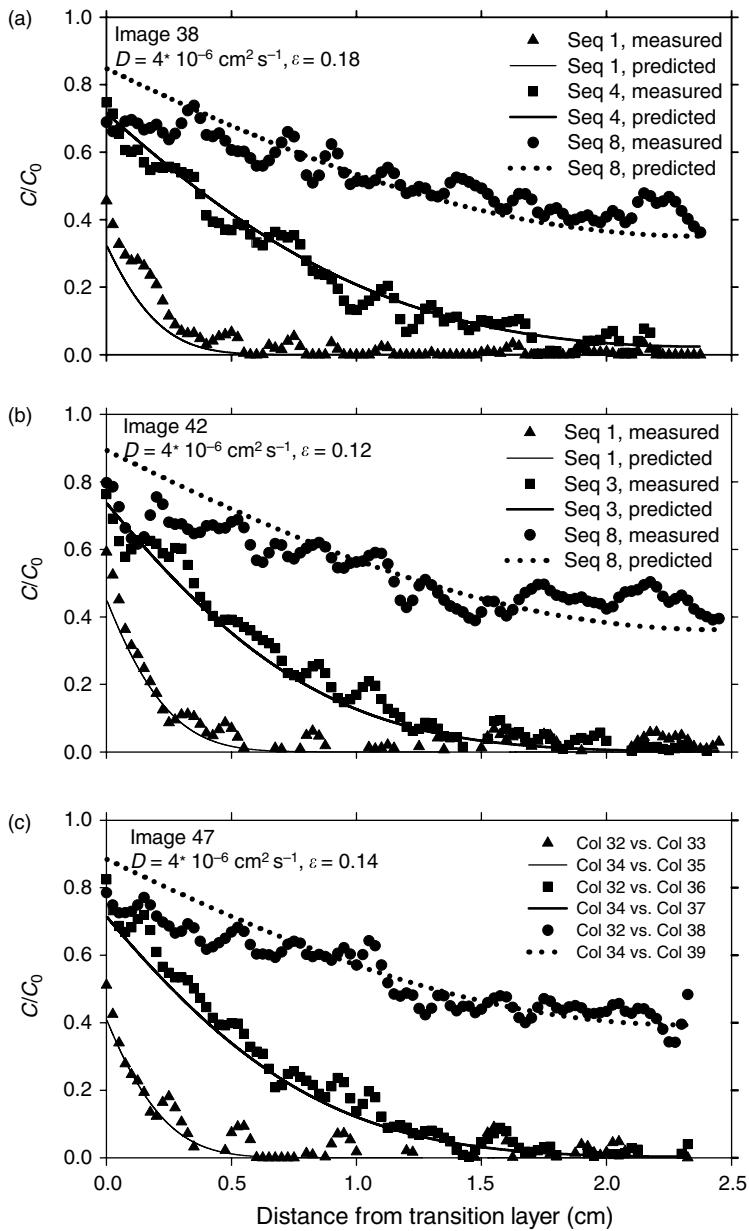


FIGURE 9.8 Measured (scanned) and modeled NaI concentrations in the diffusion experiment along image 38 (a), 42 (b), and 47 (c) in three different scan sequences. (From Polak et al., 2003. Chemical diffusion between a fracture and the surrounding matrix: Measurements by computed tomography and modeling. *Water Resour. Res.* 39(4), 1106, doi 10.1029/2001WR000813. With permission.)

Following the successful prediction of the measured data, the mathematical model was used in Polak et al. (2003a) to estimate the probable duration of remediation efforts on a field scale. The study area used for this purpose is an industrial complex in the northern Negev desert, Israel. It hosts chemical plants as well as many solid- and liquid-waste-disposal sites that have been continuous contamination sources for the 27 years of its operation. This complex was constructed on fractured chalk formations with fracture spacing of ~ 5 m. These fractures are considered the main conduits for contaminant transport to the subsurface (Nativ et al., 1999). The core used for the diffusion and back-diffusion experiments was taken

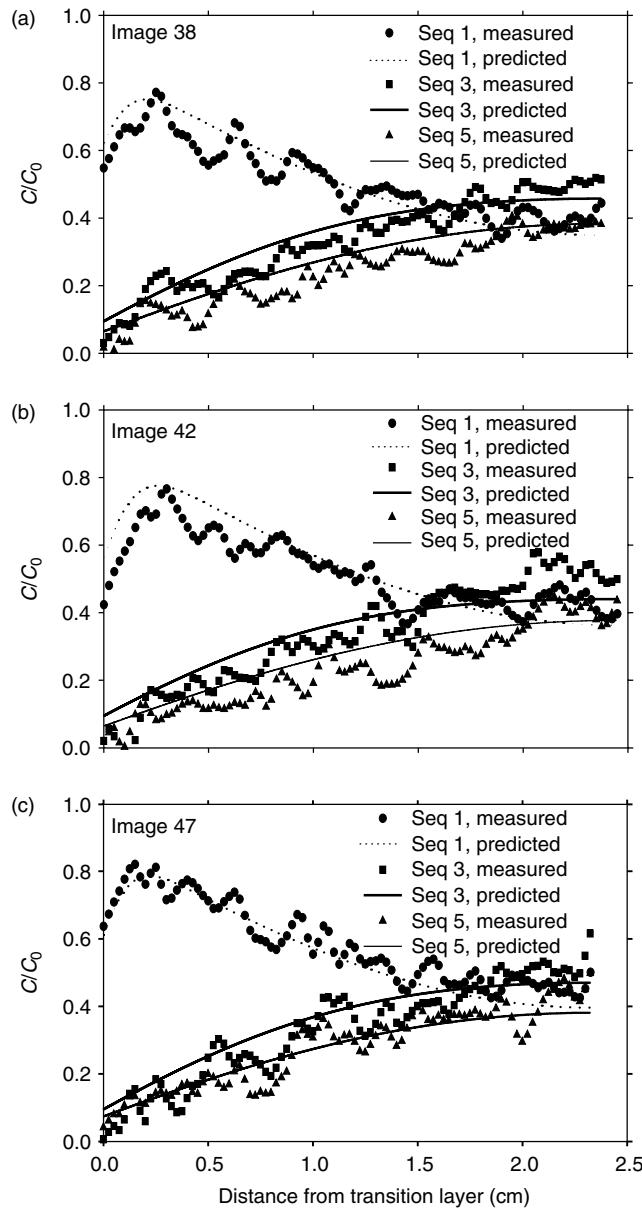


FIGURE 9.9 Measured and modeled NaI concentrations in the back-diffusion experiment along image 38 (a), 42 (b), and 47 (c) in three different scan sequences. (From Polak et al., 2003. Chemical diffusion between a fracture and the surrounding matrix: Measurements by computed tomography and modeling. *Water Resour. Res.* 39(4), 1106, doi 10.1029/2001WR000813. With permission.)

from a corehole located on site. Figure 9.10a displays the modeled tracer relative concentrations in the rock matrix following 1, 2.5, 5, 10, and 20 yr of its diffusion from the fracture. The model was applied to a lateral distance of 250 cm (half the typical fracture spacing on site) of unfractured chalk block (with a sealed boundary at its end) and with $\varepsilon = 0.15$ and $D = 4 \times 10^{-6} \text{ cm}^2 \text{ sec}^{-1}$ as the fitting parameters. The modeling results followed a typical diffusion pattern in the chalk matrix and 20 yr of tracer diffusion resulted in a calculated lateral penetration of ~ 150 cm into the matrix.

Figure 9.10b displays the modeled distribution of the relative tracer concentrations (in logarithmic scale) inside the rock matrix for the back-diffusion phase, thereby simulating the impact of flushing clean

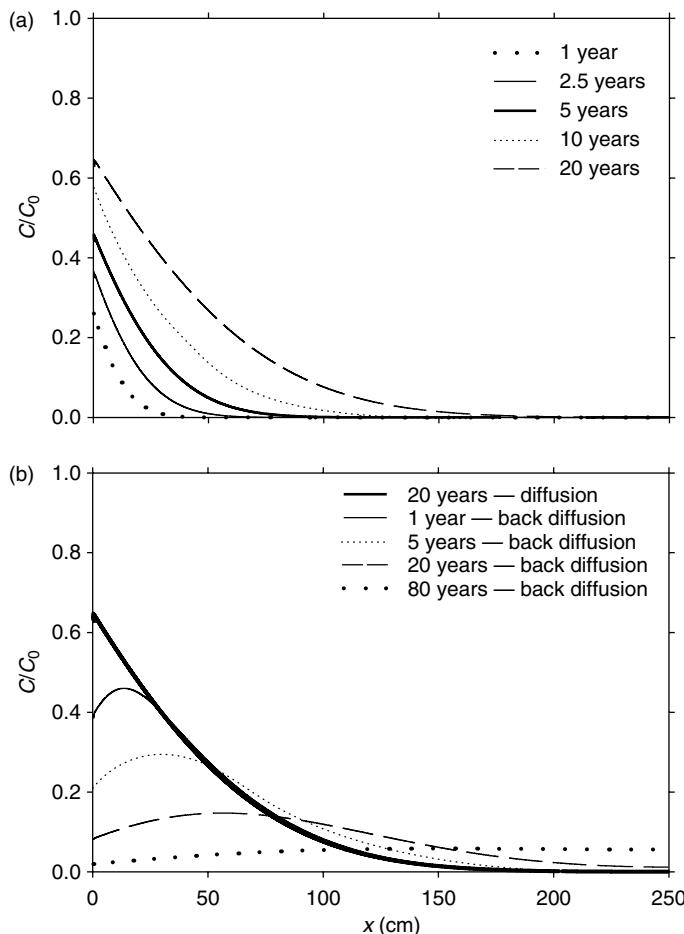


FIGURE 9.10 (a) Model prediction of tracer distribution in the 250-cm wide rock matrix following its diffusion during five different time periods, and (b) its redistribution following back diffusion during different time periods. (From Polak et al., 2003. Chemical diffusion between a fracture and the surrounding matrix: Measurements by computed tomography and modeling. *Water Resour. Res.* 39(4), 1106, doi 10.1029/2001WR000813. With permission.)

water into the fractures (e.g., pump-and-treat practice). The model parameters (ε and D) were kept identical to those used in the diffusion phase. Whereas the bold line represents the distribution of the relative tracer concentration following 20 yr of diffusion (initial condition), the other lines display the calculated distributions following 1, 5, 20, and 80 yr of back-diffusion. These simulations clearly show that while the relative tracer concentration near the fracture drops to almost zero after a long flushing, it continues to progress deeper into the rock matrix further away from the fracture. In fact, according to the model, the tracer is expected to reach the sealed boundary at 250 cm 20 yr after flushing of the fracture is initiated, thereby transforming the rock matrix into a long-term contaminant source for the fracture. Obviously, according to these calculations, remediation through clean-water injection into the fractures would take hundreds of years, rendering it practically impossible.

9.4 Dual-Porosity Models with a Distribution Layer at the Soil Surface

The models presented in the previous section assume that the soil surface layer has a structure similar to that of the subsoil and preferential flow takes place along the whole soil pedon. Recent observations

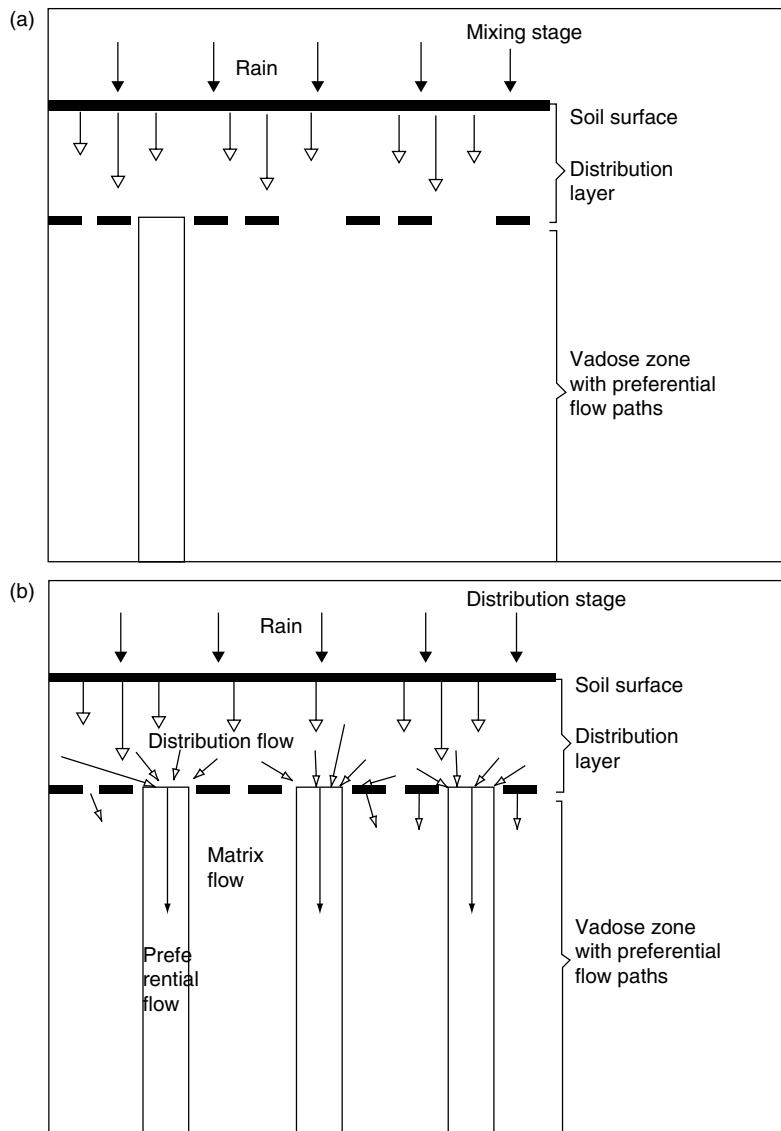


FIGURE 9.11 Conceptual framework of preferential flow model. (a) The mixing stage in the distribution layer and (b) the distribution into the preferential flow paths.

(Parlange et al., 1988; Steenhuis et al., 1994; Ritsema and Dekker, 1995; and others) have shown that a thin (2–30 cm) soil surface layer (distribution layer) exists in many cases at the soil surface where flow takes place throughout this layer. The precipitation in the distribution layer is funneled into the preferential paths in the subsoil (Figure 9.11). The distribution layer typically has a higher conductivity (especially in the horizontal direction) than the soil profile below it. Its thickness depends on various factors such as tillage, vegetation, and bioturbation. One of the approaches to model the distribution layer is to assume that it is a linear reservoir (Steenhuis et al., 1994). A linear reservoir is equivalent to a well-mixed reservoir where the momentary solute concentration is uniformly distributed throughout its depth. The lumped mass balance equation for the distribution layer is

$$\frac{dc_0}{dt'} = -\frac{J_w}{\theta_s l} c_0 \quad (9.34)$$

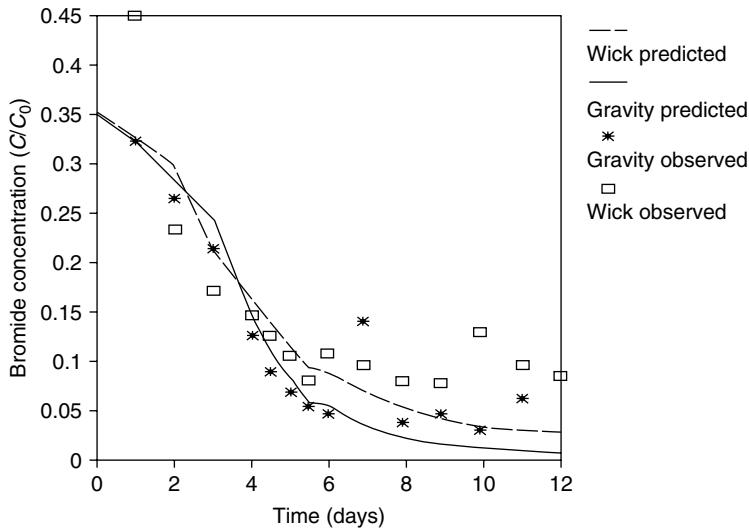


FIGURE 9.12 Observed and predicted outflow concentrations of the wick and gravity pan samplers for the Cornell Orchard site.

where θ_s is its saturated water content, J_w is the flux leaving the distribution zone, and l is the depth of the distribution layer. $c_0(t)$ is the dissolved chemical concentration in the distribution layer. Steenhuis et al. (1994) assumed that there was no interaction in preferential flow paths between the preferentially moving water and the matrix. The concentration of solutes leaving the distribution layer (and in preferential flow paths) after the solute is completely mixed in the distribution zone is

$$c_0 = \frac{M}{\theta_s l} \exp \left(-\frac{\int J_w dt}{\theta_s l} \right) \quad (9.35)$$

where M is the amount of solute applied per unit area. Figure 9.12 shows the results of an experiment carried out in the Cornell Orchard on a Rhinebeck loamy soil with a 30-cm distribution zone on top of a glacial till with macropores in the structural cracks. A pulse of chloride was added on day zero followed by daily irrigations of 3 cm. The concentration in the preferential flow paths was measured in wick and gravity pan samplers at 60 cm depth. Although the fit is quite good, it is obvious that it cannot fit the BTCs such as were obtained by Anderson and Bouma (1977). A more complicated model is needed in which the distribution layer as well as the interaction in the subsoil between macropores and matrix (introduced above) is taken into account.

In order to do so, assume a time interval, τ , when the solutes that enter the soil surface enter the preferential paths. Thus, the concentration that enters the preferential flow paths (not assuming instantaneous mixing) is

$$c_0(\tau) = c_i(1 - e^{-\alpha\tau}) \quad (9.36)$$

to be used in Equation 9.10. Following the small perturbation solution above, Wallach and Steenhuis (1998) obtained the *zeroth* and first approximations concentration in the preferential paths at point z

$$A_0(z, t) = c_i \exp \left[-\frac{kz}{\theta_m v} \right] \left\{ 1 - \exp \left[-\beta \left(t + \tau - \frac{z}{v} \right) \right] \right\} \cdot u \left(t - \frac{z}{v} \right) \quad (9.37)$$

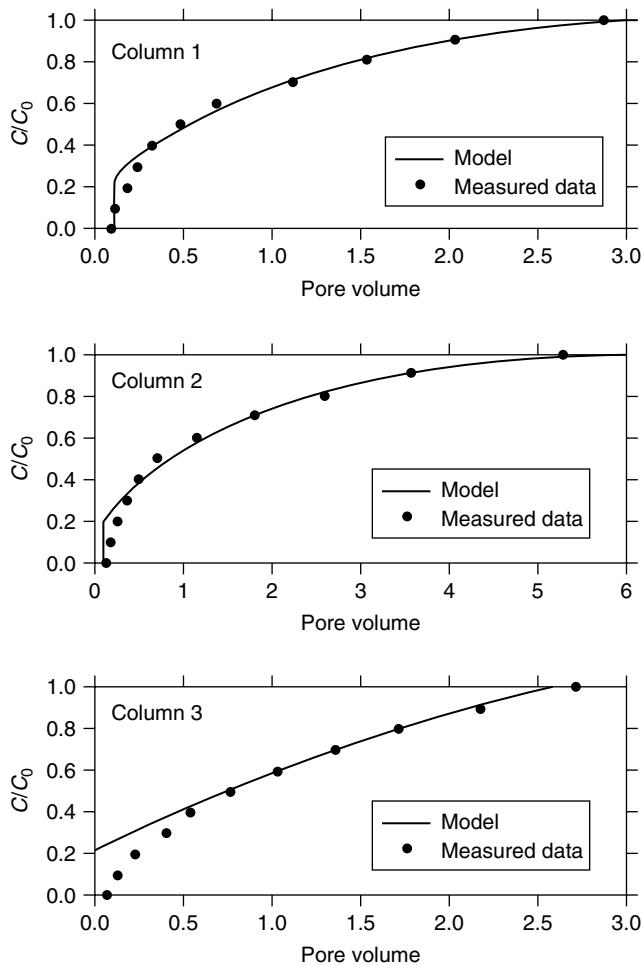


FIGURE 9.13 Measured (Data from Anderson, J. L. and Bouma, J. 1977. *Soil Sci. Soc. Am. J.* 41: 413–418.) and predicted BTCs at the outlet of three initially-saturated columns.

$$A_1(z, t) = \frac{c_i k^2}{\theta_m \theta_{im}} \exp\left[-\frac{kz}{\theta_m v}\right] \frac{z}{v} \left\{ t - \frac{z}{v} - \frac{\exp[-\beta\tau]}{\beta} \left[1 - \exp\left[-\beta\left(t + \tau - \frac{z}{v}\right)\right] \right] \right\} \cdot u\left(t - \frac{z}{v}\right) \quad (9.38)$$

This model was also validated by using BTCs that were measured by Anderson and Bouma (1977) for initially saturated columns containing structured soils. Both, measured (figure 1 in Anderson and Bouma, 1977) and predicted BTCs (Equation 9.38) are shown in Figure 9.13. The initially saturated experimental runs have much longer initial breakthrough times and BTCs tailing compared with the initially drained runs where the now air-filled large pores were drained quickly. Emptying these paths by draining the columns prior to the runs enables the infiltrating solute in the initially drained columns to flow directly through them and bypass most of the columns matrix. A very good fit is obtained between the model output and the measured BTCs. While τ expresses the travel time in the vertical direction within the distribution layer prior to entering the preferential path, the average residence time in the distribution layer, β^{-1} , expresses both the vertical flow and the horizontal travel time to the preferential paths beneath it. In all three columns (Figure 9.13) $\alpha^{-1} > \tau$, which indicates that the average residence

time of the converging flow in the distribution layer is larger than the time for downward vertical flow. Note that the dispersion in the BTCs in Figure 9.13 is only due to the interaction between the mobile and stagnant solutions and is not the result of the dispersion during the transport in the preferential paths, since this transport was modeled by the kinematic-wave model (Equation 9.4) using $D = 0$ in the convective-dispersive equation.

9.5 Multi-Pore Group Models

The main purpose of the models described above was to find the relationship between the flowing fluid in the preferential paths and the stagnant fluid within the matrix and its effect on the BTC. As will be described now, the matrix fluid, which was assumed to be immobile in the models above, may be mobile as well and appear later at the outlet with lower concentration during a longer time period. However, even if the matrix fluid flows with much lower velocities than the preferential flow, its main effect on the BTC at earlier times will be as if it were stagnant. Dye infiltration experiments by Omoti and Wild (1979) revealed rapid transport through pores too small to be considered macropores. A secondary micropore system (smaller than the macropores), also contributes to preferential flow under variably saturated-unsaturated conditions and was also observed by Germann and Beven (1981). Roth et al. (1991) found that rainfall-driven movement of a single surface-applied tracer pulse separated into a slowly moving pulse through the soil matrix and a series of fast preferential flow pulses. The sporadic preferential flow pulses resulted in rapid solute transport to 220 cm and accounted for 58% of the total mass. However, flow through the secondary pore system within the soil matrix accounted for the remaining 42% and moved the solutes to 84 cm. Jardin et al. (1990) monitored rainfall-driven vertical movement of bromide through a large soil pedon and found a rather continuous movement through a small-pore region that lasted for several days following rainfall, while rapid movement occurred as discrete pulses through the large-pore region during events. Jury and Flühler (1992) noted that substantial water flow may be occurring in the surrounding matrix as well as in the preferential flow region. They also noted that solutes partition into a rapid or preferential flow region and a slower but still mobile matrix flow region, each of which may embody a smaller but significant degree of water flow, and concluded that the dual-porosity model is unsatisfactory for representing media wherein transport occurs both in the preferential flow region and in the bulk matrix.

To model preferential flow through a soil profile with multi-pore groups is to assume that the soil is a composite of flow paths with different velocities where mixing may or may not take place among the different flow paths. Dependent on the scale of observation, such flow regions may be pores, fissures, wetted fingers, fine porous matrix, or any region with a relatively well-defined velocity. If the flow regions are non-interacting, meaning that the boundaries of each region are impermeable for solute transfer, an individual solute parcel remains for its entire travel time within the region it entered. If the particle can move laterally through the boundary of its flow region, it experiences the velocities of other flow regions. The velocity, assumed constant within a flow region, can be a deterministic or a random variable distributed over the entire flow cross-section. The fate of a particle that enters a bundle of hypothetically parallel flow regions at the inlet end is now considered.

The straightforward extension of the mobile-immobile model is a two-region convective-dispersive model in which mass is transported from one region to the other in response to differences in the residence fluid concentration. Approximate solutions to this problem have been obtained by Skopp et al. (1981), Van Duin and van der Zee (1986), and Gerke and van Genuchten (1993). Skopp et al. (1981) developed a model where the liquid-filled pores are partitioned into two distinct pore size classes. One region represents macro- or inter-aggregate porosity, and the other represents a mobile matrix porosity. The regions have different hydraulic and transport characteristics and an interaction coefficient characterizes the linear transfer between them. A small interaction coefficient was assumed, and a regular perturbation method was used to solve the model equations. The leading-order solution ignores interaction between the two regions, and flow and transport take place in each one of them independently. The interaction between

the two regions appears in the higher order approximations. If the interaction coefficient is sufficiently large, the convective-dispersive equation is obtained. Comparing the model's output to the measured data of Anderson and Bouma (1977) showed that the interaction term between the two regions is higher than anticipated and the concentration distribution just after the breakthrough time was not well described. Gerke and van Genuchten (1993) also used the two-mobile porosity concept to separate non-steady water flow and transport equations for the matrix and macropore. The chemical exchange between the two regions is described by transfer coefficients which are often functions of the pore size.

Lindstrom and Boersma (1971) separated the pore space of a saturated soil into a finite number of subdomains with characteristic pore sizes. They estimated the flow velocity in each pore domain by Poiseuille's law and calculated total transport by superimposing analytical solutions of the convection-dispersion equation. Their model was not successful in describing experimental breakthrough curves (Rao et al., 1976) as it did not include any mixing of solutes between the pore classes. Steenhuis et al. (1990) followed a similar conceptual framework and proposed a mathematical model that considers many flow and transport domains. Extensions and improvements to the model were later proposed by Stagnitti et al. (1991, 1995) and Durner and Flühler (1996). The preferential flow paths were characterized by taking piecewise linear approximations of the hydraulic conductivity, resulting in a number of pore groups with a mobile transport velocity in each.

Exchange processes between pore groups require a calibration to a particular field or laboratory experiment. Unlike the usual modeling assumptions applied in the CDE, the concentration of solutes in the percolating water is dependent on the varying rate of applied water and the time period between rainfall and chemical application. This is achieved by relating the solute flux to the water flux. Thus, transient field conditions can be simulated. The model can be applied to both large-scale field experiments and small-scale laboratory experiments. As the model has been fully described elsewhere (Steenhuis et al., 1990; Stagnitti et al., 1995; Parlange et al., 1996), we will present only a brief description here.

The total amount of moisture, $\theta(x, t)$ in the soil at time, t , and point, x , is the sum of all individual moisture contents for each capillary bundle, p

$$\theta(x, t) = \sum_{p=0}^N \theta_p(x, t) \quad (9.39)$$

where θ_p is the individual moisture content for the p th pore group. The maximum amount of moisture that each group can hold and transmit is $\Delta M_p = M_p - M_{p-1}$, where M_p and M_{p-1} are various moisture contents representing upper and lower limiting values for the p th group and are a function of the size of pores in each group. When $\theta_p = \Delta M_p$ then all the pathways for the p th group are completely saturated with soil moisture. The group's moisture content, θ_p , is a function of the vertical percolation rate, q_p , effects of precipitation and evapotranspiration, and loss or gain of moisture from interaction and exchanges with other groups. The mass balance equations for the water and solute in the different pore groups and the algorithm to get the moisture and BTCs for different scenarios are presented in detail in Stagnitti et al. (1995). As the networks of flow paths in the soil are interconnected, the exchange of moisture and solutes between the various capillary bundles ranges from a complete exchange of moisture and solute to no exchange or mixing. The basic idea behind the different mixing degrees is the existence of a virtual "common" pool where moisture can be extracted from any group and placed into it, while any fraction of moisture may move back from the common pool into the p th group after mixing. The values of the mixing coefficients depend on the soil type and would normally have to be found by careful experimentation. The preferential flow model has been successfully applied to a number of field and laboratory experiments (e.g., Steenhuis et al., 1990; Nijssen et al., 1991; Stagnitti et al., 1991, 1995; Parlange et al., 1996).

Recently, Skopp and Gardner (1992), who generalized the model of Skopp et al. (1981), and Durner (1992) presented an analytical model for a continuous distribution of local flow velocities perpendicular on an axis orthogonal to the main transport direction. Skopp and Gardner (1992) used the method of moments to relate the dispersion coefficient to flow velocity, which can then be substituted into the

convective-dispersive equation for large interaction between the flow domain perpendicular to the main flow direction.

Durner and Flühler (1996) expanded the previous model by presenting a numerical solution to a conceptual multi-domain model for the transport of solutes in saturated-unsaturated soils. The pore space in this model is represented by a continuous distribution of pore size on a virtual structural coordinate, orthogonal to the spatial coordinate. The chemical transport is simulated by the CDE. Following the method of Steenhuis et al. (1990) a two-step procedure for convective transport and mixing was applied. In the first step, the transport in each pore domain along the spatial axis is solved. In the second step, solute exchange between the pore domains is modeled by a finite difference approximation of the diffusion equation. The model presents the gradual change of convection dominated transport to convective-dispersive transport with time and depth. The length scale where this change takes place depends on the pore size distribution of the porous medium, on the intensity of lateral mixing, and on the degree of saturation.

Symbols

α	rate coefficient describing mass-transfer between the mobile and immobile regions
a, b	Laplace transform of A, B
β^{-1}	average residence time in distribution layer
b	fracture/crack width
c	solute concentration
d	finger diameter
D	dispersion coefficient
$\varepsilon, \varepsilon'$	small dimensionless parameter
J	mass flux
κ	π or 4.8 in 2 and 3 dimensions, respectively
k	rate coefficient describing mass-transfer between mobile and mobile regions in Wallach and Steenhuis (1998) model, mass-transfer coefficient per unit crack width in Wallach and Parlange (1998, 2000)
K	soil hydraulic conductivity
l	depth of distribution layer
L	column/fracture length
M	applied mass per unit area
μ	fluid dynamic viscosity
N	number of pore groups
θ	moisture content, matrix porosity
q	solute flux
Q	imposed surface water flux
ρ	fluid density
Re	Reynolds number
s	Laplace variable
S	sorptivity
Sc	Schmidt number
Sh	Sherwood number
τ	mixing zone
t	time
u	step-function
v	pore water velocity
x	direction perpendicular to flow
z	direction of flow
Sc	$=\mu/\rho D$ (Schmidt number)

Subscripts

e	effective porosity
f	fracture
F	finger tip
I	initial
im	immobile
m	mobile
n	order of expansion
p	pore group
w	water
0	zeroth-order approximation
1	first-order approximation, transition layer at the fracture–matrix interface
2	matrix

References

- Anderson, J. L. and Bouma, J. 1977. Water movement through pedal soil. I. Saturated flow. *Soil Sci. Soc. Am. J.* 41: 413–418.
- Bahr, J. M. and Rubin, J. 1987. Direct comparison of kinetic and local equilibrium formulations for solute transport affected by surface reactions. *Water Resour. Res.* 23: 438–452.
- Baker, R. S. and Hillel, D. 1990. Laboratory tests of a theory of fingering during infiltration into layered soils. *Soil Sci. Soc. Am. J.* 54:20–29.
- Barenblatt, G. I., Zheltov, I. P., and Kochina, I. N. 1960. Basic concepts in the theory of seepage of homogeneous liquids in fissured rocks (strata). *Prikl. Mat. Mekh.* 24: 852–864; *J. Appl. Math. Mech.* 24: 1286–1303.
- Bear, J. 1972. *Dynamics of Fluids in Porous Media*. Elsevier, New York.
- Beven, K. and Germann, P. 1981. Water flow in soil macropores. II. A combined flow model. *J. Soil Sci.* 32: 15–29.
- Biggar, J. W. and Nielsen, D. R. 1962. Miscible displacement. II. Behavior of tracers. *Soil Sci. Soc. Am. Proc.* 26: 125–128.
- Biggar, J. W. and Nielsen, D. R. 1967. Miscible displacement and leaching phenomenon. *Agron. Monogr.* 11: 254–274.
- Boll, J., Selker, J. S., Shalit, G., and Steenhuis, T. S. 1997. Frequency distribution of water and solute transport properties derived from pan sampler data. *Water Resour. Res.* 33: 2655–2664.
- Brusseau, M. L., and Rao, P. S. C. 1989. Modeling solute transport in structured soil: A review. *Geoderma.* 46: 169–192.
- Brusseau, M. L., Gerstl, Z., Augustijn, D., and Rao, P. S. C. 1994. Simulating solute transport in an aggregated soil with the dual-porosity model: Measured and optimized parameter values. *J. Hydrology* 163: 187–193.
- Chen, C. and Wagenet, R. J. 1992. Simulation of water and chemicals in macropore soils. Part 2. Application of linear filter theory. *J. Hydrology* 130: 127–149.
- Chow, V. T. 1964. *Handbook of Applied Hydrology*. McGraw-Hill, Inc., New York.
- Crittenden, J.C., Hutzler, N.J., and Geyer, D.G. 1986. Transport of organic compounds with saturated ground water flow: Model development and parameter sensitivity. *Water Resour. Res.* 22: 271–284.
- Dekker, L. W. and Ritsema, C. J. 1994. Fingered flow: The creator of sand columns in dune and beach sands. *Earth Surf. Process. Landforms* 19: 153–164.
- Durner, W. 1992. New concepts for the functional description of soil hydraulic properties and for the modeling of pollutant transport during stormwater infiltration into multi porosity soils, in *Hydrological and Pollutational Aspects of Stormwater Infiltration*. Eds. M. Grottger and W. Schilling, Proc. Eur. Workshop, Schriftenreihe der EAWAG, 3, CH-8600 Diibendorf, pp. 125–143.

- Durner, W. and Flühler, H. 1996. Multi-domain model for pore-size dependent transport of solutes in soils, in *Fingered Flow in Unsaturated Soil: From Nature to Model*. Eds. T. S. Steenhuis, C. J. Ritsema, and L. W. Dekker. Special Issue of *Geoderma*. 70: 281–298.
- Ela, S. D., Gupta, S. C., Rawls, W. J., and Moncrief, J. F. 1991. Role of earthworm macropores formed by *Aporrectodea tuberculata* on preferential flow of water through a typic hapludoll, in *Preferential Flow*. Eds. T. J. Gish and A. Shirmohammadi, Proc. Am. Soc. Agricultural Engineers National Symp., December 16–17, 1991, Chicago, IL, pp. 68–76.
- Flühler, H., Durner, W., and Flury, M. 1996. Lateral solute mixing processes — A key for understanding field-scale transport of water and solutes, in *Fingered Flow in Unsaturated Soil: From Nature to Model*. Eds. T. S. Steenhuis, C. J. Ritsema, and L. W. Dekker. Special Issue of *Geoderma* 70: 165–183.
- Fujikawa, Y. and Fukui, M. 1990. Adsorptive solute transport in fractured rock: Analytical solutions for delta-type source conditions. *J. Contam. Hydrol.* 6: 85–102.
- Gerke, H. H. and van Genuchten, M. T. 1993. A dual-porosity model for simulating preferential movement of water and solutes in structured porous media. *Water Resour. Res.* 29: 305–319.
- Germann, P. and Beven, K. 1981. Water flow in soil macropores. *J. Soil Sci.* 32: 1–13.
- Glass, R. J., Oosting, G. H., and Steenhuis, T. S. 1989a. Preferential solute transport in layered homogeneous sands as a consequence of wetting front instability. *J. Hydrology*. 110: 87–105.
- Glass, R. J., Parlange, J.-Y., and Steenhuis, T. S. 1989b. Wetting front instability. 1. Theoretical discussion and dimensional analysis. *Water Resour. Res.* 25: 1187–1194.
- Glass, R. J., Steenhuis, T. S., and Parlange, J.-Y. 1989c. Wetting front instability. 2. Experimental determination of relationships between system parameters and two-dimensional unstable flow field behavior in initially dry porous media. *Water Resour. Res.* 25: 1195–1207.
- Glass, R. J., Parlange, J.-Y., and Steenhuis, T. S. 1991. Immiscible displacement in porous media: Stability analysis of three-dimensional, axisymmetric disturbances with application to gravity-driven wetting front instability. *Water Resour. Res.* 27: 1947–1956.
- Glass, R. J. and Nichol, M. J. 1996. Physics of gravity fingering of immiscible fluids within porous media: An overview of current understanding and selected complicating factors, in *Fingered Flow in Unsaturated Soil: From Nature to Model*. Eds. T. S. Steenhuis, C. J. Ritsema, and L. W. Dekker. Special Issue of *Geoderma* 70: 117–132.
- Grisak, G. E. and Pickens, J. F. 1980. Solute transport through fractured media. 1. The effect of matrix diffusion. *Water Resour. Res.* 16: 719–730.
- Grisak, G. E. and Pickens, J. F. 1981. An analytical solution for solute transport through fractured media with matrix diffusion. *J. Hydrology* 52: 47–57.
- Heijs, A. W. J., Ritsema, C. J., and Dekker, L. W. 1996. Three-dimensional visualization of preferential flow patterns in two soils, in *Fingered Flow in Unsaturated Soil: From Nature to Model*. Eds. T. S. Steenhuis, C. J. Ritsema, and L. W. Dekker. Special Issue of *Geoderma* 70: 87–101.
- Hendrickx, J. M. H. and Dekker, L. W. 1991. Experimental evidence of unstable wetting fronts in homogeneous non-layered soils, in *Preferential Flow*. Eds. T. J. Gish and A. Shirmohammadi, Proc. Am. Soc. Agricultural Engineers National Symp., December 16–17, 1991, Chicago, IL, pp. 22–31.
- Hill, D. E. and Parlange, J.-Y. 1972. Wetting front instability in layered soils. *Soil Sci. Soc. Am. J.* 36: 697–702.
- Hillel, D. and Baker, R. S. 1988. A descriptive theory of fingering during infiltration into layered soils. *Soil Sci.* 146: 51–56.
- Huang, J. I. and Kim, K. 1993. Fracture process zone development during hydraulic fracturing. *Int. J. Rock Mech. Min. Sci.* 30: 1295–1298.
- Jardin, P. M., Wilson, G. V., and Luxmoore, R. J. 1990. Unsaturated solute transport through a forest soil during rain storm events. *Geoderma* 46: 103–118.
- Jennings, A. A. and Kirkner, D. J. 1984. Instantaneous equilibrium approximation analysis. *J. Hydraul. Eng.* 110: 1700–1717.
- Jury, W.A., Gardner, W. R., and Gardner, W. H. 1991. *Soil Physics*. John Wiley & Sons, New York.

- Jury, W. A. and Flühler, H. 1992. Transport of chemicals through soil: Mechanisms, models and field applications. *Adv. Agron.* 47: 141–201.
- Kluitenberg, G. J. and Horton, R. 1990. Effect of solute application method on preferential transport of solutes in soil. *Geoderma* 46: 283–297.
- Krapp, H. K. and Elrick, D. E. 1969. Density effects in miscible displacement experiments. *Soil Sci.* 107: 372–380.
- Kung, K.-J. S. 1990a. Preferential flow in a sandy vadose soil. 1. Field observations. *Geoderma* 46: 51–58.
- Kung, K.-J. S. 1990b. Preferential flow in a sandy vadose soil. 2. Mechanism and implications. *Geoderma* 46: 59–71.
- Labuz, J. F., Shah, S. P., and Dowding, C. H. 1987. Fracture process zone in granite: Evidence and effect. *Int. J. Rock Mech. Min. Sci.* 24: 235–246.
- Lindstrom, E. T. and Boersma, L. 1971. A theory on the mass transport of previously distributed chemicals in water saturated sorbing porous medium. *Soil Sci.* 111: 192–199.
- Liu, Y., Bierck, B. R., Selker, J. S., Steenhuis, T. S., and Parlange, J.-Y. 1993. High intensity x-ray and tensiometer measurements in rapidly changing preferential flow fields. *Soil Sci. Soc. Am. J.* 57: 1188–1192.
- Liu, Y., Steenhuis, T. S., and Parlange, J.-Y. 1994a. Closed-form solution for finger width in sandy soils at different water contents. *Water Resour. Res.* 30: 949–952.
- Liu, Y., Steenhuis, T. S., and Parlange, J.-Y. 1994b. Formation and persistence of fingered flow fields in coarse grained soils under different moisture contents. *J. Hydrology.* 159: 187–195.
- Maloszewski, P. and Zuber, A. 1990. Mathematical modeling of tracer behavior in short-term experiments in fissured rocks. *Water Resour. Res.* 26: 1517–1528.
- Moreno, L. and Rasmussen, A. 1986. Contaminant transport through a fractured porous rock: Impact of the inlet boundary condition on the concentration profile in the rock matrix. *Water Resour. Res.* 22: 1728–1730.
- Nativ, R., Adar, E., and Becker, A. 1999. A monitoring network for groundwater in fractured media. *Ground Water* 37: 38–47.
- Neretnieks, I., Eriksen, T., and Tahtinen, P. 1982. Tracer movement in a single fissure in granitic rock: Some experimental results and their interpretation. *Water Resour. Res.* 18: 849–858.
- Neretnieks, I. 1983. A note on fracture flow dispersion mechanisms in the ground. *Water Resour. Res.* 19: 364–370.
- Nijssen, B. M., Steenhuis, T. S., Kluitenberg, G. J., Stagnitti, E., and Parlange, J.-Y. 1991. Moving water and solutes through the soil: Testing of a preferential flow model, in *Preferential Flow*. Eds. T. J. Gish and A. Shirmohammadi. Proc. Am. Soc. Agricultural Engineers National Symp., December 16–17, 1991, Chicago, IL, pp. 223–232.
- Nkedi-Kizza, P., Biggar, J. W., van Genuchten, M. Th., Wierenga, P. J., Selim, H. M., Davidson, J. M., and Nielsen, D. R. 1983. Modeling tritium and chloride transport through an aggregated oxisol. *Water Resour. Res.* 19: 691–700.
- Omoti, U. and Wild, A. 1979. Use of fluorescent dyes to mark the pathways of solute movement through soils under leaching conditions. 2. Field experiments. *Soil Sci.* 128: 98–104.
- Parker, J. C. and van Genuchten, M. T. 1984. Determining transport parameters from laboratory and field tracer experiments. *Bulletin 84-3*. Virginia Agricultural Experimental Station, Blacksburg, VI.
- Parker, J. C. and Valocchi, A. J. 1986. Constraints on the validity of equilibrium and first order kinetic transport models in structured soils. *Water Resour. Res.* 22: 399–407.
- Parlange, J.-Y. and Hill, D. E. 1976. Theoretical analysis of wetting front instability in soils. *Soil Sci.* 122: 236–239.
- Parlange, J.-Y., Steenhuis, T. S., Glass, R. J., Pickering, N. B., Waltman, W. J., Bailey, N. O., Andreini, M. S., and Throop, J. A. 1988. The flow of pesticides through preferential paths in soils. *NY Food Life Sci. Q.* 18: 20–23.

- Parlange, J.-Y., Steenhuis, T. S., Stagnitti, F., Simmelman, E., and Nijssen, B. 1996. Recent advances in modelling vadose zone transport, in *Subsurface Water Hydrology*. Eds. V. P. Singh and B. Kumar, Kluwer Academic Publishers, Netherlands, pp. 127–151.
- Parlange, J.-Y., Steenhuis, T. S., Haverkamp, R., Barry, D. A., Culligan, P. J., Hogarth, W. L., Parlange, M. B., and Ross, P. 1998. Soil properties and water movement, in *Vadose Zone Hydrology — Cutting Across Disciplines*. Eds. M. B. Parlange and J. W. Hopmans. Oxford University Press, New York.
- Passioura, J. B. 1971. Hydrodynamic dispersion in aggregate media. 1. Theory. *Soil Sci.* 111: 339–334.
- Philip, J. R. 1975. Stability analysis of infiltration. *Soil Sci. Soc. Am. Proc.* 39: 1042–1049.
- Piquemal, J. 1993. On the modelling conditions of mass transfer in porous media presenting capacitance effects by a dispersion–convection equation for the mobile fluid and a diffusion equation for the stagnant fluid. *Transp. Porous Media* 10: 271–283.
- Polak, A., Nativ, R., and Wallach, R. 2002. Matrix diffusion in chalk formation and its correlation to porosity. *J. Hydrology* 268: 203–213.
- Polak, A., Grader, A. S., Wallach, R., and Nativ, R., 2003a. Chemical diffusion between a fracture and the surrounding matrix: Measurement by computed tomography and modeling. *Water Resour. Res.* 39: 1106–1119.
- Polak, A., Grader, A. S., Wallach, R., and Nativ, R. 2003b. Tracer diffusion from a horizontal fracture into surrounding matrix: measurement by computed tomography. *J. Contam. Hydrol.* 67: 95–112.
- Raats, P. A. C. 1973. Unstable wetting fronts in uniform and nonuniform soils. *Soil Sci. Soc. Am. Proc.* 37: 681–685.
- Rao, P. S. C., Jessup, R. E., Ahuja, L. R., and Davidson, J. 1976. Evaluation of a capillary bundle model for describing solute dispersion in aggregate soils. *Soil Sci. Soc. Am. J.* 40: 815–820.
- Rao, P. S. C., Rolston, D. E., Jessup, R. E., and Davidson, J. M. 1980. Solute transport in aggregated porous media: Theoretical and experimental evaluation. *Soil Sci. Soc. Am. J.* 44: 1139–1146.
- Rice, R. C., Jaynes, D. B., and Bowman, R. S. 1991. Preferential flow of solutes and herbicide under irrigated fields. *Trans. Am. Soc. Agric. Eng.* 34: 914–918.
- Ritsema, C. J. and Dekker, L. W. 1993a. Preferential flow mechanism in a water repellent sandy soil. *Water Resour. Res.* 29: 2183–2193.
- Ritsema, C. J. and Dekker, L. W. 1993b. Soil moisture and dry bulk density patterns in bare dune sands. *J. Hydrology* 154: 107–131.
- Ritsema, C. J. and Dekker, L. W. 1995. Distribution flow: A general process in water repellent soils. *Water Resour. Res.* 31: 1187–1200.
- Ritsema, C. J., Steenhuis, T. S., Parlange, J.-Y., and Dekker, L. W. 1996. Predicted and observed finger diameters in field soils, in *Fingered Flow in Unsaturated Soil: From Nature to Model*. Eds. T.S. Steenhuis, C.J. Ritsema, and L.W. Dekker. Special Issue of *Geoderma* 70: 185–196.
- Roth, K., Jury, W. A., Flühler, H., and Attinger, W. 1991. Transport of chloride through an unsaturated field soil. *Water Resour. Res.* 27: 2533–2541.
- Rubin, J. 1983. Transport of reacting solutes in porous media: Relation between mathematical nature of problems formulation and chemical nature of reaction. *Water Resour. Res.* 19: 1231–1252.
- Selker, J. S., Leclercq, P., Parlange, J.-Y., and Steenhuis, T. S. 1992a. Fingered flow in two dimensions. Part 1. Measurement of matric potential. *Water Resour. Res.* 28: 2513–2521.
- Selker, J. S., Parlange, J.-Y., and Steenhuis, T. S. 1992b. Fingered flow in two dimensions. Part 2. Predicting finger moisture profile. *Water Resour. Res.* 28: 2523–2528.
- Selker, J. S., Steenhuis, T. S., and Parlange, J.-Y. 1993. Wetting front instability in homogeneous sandy soils under continuous infiltration. *Soil Sci. Soc. Am. J.* 56: 1346–1350.
- Seyfried, M. S. and Rao, P. S. C. 1987. Solute transport in undisturbed columns of an aggregated tropical soil: Preferential flow effects. *Soil Sci. Soc. Am. J.* 51: 1434–1443.
- Sidle, R. C., Nilsson, B., Hansen, M., and Fredericia, J. 1998. Spatially varying hydraulic and solute transport characteristics of a fractured till determined by field tracer tests, Funen, Denmark. *Water Resour. Res.* 34: 2515–2527.

- Skopp, J., Gardner, W. R., and Taylor, E. J. 1981. Solute movement in structured soils: Two-regions model with small interaction. *Soil Sci. Soc. Am. J.* 45: 837–842.
- Skopp, J. and Gardner, W. R. 1992. Miscible displacement: An interacting flow region model. *Soil Sci. Soc. Am. J.* 56: 1680–1686.
- Smettem, K. R. J. and Collis-George, N. 1985. Statistical characterization of soil biopores using a soil peel method. *Geoderma*. 36: 27–36.
- Sposito, G., Jury, W. A., and Gupta, V. K. 1986. Fundamental problems in the stochastic convective dispersion model of solute transport in aquifers and field soils. *Water Resour. Res.* 22:77–88.
- Stagnitti, F., Steenhuis, T. S., Parlange, J.-Y., Nijssen, B. M., and Parlange, M. B. 1991. Preferential solute and moisture transport in hillslopes. *Inst. Engrs, Aust.* 22: 919–924.
- Stagnitti, F., Parlange, J.-Y., Steenhuis, T. S., Boll, J., Pivetz, B., and Barry, D. A. 1995. Transport of moisture and solutes in the unsaturated zone by preferential flow, in *Environmental Hydrology*. Ed. V.P. Singh. Kluwer Academic Publishers, Netherlands, 193–224.
- Starr, J. L. and Parlange, J.-Y. 1977. Plate-induced tailing in miscible displacement experiments. *Soil Sci.* 124: 56–60.
- Starr, J. L., DeRoo, H. C., Frink, C. R., and Parlange, J.-Y. 1978. Leaching characteristics of a layered field soil. *Soil Sci. Soc. Am. J.* 42: 376–391.
- Steenhuis, T. S., Parlange, J.-Y., and Andreini, M. S. 1990. A numerical model for preferential solute movement in structured soils. *Geoderma* 46: 193–208.
- Steenhuis, T. S., Boll, J., Shalit, G., Selker, J. S., and Merwin, I. A. 1994. A simple equation for predicting preferential flow solute concentrations. *J. Env. Quality.* 23: 1058–1064.
- Steenhuis, T. S., Vandenheuvel, K., Weiler, K. W., Boll, J., Dalipathy, J., Herbert, S., and Kung, K.-J. S. 1998. Mapping and interpreting soil textural layers to assess agri-chemical movement at several scales along the eastern seaboard (USA). *Nutr. Cycling Agroecosyst.* 50: 91–98.
- Stehfest, H. 1970. Numerical inversion of Laplace transforms. *J. Assoc. Comput. Mach.* 13: 47–49, 1970.
- Sudicky, E. A. and Frind, E. O. 1982. Contaminant transport in fractured porous media: analytical solutions for a system of parallel fractures. *Water Resour. Res.* 18: 1634–1642.
- Tang, D. H., Frind, E. O., and Sudicky, E. A. 1981. Contaminant transport in fractured porous media: Analytical solution for a single fracture. *Water Resour. Res.* 17: 555–564.
- Valocchi, A. J. 1985. Validity of the local equilibrium assumption for modeling sorbing solute transport through homogeneous soils. *Water Resour. Res.* 21: 808–820.
- Van Duin, C. J. and van der Zee, S. E. A. T. M. 1986. Solute transport parallel to an interface separating two different porous materials. *Water Resour. Res.* 22: 1779–1789.
- van Genuchten, M. Th. and Wierenga, P. J. 1977. Mass transfer studies in sorbing porous media: Experimental evaluation with tritium. *Soil Sci. Soc. Am. J.* 41: 278–285.
- van Genuchten, M. Th. and Alves, W. J. 1982. Analytical solutions of the one-dimensional convection dispersion solute transport equation. *Tech. Bull.* 1661. U.S. Department of Agriculture, pp. 1–151.
- van Genuchten, M. Th. and Dalton, F. N. 1986. Models for simulating salty movement in aggregated field soils. *Geoderma* 38: 165–183.
- Wallach, R. 1998. A small perturbations solution for nonequilibrium chemical transport through soils with relatively high desorption rate. *Water Resour. Res.* 34: 149–154.
- Wallach, R. and Parlange, J.-Y. 1998. Modeling transport in a single crack by the dual-porosity concept with a boundary layer at the interface. *J. Contam. Hydrol.* 34: 121–138.
- Wallach, R. and Steenhuis, T. S. 1998. A model for non-reactive solute transport in structured soils with continuous preferential flow paths. *Soil Sci. Soc. Am. J.* 123: 734–737.
- Wallach, R. and Parlange, J.-Y. 2000. Applying the boundary layer concept to model transport of dissolved chemicals in preferential flow paths. *Water Resour. Res.* 36: 2845–2851.
- Wang, D., McSweeney, K., Norman, J. M., and Lowery, B. 1991. Preferential flow in soils with ant burrows, in *Preferential Flow*. Eds. T. J. Gish and A. Shirmohammadi. Proc. Am. Soc. Agricultural Engineers National Symp., December 16–17, 1991, Chicago, IL, pp. 183–191.

- Wang, D., Norman, J. M., Lowery, B., and McSweeney, K. 1994. Nondestructive determination of hydro geometrical characteristics of soil macropores. *Soil Sci. Soc. Am. J.* 58: 294–303.
- Warren, J. E., and Root, P. J. 1963. The behaviour of naturally fractured reservoirs. *Soc. Petrol. Eng. J.* 3: 245–255.
- Weber, W. J. Jr. and DiGiano, F. A. 1996. *Process Dynamics in Environmental Systems*, John Wiley & Sons, Inc., New York.
- White, R. E., Thomas, G. W., and Smith, M. S. 1984. Modeling water flow through undisturbed soil cores using a transfer function model derived from ^{3}HOH and Cl transport. *J. Soil Sci.* 35: 159–168.
- White, R. E., Dyson, J. S., Gerstl, Z., and Yaron, B. 1986. Leaching of herbicides in the undisturbed cores of a structured clay soil. *Soil Sci. Soc. Am. J.* 50: 277–283.

Further Information

Preferential flow is a fast-developing science. Summaries of the work can be found in conference proceedings and special issues of journals. A special issue of *Geoderma* (Steenhuis et al., 1996) covers the range from field experiments, where preferential flow was found, to theoretical development. In an older publication, Gish and Shirmohammadi (1991) summarize the preferential flow research up to 1991. A very good overview article about the occurrence of preferential flow was written by Flury (1996). The other references on the different types of preferential flow are mentioned in the text.

- Gish, T. J. and Shirmohammadi, A., Eds. 1991. *Preferential Flow*. Proceedings of American Society of Agricultural Engineers National Symposium, December 16–17, 1991, Chicago, IL.
- Flury, M. 1996. Experimental evidence of transport of pesticides through field soils. A review. *J. Env. Qual.* 25: 25–45.
- Steenhuis, T. S., Ritsema, C. J., and Dekker, L. W., Eds. 1996. *Fingered Flow in Unsaturated Soil: From Nature to Model*. Special Issue of *Geoderma*, 70.
- Haws, N. W., Rao, P. S. C., Simunek, J., and Poyer, I. C. 2005. Single-porosity and dual-porosity modeling of water flow and solute transport in subsurface-drained fields using effective field-scale parameters. *J. Hydrology*. 313: 257–273.

10

Well Hydraulics and Aquifer Tests

10.1	Introduction.....	10-1
	Aquifer Types • Physical Properties	
10.2	Well-Flow Equations.....	10-4
	Confined Aquifers • Unconfined Aquifers • Leaky Aquifers • Partial Penetration • Recovery Well-Flow Equations	
10.3	Performance of an Aquifer Test	10-12
	Measurements • Duration of a Pumping Test • Processing the Data • Interpretation of the Data	
10.4	Time–Drawdown Analyses	10-17
	Aquifer Test in a Confined Aquifer • Aquifer Test in an Unconfined Aquifer • Aquifer Test in a Leaky Aquifer • Single-Well Tests	
10.5	Distance–Drawdown Analyses	10-26
	Confined and Unconfined Aquifers • Leaky Aquifers	
10.6	Interpretation of Heterogeneous Aquifers	10-30
10.7	Concluding Remarks	10-31
	Glossary.....	10-31
	References	10-32
	Further Information	10-33
	Appendix 10.1	10-34

Hans Boonstra and
Richard Soppe
Alterra-ILRI, The Netherlands

10.1 Introduction

Estimating the physical properties of water-bearing layers is an essential part of groundwater studies. One of the most effective ways of determining these properties is to conduct and analyze aquifer tests. The principle of an aquifer test is simple: water is pumped from a well tapping the *aquifer* and the discharge of the well and the changes in water levels in the well and in *piezometers* at known distances from the well are measured. The change in water level induced by the pumping is known as the *drawdown*. In the literature, aquifer tests based on the analysis of drawdowns during pumping are commonly referred to as “pumping tests.”

The physical properties can also be found from a recovery test. In such a test, a well that has been discharging for some time is shut down, and thereafter the recovery of the aquifer’s hydraulic head is measured in the wells; the change in water level during this recovery period is known as the *residual drawdown*.

Owing to the high costs of aquifer tests, the number that can be performed in most groundwater studies has to be restricted. Nevertheless, one can perform an aquifer test without using piezometers, thereby cutting costs, although one must then accept a certain, sometimes appreciable, error. To distinguish such tests from normal aquifer tests, they are called single-well tests. In these tests, measurements are only taken inside the pumped well.

Analyzing and evaluating aquifer-test data is as much an art as a science. It is a science as it is based on theoretical models that the geologist or engineer must understand and on thorough investigations that he must conduct into the geological formations in the area of the test. It is an art as different types of aquifers can exhibit similar drawdown behaviors that demand interpretational skills on the part of the geologist or engineer.

The well-flow equations presented in Section 10.2 are from well hydraulics. Any lengthy derivations of the equations have been omitted as these can be found in the original publications listed in the references. The equations are presented in their final form, emphasizing the assumptions and conditions that underlie them. The essentials of how to conduct an aquifer test are summarized in Section 10.3. The remaining sections deal with the analysis and evaluation of aquifer-test data from a variety of aquifer types and analysis methods, outlining the procedures that are to be followed for their successful application. All these procedures are illustrated with field data.

10.1.1 Aquifer Types

In the analysis and evaluation of aquifer-test data, three main types of aquifer are distinguished: the confined, unconfined, and leaky aquifer (Figure 10.1).

A confined aquifer is a completely saturated aquifer whose upper and lower boundaries are *aquiclude*s. In confined aquifers the pressure of the water is usually higher than that of the atmosphere and the water level in wells tapping such aquifers stands above the top of the aquifer (see Figure 10.1a). The piezometric surface is the imaginary surface to which the water will rise in wells penetrating the aquifer. When the water level in wells tapping such aquifers stands above the ground surface, they are called free-flowing wells or artesian wells.

An unconfined aquifer is a partly saturated aquifer bounded below by an aquiclude and above by the free water table or phreatic surface (see Figure 10.1b). At the free water table, the pressure of the groundwater equals that of the atmosphere. The water level in a well penetrating an unconfined aquifer does not, in general, rise above the water table, except when there is a vertical flow.

A leaky aquifer, also known as a semi-confined aquifer, is a completely saturated aquifer that is bounded below by an aquiclude and above by an *aquitard*. The overlying aquitard may be partly saturated when it extends to the land surface (see Figure 10.1c) or is fully saturated when it is overlain by an unconfined aquifer that is bounded above by the water table (see Figure 10.1d). The piezometric level in a well tapping a leaky aquifer may coincide with the water table if there is a hydrologic equilibrium between the aquifer's recharge and discharge; it may rise or fall below the water table in areas with upward or downward flow, in other words, in discharge or recharge areas.

A multilayered aquifer is a special case of the leaky aquifer. The aquiclude in Figure 10.1c and Figure 10.1d are then aquitards with other aquifers occurring at larger depths. In deep sedimentary basins such an interbedded system of permeable and less permeable layers is very common.

10.1.2 Physical Properties

This section summarizes the physical properties and derived parameters of aquifers and aquitards that appear in the various well-flow equations.

The hydraulic conductivity is the constant of proportionality in Darcy's law and is defined as the volume of water that will move through a porous medium in unit time under a unit hydraulic gradient through a unit area measured at right angles to the direction of flow. The hydraulic conductivity of an aquifer

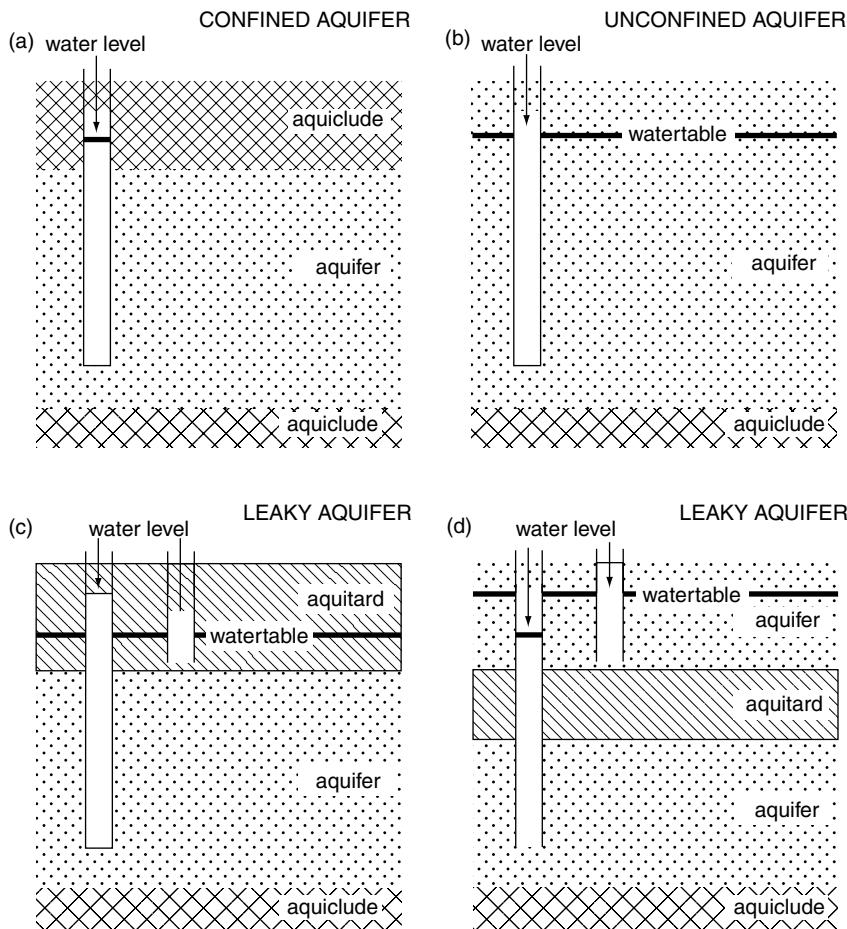


FIGURE 10.1 Different types of aquifers.

is denoted by the symbol K . Common orders of magnitude of the hydraulic conductivities for different aquifer materials range from 10^{-8} m/d for clay deposits to 10^3 m/d for coarse sands.

For confined aquifers, the saturated thickness is equal to the physical thickness of an aquifer between the aquiclude above and below it (see Figure 10.1a). The same is true for the confined parts of a leaky aquifer bounded by an aquitard and an aquiclude (see Figures 10.1c and 10.1d). In both these cases, the saturated thickness is a constant.

For unconfined aquifers, the saturated thickness is equal to the difference between the free water table and the aquiclude (see Figure 10.1b). As the water table changes its position with time, the saturated thickness of an unconfined aquifer is not constant, but variable. Whether constant or variable, the saturated thickness of an aquifer is denoted by the symbol D . Its order of magnitude can range from several meters to hundreds or even thousands of meters.

For aquitards in leaky aquifers, the saturated thickness can be variable or constant. In Figure 10.1c, the aquitard is partly saturated and has a free water table. Its saturated thickness depends upon the position of the water table. In Figure 10.1d, the aquitard is bounded by two aquifers and is fully saturated. Its saturated thickness is physically determined and thus constant. The saturated thickness of an aquitard is denoted by the symbol D' . It may range from a few meters to hundreds of meters.

The transmissivity is the product of the average hydraulic conductivity and the saturated thickness of the aquifer. Consequently, the transmissivity is the rate of flow under a hydraulic gradient equal to unity through a cross-section of unit width over the whole saturated thickness of the water-bearing

layer. The transmissivity of an aquifer is denoted by the symbol T . Its order of magnitude can be derived from those of K and D , and may range from hundreds to tens of thousands of square meter per day.

The specific storage of a saturated confined aquifer is the volume of water that a unit volume of aquifer releases from storage under a unit decline in head. This release of water under conditions of decreasing hydraulic head stems from the compaction of the aquifer due to increasing effective stress and from the expansion of the water due to decreasing water pressure. The specific storage depends on the elasticity of both the aquifer material and the water. For a certain location it can be regarded as a constant. It is denoted by the symbol S_s ; its order of magnitude is 10^{-6} to 10^{-4} d^{-1} .

The storativity of a saturated confined aquifer of thickness D is defined as the volume of water released from storage per unit surface area of the aquifer per unit decline in the component of hydraulic head normal to that surface. In a vertical column of unit area extending through the confined aquifer, the storativity equals the volume of water released from the aquifer when the piezometric surface drops over a unit distance. Storativity is thus defined as $S = S_s D$. The storativity of a saturated aquifer is a function of its thickness. Storativity is a dimensionless quantity, as it involves a volume of water per volume of aquifer. Its values in confined aquifers range from 5×10^{-5} to 5×10^{-3} .

The specific yield is the volume of water that an unconfined aquifer releases from storage per unit surface area of aquifer per unit decline of the water table. In unconfined aquifers, the effects of the elasticity of the aquifer material and the water are negligible, except for a short time after the start of pumping. Values of the specific yield of unconfined aquifers are much higher than the storativities of confined aquifers. The orders of magnitude of the specific yield denoted by the symbol S_y range from 0.01 for clay deposits to 0.47 for coarse sands. Specific yield is also referred to as effective porosity, unconfined storativity, or drainable pore space. Small interstices do not contribute to the effective porosity, as the retention forces in them are greater than the weight of water. Hence, no groundwater will be released from these small interstices by gravity drainage.

The hydraulic resistance characterizes the resistance of an aquitard to vertical flow, either upward or downward. It is the ratio of the saturated thickness of the aquitard D' and its hydraulic conductivity for vertical flow K' and is thus defined as $c = D'/K'$. Its order of magnitude may range from a few tens to thousands of days. Aquitards having c -values of 2500 d or more are regarded to act as aquiclude, although theoretically an aquiclude has an infinitely high c -value.

The leakage factor describes the spatial distribution of leakage through one or two aquitards into a leaky aquifer or vice versa. It is the square root of the product of the aquifer's transmissivity and the aquitard's hydraulic resistance and is thus defined as $L = \sqrt{(Tc)}$; it is expressed in meters. Large values of L originate from a high transmissivity of the aquifer and a high hydraulic resistance of the aquitard. In both cases, the influence of leakage will be small and the area over which leakage takes place, large.

10.2 Well-Flow Equations

The well-flow equations presented in this section were developed under the following common assumptions and conditions (1) the aquifer has a seemingly infinite areal extent; (2) the aquifer is *homogeneous, isotropic*, and of uniform thickness over the area influenced by the test; (3) prior to pumping, the hydraulic head is horizontal (or nearly so) over the area that will be influenced by the test; (4) the pumped well penetrates the entire thickness of the aquifer and thus receives water by horizontal flow; (5) the aquifer is pumped at a constant-discharge rate; (6) the water removed from storage is discharged instantaneously with decline of head; and (7) the diameter of the pumped well is small, that is, the storage inside the well can be neglected.

The above mentioned simplifying assumptions and limiting conditions cannot always be fulfilled in the field. In these cases, special well-flow equations for bounded, sloping, or anisotropic aquifers or for large-diameter wells can be used; these can be found in Kruseman and de Ridder (1994).

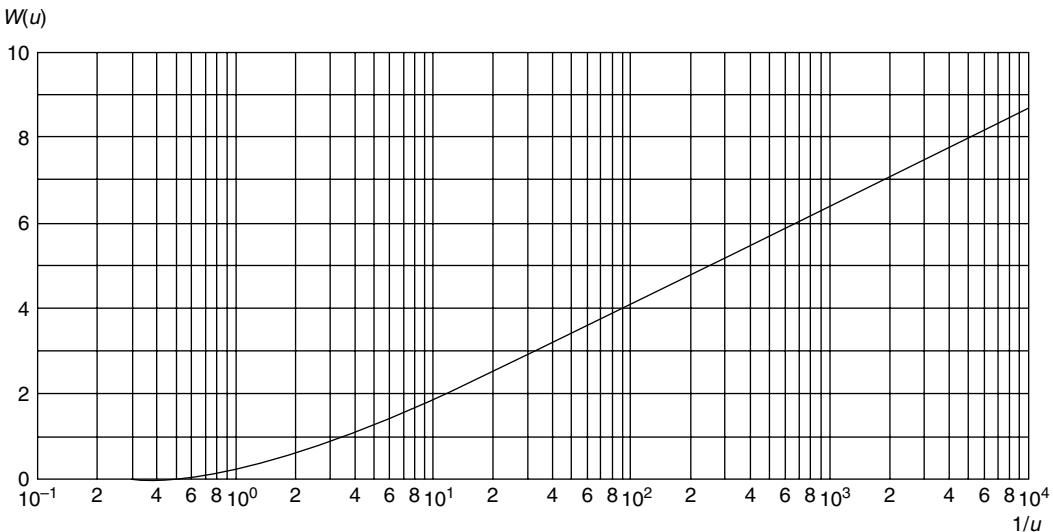


FIGURE 10.2 Theis well function $W(u)$ vs. $1/u$ for fully-penetrating wells in confined aquifers.

10.2.1 Confined Aquifers

Theis (1935) was the first to develop an equation for unsteady-state flow that introduced the time factor and the storativity. He noted that when a fully penetrating well pumps an extensive confined aquifer at a constant rate, the influence of the discharge extends outward with time. The rate of decline of head, multiplied by the storativity and summed over the area of influence, equals the discharge. The Theis equation, which was derived from the analogy between the flow of groundwater and the conduction of heat, is written as

$$s(r, t) = \frac{Q}{4\pi T} \int_u^\infty \frac{e^{-y}}{y} dy = \frac{Q}{4\pi T} W(u) \quad (10.1)$$

with

$$u = \frac{r^2 S}{4Tt} \quad (10.2)$$

where $s(r, t)$ is the drawdown in meters measured in a well, Q is the constant well discharge in m^3/d , T is the transmissivity of the aquifer in m^2/d , $W(u)$ is the dimensionless Theis well function, r is the distance in meters from the pumped well, S is the dimensionless storativity of the aquifer, and t is the time in days since pumping started. In Figure 10.2, the Theis well function $W(u)$ is plotted vs. $1/u$ on semi-log paper. This figure shows that, for large values of $1/u$, the Theis well function exhibits a straight-line segment. The Jacob method is based on this phenomenon. Cooper and Jacob (1946) showed that, for the straight-line segment, Equation 10.1 can be approximated by

$$s(r, t) = \frac{2.3Q}{4\pi T} \log \frac{2.25Tt}{r^2 S} \quad (10.3)$$

with an error less than 1%, 2%, 5%, and 10% for $1/u$ larger than 30, 20, 10, and 7, respectively.

In most handbooks on this subject, the condition to use Equation 10.3 is taken as $1/u > 100$. From personal experience, this limiting condition can often be relaxed to $1/u > 10$.

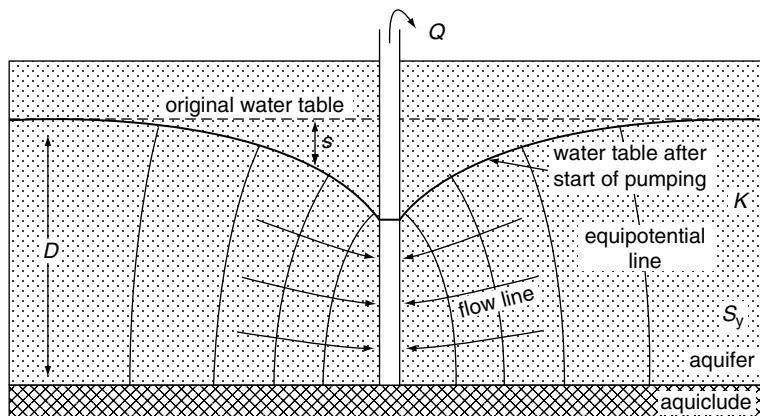


FIGURE 10.3 Schematic cross-section of a pumped unconfined aquifer.

When a confined aquifer is pumped, the cone of depression will continuously deepen and expand. Even at late pumping times, the water levels in the piezometers will never stabilize to a real steady state. Although the water levels continue to drop, the cone of depression will eventually deepen uniformly over the area influenced by the pumping. At that stage, the hydraulic gradient has become constant; this phenomenon is called pseudo-steady state. For this situation, Thiem (1906), using two or more piezometers, developed an equation, the so-called Thiem–Dupuit equation that can be written as

$$s(r_1, t) - s(r_2, t) = \frac{2.3Q}{2\pi T} \log \frac{r_2}{r_1} \quad (10.4)$$

Equation 10.4 can also be derived by applying Equation 10.3 to two piezometers at distances r_1 and r_2 at large times.

10.2.2 Unconfined Aquifers

Figure 10.3 shows a pumped unconfined aquifer underlain by an aquiclude. These are the following basic differences between unconfined and confined aquifers when they are pumped: (1) an unconfined aquifer is partly dewatered during pumping resulting in a decreasing saturated aquifer thickness, whereas a confined aquifer remains fully saturated, (2) the water produced by a well in an unconfined aquifer comes from the physical dewatering of the aquifer, whereas in a confined aquifer it comes from the expansion of the water in the aquifer due to a reduction of the water pressure (although in some hydrological theories, water is assumed to be an incompressible fluid), and from the compaction of the aquifer due to increased effective stresses, and (3) the flow toward a well in an unconfined aquifer has clear vertical components near the pumped well, whereas there are no such vertical flow components in a confined aquifer, provided, of course, that the well is fully penetrating.

In an unconfined aquifer, the water levels in piezometers near the well often tend to decline at a slower rate than that described by the Theis equation. Time–drawdown curves on semi-log paper therefore usually show a typical S shape: a relatively steep early-time segment, a flat intermediate segment, and a relatively steep segment again at later times as is shown in Figure 10.4. Nowadays, the widely used explanation of this S-shaped time–drawdown curve is based on the concept of delayed yield (Neuman, 1972). It is caused by a time lag between the early elastic response of the aquifer and the subsequent downward movement of the water table due to gravity drainage.

During the early stage of an aquifer test — a stage that may last for only a few minutes — the discharge of the pumped well is derived uniquely from the elastic storage within the aquifer. Hence, the reaction of the unconfined aquifer immediately after the start of pumping is similar to the reaction of a confined aquifer as described by the flow equation of Theis.

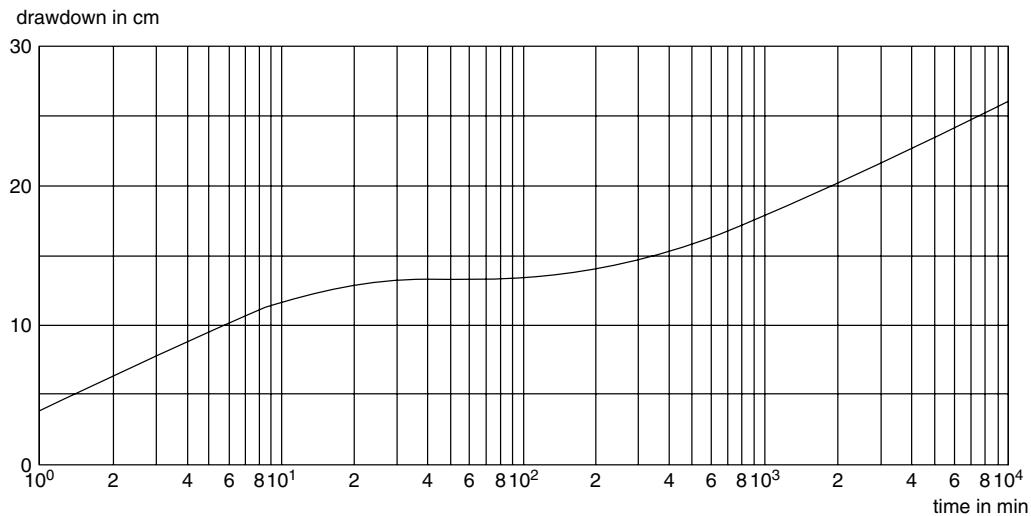


FIGURE 10.4 Time–drawdown plot in an unconfined aquifer showing delayed yield.

Only later the water table starts to fall and the effect of the delayed yield becomes apparent. The influence of the delayed yield is comparable to that of leakage: the average drawdown slows down with time and no longer conforms to the Theis curve. After a few minutes to a few hours of pumping, the time–drawdown curve approaches a horizontal position.

The late-time segment of the time–drawdown curve may start from several minutes to several days after the start of pumping. The declining water table can now keep pace with the increase in the average drawdown. The flow in the aquifer is essentially horizontal again and as in the early pumping time the time–drawdown curve approaches the Theis curve.

Jacob (1950) showed that if the drawdowns in an unconfined aquifer are small compared to the initial saturated thickness of the aquifer, the condition of horizontal flow toward the well is approximately satisfied, so that Equation 10.1 to Equation 10.3 can also be applied to determine the physical properties. The only changes required are that the storativity S be replaced by the specific yield S_y of the unconfined aquifer, and that the transmissivity T be defined as the transmissivity of the initial saturated thickness of the aquifer.

When the drawdowns in an unconfined aquifer are large compared with the aquifer's original saturated thickness, the observed drawdowns need to be corrected before this equation can be used. Jacob (1944) proposed the following correction

$$s_c(r, t) = s(r, t) - \frac{s^2(r, t)}{2D} \quad (10.5)$$

where $s_c(r, t)$ is the corrected drawdown in meters, $s(r, t)$ is the observed drawdown in meters, and D is the saturated aquifer thickness in meters prior to pumping. This correction is only needed when the maximum drawdown at the end of the test is larger than 5% of the original saturated aquifer thickness.

As with confined aquifers, the cone of depression will continuously deepen and expand in unconfined aquifers. For pseudo-steady state conditions, Equation 10.4 can also be used to describe the drawdown behavior in two piezometers in an unconfined aquifer, provided that the observed drawdowns are corrected according to Equation 10.5.

10.2.3 Leaky Aquifers

When a leaky aquifer is pumped (Figure 10.5), the piezometric level of the aquifer in the well is lowered. This lowering spreads radially outward as pumping continues, creating a difference in hydraulic head

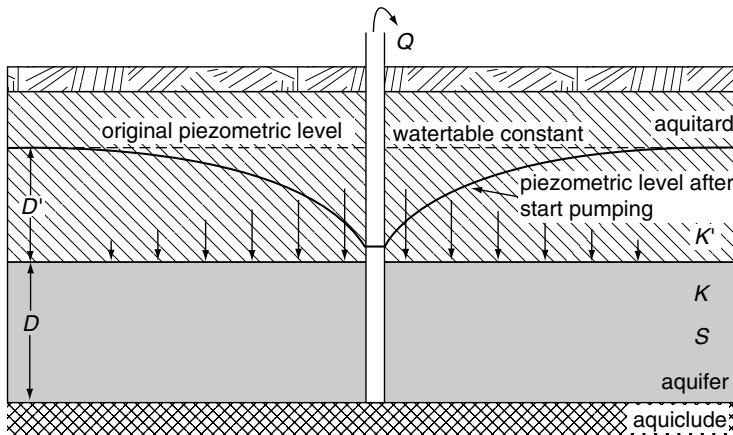


FIGURE 10.5 Schematic cross-section of a pumped leaky aquifer.

between the aquifer and the aquitard. Consequently, the groundwater in the aquitard will start moving vertically downward to join the water in the aquifer. The aquifer is thus partially recharged by downward percolation from the aquitard. As pumping continues, the percentage of the total discharge derived from this percolation increases. After a certain period of pumping, equilibrium will be established between the discharge rate of the pump and the recharge rate by vertical flow through the aquitard. This steady state will be maintained as long as the water table in the aquitard is kept constant.

According to Hantush and Jacob (1955), the drawdown due to pumping a leaky aquifer can be described by the following equation

$$s(r, t) = \frac{Q}{4\pi T} \int_u^{\infty} \frac{1}{y} \exp\left(-y - \frac{r^2}{4L^2 y}\right) dy = \frac{Q}{4\pi T} W(u, r/L) \quad (10.6)$$

where $W(u, r/L)$ is the dimensionless Hantush well function, $L (= \sqrt{Tc})$ is the leakage factor or characteristic length in meters, $c (= D'/K')$ is the hydraulic resistance of the aquitard in d, D' is the saturated thickness of the aquitard in meters, K' is the vertical hydraulic conductivity of the aquitard in m/d, and the other symbols as defined earlier.

Equation 10.6 has the same form as the Theis equation (Equation 10.1), but there are two parameters in the integral: u and r/L . Equation 10.6 approaches the Theis equation for large values of L , when the exponential term $r^2/(4L^2 y)$ approaches zero. In Figure 10.6, the Hantush well function $W(u, r/L)$ is plotted vs. $1/u$ on semi-log paper for an arbitrary value of r/L . This figure shows that the Hantush well function exhibits an S shape and, for large values of $1/u$, a horizontal straight-line segment indicating steady-state flow.

For the steady-state drawdown in a leaky aquifer, De Glee (1930, 1951) derived the following equation

$$s_m(r) = \frac{Q}{2\pi T} K_0\left(\frac{r}{L}\right) \quad (10.7)$$

where $s_m(r)$ is the steady-state, stabilized drawdown in meters and $K_0(r/L)$ is the dimensionless modified Bessel function of the second kind and of zero order (Hankel function). Hantush (1956, 1964) noted that if r/L is small ($r/L < 0.05$) and $L > 3D$, Equation 10.7 can, for all practical purposes, be approximated by

$$s_m(r) = \frac{2.3Q}{2\pi T} \log \frac{1.12L}{r} \quad (10.8)$$

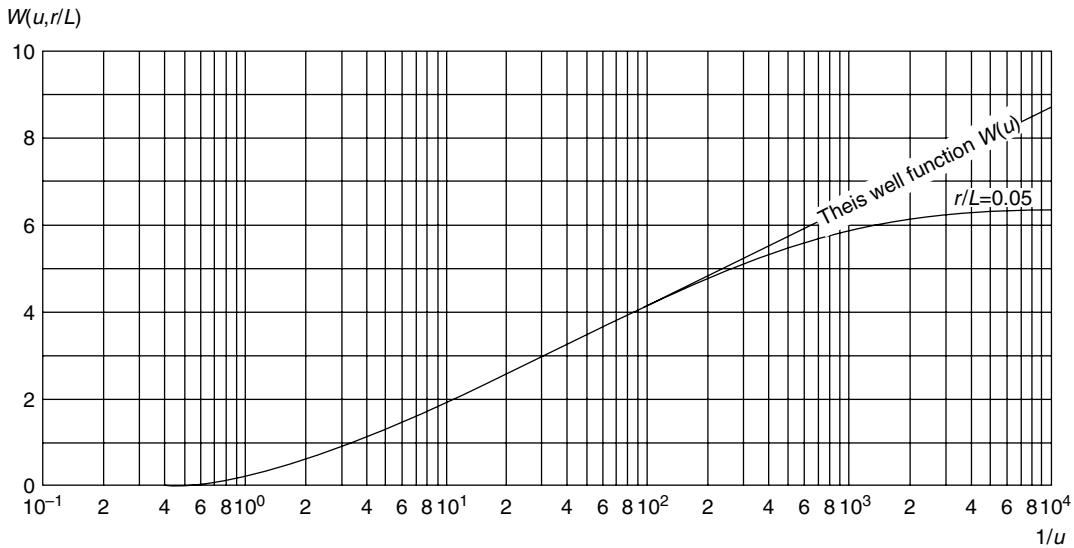


FIGURE 10.6 Hantush well function $W(u, r/L)$ vs. $1/u$ for fully penetrating wells in leaky aquifers.

It is important to note that the flow system in a pumped leaky aquifer consists of a vertical component in the overlying aquitard and a horizontal component in the aquifer. In reality, the flow lines in the aquifer are not horizontal but curved (i.e., there are both vertical and horizontal-flow components in the aquifer). The above equations can therefore only be used when the vertical-flow component in the aquifer is so small compared to the horizontal-flow component that it can be neglected. In practice, this condition is fulfilled when $L > 3D$.

10.2.4 Partial Penetration

When an aquifer is pumped by a partially penetrating well, the assumption that the well receives water from horizontal flow is no longer valid. Hence, the previous equations cannot be used to describe the flow to pumped wells. Due to a contraction of flow lines, partial penetration causes the flow velocity in the immediate vicinity of the well to be higher than it would be otherwise, leading to an extra loss of head. This effect is strongest at the well face, and decreases with increasing distance from the well. According to Hantush (1962) the drawdown due to pumping in a confined aquifer can be described by the following equation

$$s(r, t) = \frac{Q}{4\pi T} [W(u) + f_s] \quad (10.9)$$

and

$$f_s = \frac{2D}{\pi(b-d)} \sum_{n=1}^{\infty} \left(\frac{1}{n} \right) W \left(u, \frac{n\pi r}{D} \right) \times \left[\sin \left(\frac{n\pi b}{D} \right) - \sin \left(\frac{n\pi d}{D} \right) \right] \left(\cos \frac{n\pi z}{D} \right) \quad (10.10)$$

where $W(u, n\pi r/D)$ is the dimensionless Hantush well function, b and b' are the penetration depths in meters of the pumped well and of the piezometer, d and d' are the non-screened parts in meters of the pumped well and of the piezometer, and $z = (b' + d')/2$ (see Figure 10.7).

In Figure 10.8, the expression $W(u) + f_s$ is plotted vs. $1/u$ on semi-log paper. This figure shows that, for large values of $1/u$, this expression exhibits a straight-line segment. For this segment Equation 10.10

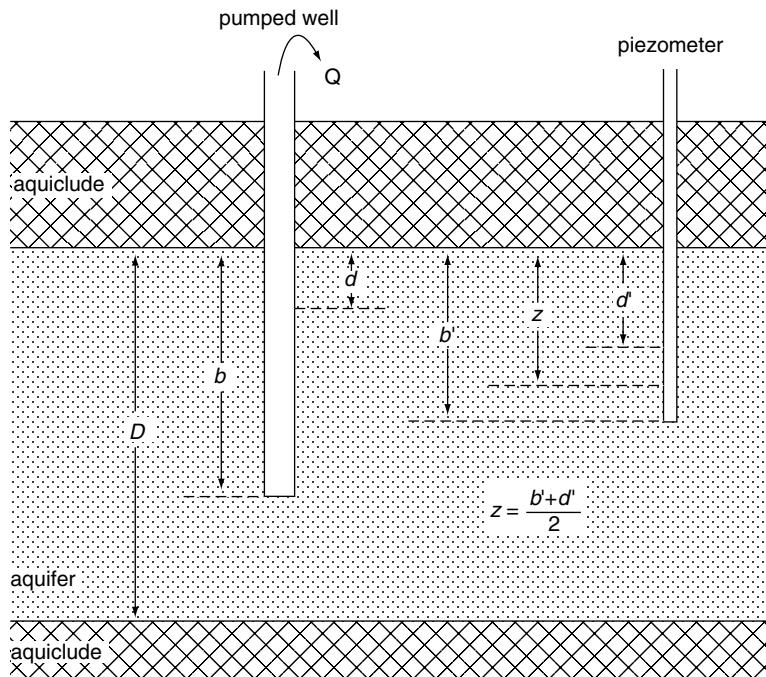


FIGURE 10.7 Schematic illustration of the parameters used for the analysis of partially penetrating wells.

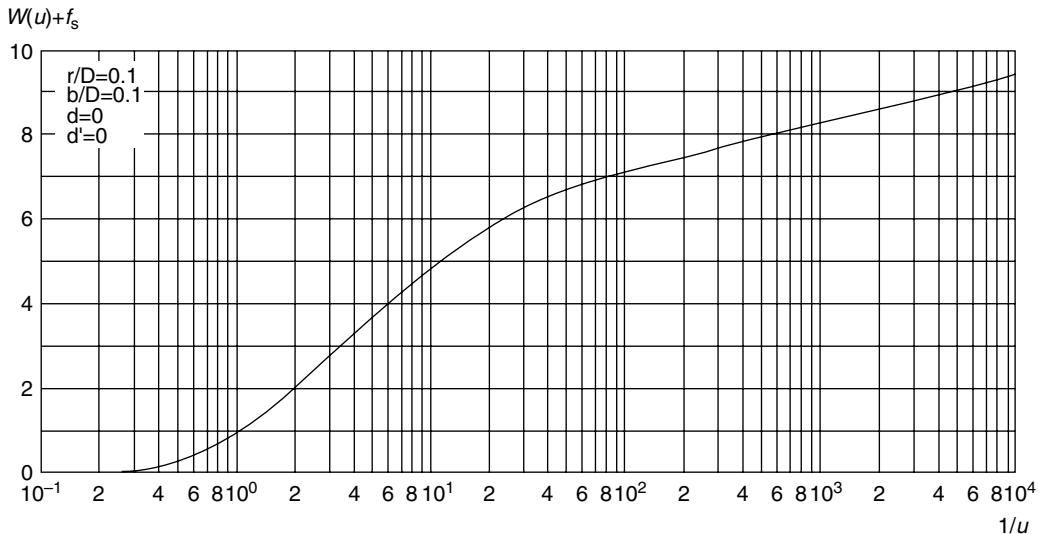


FIGURE 10.8 Hantush well function $W(u) + f_s$ vs. $1/u$ for partially penetrating wells in confined aquifers.

reduces to

$$f_s = \frac{4D}{\pi(b-d)} \sum_{n=1}^{\infty} \left(\frac{1}{n} \right) K_0 \left(\frac{n\pi r}{D} \right) \left[\sin \left(\frac{n\pi b}{D} \right) - \sin \left(\frac{n\pi d}{D} \right) \right] \left[\cos \left(\frac{n\pi z}{D} \right) \right] \quad (10.11)$$

For a particular well/piezometer configuration Equation 10.11 yields constant values for f_s . The Hantush's modification of the Jacob equation was based on this phenomenon. Hantush (1962) showed that for the

straight-line segment Equation 10.9 can be approximated by

$$s(r, t) = \frac{2.3Q}{4\pi T} \log \frac{2.25Tt}{r^2 S} e^{f_s} \quad (10.12)$$

Equation 10.12 can be used to describe the drawdown behavior in a confined aquifer with a partially penetrating well, provided that $t > D^2 S / 2T$. Equation 10.12 can also be applied to unconfined aquifers. It is then assumed that due to partial penetration the drawdowns are small compared with the initial saturated thickness of the aquifer. The only changes required are that the storativity S be replaced by the specific yield S_y of the unconfined aquifer, and that the transmissivity T be defined as the transmissivity of the initial saturated thickness of the aquifer.

For leaky aquifers, the drawdown behavior can be described by Weeks (1969)

$$s(r, t) = \frac{Q}{4\pi T} [W(u, r/L) + f_s] \quad (10.13)$$

The pumping time should also be long to apply Equation 10.13 ($t > D^2 S / 2T$).

10.2.5 Recovery Well-Flow Equations

The well-flow equations describing the drawdown behavior during the recovery period are based on the principle of superposition. Applying this principle, it is assumed that, after the pump has been shut down, the well continues to be pumped at the same discharge as before, and that an imaginary recharge, equal to the discharge, is injected into the well. The recharge and discharge thus cancel each other, resulting in an idle well as is required for the recovery period. For any of the well-flow equations presented in the previous sections, a corresponding recovery equation can be formulated.

On the basis of Equation 10.1, Theis (1935) developed his recovery method for confined aquifers. After a constant-rate pumping test, the residual drawdown during the recovery period is given by

$$s'(r, t/t') = \frac{Q}{4\pi T} \{W(u) - W(u')\} \quad (10.14)$$

where $s'(r, t)$ is the residual drawdown in meters, $u' = r^2 S' / (4Tt')$, S' is the dimensionless storativity of the aquifer during recovery, and t' is the time in days since pumping stopped. In Figure 10.9, the Theis recovery well function $W(u) - W(u')$ is plotted vs. u'/u on semi-log paper. This figure shows that, for small values of u'/u , the expression exhibits a straight-line segment. For this segment, Equation 10.14 can be approximated by

$$s'(r, t/t') = \frac{2.3Q}{4\pi T} \log \frac{S't}{S't'} \quad (10.15)$$

Neuman (1975) showed that Equation 10.15 can also be used to describe the residual drawdown behavior in unconfined aquifers, as its delayed water table response to pumping is fully reversible (no hysteresis effects). The only changes required are that the storativity S be replaced by the specific yield S_y of the unconfined aquifer, and that the transmissivity T be defined as the transmissivity of the initial saturated thickness of the aquifer.

For leaky aquifers, the principle of superposition can also be applied. This results in the following equation

$$s'(r, t') = \frac{Q}{4\pi T} \left[W\left(u, \frac{r}{L}\right) - W\left(u', \frac{r}{L}\right) \right] \quad (10.16)$$

Equation 10.16 can only be solved by numerical simulation. Hantush (1964) showed that, if pumping and recovery times are short, that is, if $t + t' < (L^2 S) / 20T$ or $t + t' < cS / 20$, Equation 10.14 can also be used.

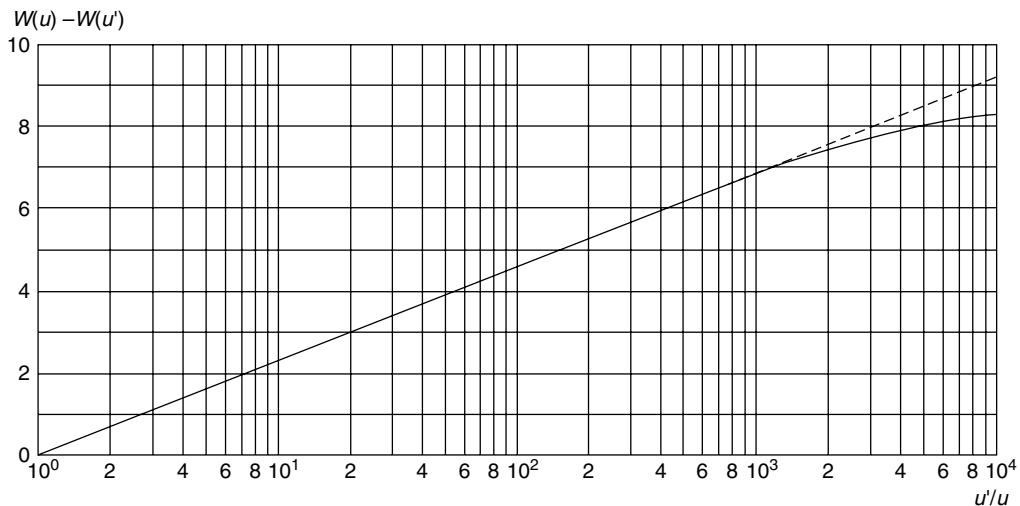


FIGURE 10.9 Theis recovery well function $W(u) - W(u')$ vs. u'/u for fully penetrating wells in confined aquifers.

10.3 Performance of an Aquifer Test

The changes of water levels observed in both aquifer tests and single-well tests can be analyzed by using the time–drawdown and time–residual–drawdown relationships of the well-flow equations of Section 10.2. Figure 10.10a gives an example of the time–drawdown relationship for the pumped well or a piezometer during a pumping test followed by a recovery test. With aquifer tests, the above relationships can be used to make time–drawdown analyses (Section 10.4) for each piezometer separately. In addition, the distance–drawdown relationship (see Figure 10.10b) can be used to make a distance–drawdown analysis (Section 10.5), provided that drawdown data in at least two piezometers are available. The advantage of aquifer tests is thus that the various estimates of the aquifer’s physical properties obtained from the time–drawdown analyses of each piezometer can be compared with an independent estimate of the physical properties obtained from a distance–drawdown analysis.

With single-well tests, only a single time–drawdown analysis based on the drawdowns of the pumped well can be made, that is, no comparison of results from different time–drawdown analyses is possible, and no distance–drawdown analysis can be made. Consequently, the results of aquifer tests will be more accurate than the results of single-well tests. Moreover, aquifer-test results are representative of a larger volume of the aquifer than single-well test results.

The question of how many piezometers should be employed depends not only on the amount of information desired and the required degree of accuracy, but also on the funds available for the test. It is always best to have as many piezometers as conditions permit and to place them in various directions and distances from the pumped well.

Although no fixed rule can be given, placing piezometers at distances of between 10 and 100 m from the well will usually give reliable data. For thick or stratified confined aquifers, these distances must be greater, say between 100 and 250 m or more. An optimum result in determining the physical properties of the aquifer can be obtained by placing the screen of the piezometer at half the depth of the screen of the pumped well.

The design and construction of the pumped well and the piezometers will be discussed in Chapter 11; for the site selection, reference is made to Kruseman and de Ridder (1994).

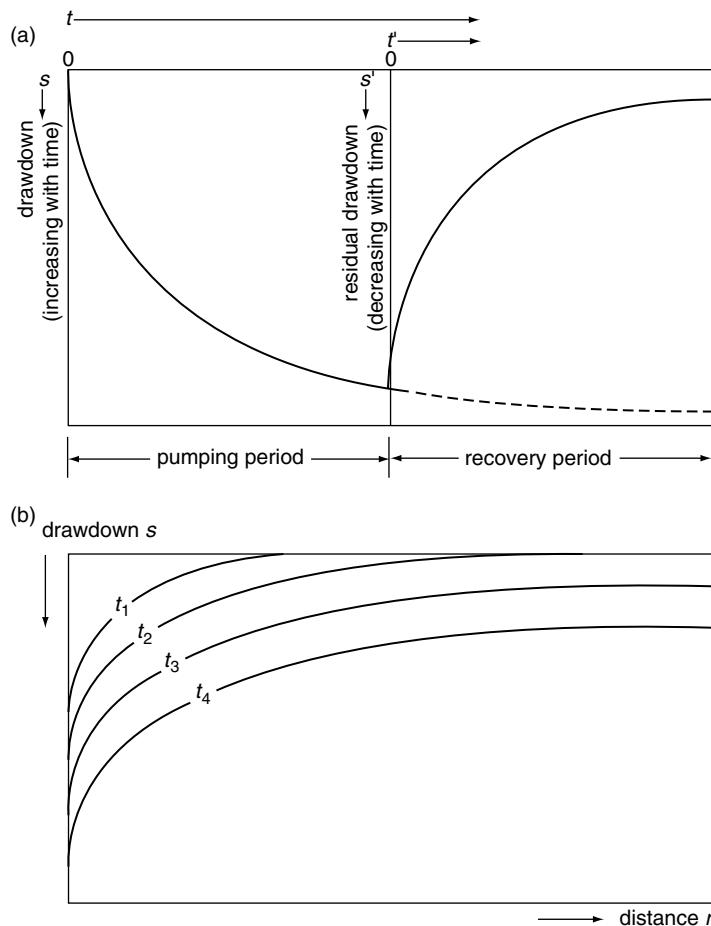


FIGURE 10.10 (a) Time–drawdown relationship during a pumping test, followed by a recovery test; (b) distance–drawdown relationship during a pumping test.

10.3.1 Measurements

Ideally, an aquifer test should be performed under the natural conditions of a stable water table. This is, however, not always possible. Water tables rise and fall due to natural recharge and discharge of the groundwater reservoir (precipitation and evaporation), manmade recharge and discharge of the groundwater reservoir (irrigation losses and pumping from wells), changes in barometric pressure, and tidal movements in coastal aquifers. Such short-term variations of the water table have an adverse effect on the drawdown and recovery of the water table during testing. Hence, for some days prior to the actual test, the water levels in the well and the piezometers should be measured, say twice a day. For each observation point, a time-vs.-water-level curve, or hydrograph, should be drawn. From these, the trend and rate of water-level changes can be read. At the end of the test, that is, after complete recovery, water-level readings should be continued at the observation points for one or two days. With these data, the hydrographs can be completed and the rate of water-level change during the test can be determined. This information can then be used to correct the drawdowns observed during the test itself.

The water-level measurements must be taken many times during the course of a test, and with as much accuracy as possible. As water levels drop fast during the first hour or two of the test, readings should first be taken at brief intervals. These intervals should be gradually increased as pumping continues. As in all the analysis procedures the time is plotted on a logarithmic scale, it is recommended to have the same

TABLE 10.1 Typical Time Sequence for Taking Water-Level Measurements in an Aquifer Test

Time (sec)	Time (min)	Time (min)	Time (h)	Time (h)
10	2.5	20	2.5	22
20	3.0	25	3.0	27
30	4.0	30	4.0	33
40	5.0	40	5.0	42
50	6.5	50	7.0	53
60	8.0	65	8.5	67
80	10.0	80	11.0	103
100	13.0	90	13.0	108
120	16.0	120	17.0	133

Note: Aim is to get equal number of observations in each log cycle of time.

number of readings in each log cycle of time. Table 10.1 shows an example of the sequence in time for taking water-level measurements, based on ten readings in each log cycle and resulting in approximately equidistant plotting positions. For piezometers far from the well and for those in aquitards above or below the aquifer, the brief time intervals in the first minutes of the pumping test can be disregarded.

After the pump has been shut down, the water levels in the well and the piezometers will start to rise. In the first hour, they will rise rapidly, but as time goes on, the rate of rise decreases. These rises can be measured in a recovery test. If the yield of the well was not constant throughout the pumping test, recovery-test data are more consistent than the drawdown data collected during pumping. Recovery-test data can thus be used as a check on the calculations that are based on the drawdown data. The schedule for recovery measurements is the same as that for the pumping period.

Water-level measurements can be taken in various ways, that is, wetted-tape method, mechanical sounder, electric water-level indicator, floating-level indicator or recorder, pressure gauge, or pressure logger. For detailed information on these devices, reference is made to Kruseman and de Ridder (1994) and Driscoll (1986). Fairly accurate measurements of water levels can be made manually, but then the instant of each reading should be recorded with a chronometer. Experience has shown that it is possible to measure the depth to water within 2 mm. For piezometers close to the well, the wetted-tape method and the mechanical sounder cannot be used due to the rapid water-level changes and the noise of the pump, respectively.

Although the pressure-gauge method is less accurate than the other methods (within 6 cm), it is the most practical method for measuring water levels in a pumped well. It should, however, not be used for measuring water levels in piezometers, as the drop in water levels during the initial stage of pumping will be small in these wells. Electronic self-logging pressure transducers with high sampling interval (20 sec) and high accuracy (1 to 6 mm) can be used when funds are available.

Among the arrangements to be made for a pumping test is the control of the discharge rate. To avoid complicated calculations later, the discharge rate should preferably be kept constant throughout the test. The discharge should be kept constant by manipulating a valve in the discharge pipe. This gives more accurate control than changing the speed of the pump. During pumping tests, the discharge rate should be measured at least once every hour, and adjustments necessary to keep it constant should be made. The discharge rate can be measured with various devices, that is, commercial water meter, flume, container, orifice weir, orifice bucket, or jet-stream method. For detailed information on these devices, reference is also made to Kruseman and de Ridder (1994) and Bos (1989). The water delivered by the well should be prevented from returning to the aquifer. This can be done by conveying the water through a large-diameter pipe over a convenient distance, at least 300 m, depending on the location of the piezometers, and then discharging it into a canal or natural channel. Preferably, the water should be discharged

away from the line of piezometers. The pumped water can also be conveyed through a shallow ditch, but precautionary measures should be taken to seal the bottom of the ditch with clay or plastic sheets to prevent leakage.

10.3.2 Duration of a Pumping Test

The question of how long a pumping test should last is difficult to answer as the period of pumping depends on the type of aquifer and the degree of accuracy desired in establishing its physical properties. Economizing on the pumping period is not recommended as the costs of running the pump a few extra hours are low compared with the total costs of the test. Moreover, better and more reliable data are obtained if pumping continues until the cone of depression has stabilized and does not seem to be expanding further as pumping continues. At the beginning of the test, the cone develops quickly as the pumped water is initially derived from the aquifer storage immediately around the well. But, as pumping continues, the cone expands and deepens more slowly as, with each additional meter of horizontal expansion, a larger volume of stored water becomes available. This may often lead inexperienced observers to conclude that the cone has stabilized, or in other words that steady state has been reached. Inaccurate measurements of the drawdowns in the piezometers — drawdowns that are becoming smaller and smaller as pumping continue — can also lead to this wrong conclusion. In reality, the depression cone will continue to expand until the recharge of the aquifer, if any, equals the discharge.

The unsteady-state flow, also known as non-equilibrium flow, is time dependent, that is, the water level as observed in piezometers, changes with time. During a pumping test, the unsteady-state flow condition occurs from the moment pumping starts till the steady state is reached. Theoretically, in an infinite, horizontal, completely confined aquifer of constant thickness that is pumped at a constant rate, there will always be an unsteady state, as the aquifer is not recharged by an outside source. In practice, well flow is considered to be in unsteady state as long as the changes of the water level in the piezometers are measurable, or as long as the hydraulic gradient changes in a measurable way.

The steady-state flow, also known as equilibrium flow, is independent of time, that is, the water level, as observed in piezometers, does not change with time. It occurs, for instance, when there is equilibrium between the discharge of a pumped well and the recharge of the pumped aquifer by an outside source. Such outside sources may be recharge from surface water of nearby rivers, canals, or lakes, or recharge from groundwater of an unconfined aquifer with constant water table overlying an aquitard that covers a pumped leaky aquifer. As real steady-state conditions seldom occur, it is said in practice that a steady-state condition is reached when the changes of the water level as observed in piezometers are negligibly small, or when the hydraulic gradient has become constant.

To establish whether unsteady or steady-state conditions prevail, the changes in head during the pumping test should be plotted. Figure 10.11 shows possible shapes of time–drawdown plots and their interpretations.

In some wells, a steady state occurs a few hours after the start of pumping; in others, it does not occur until after a few days or weeks, whereas in yet other wells it never occurs, even though pumping continues for years. Kruseman and de Ridder (1994) suggest that, under average conditions, steady-state flow is generally reached in leaky aquifers after 15 to 20 h of pumping, and in a confined aquifer, after 24 h. In an unconfined aquifer, the cone of depression expands more slowly, so a longer period of pumping is required, say, three days.

Preliminary plotting of drawdown data during the test will often show what is happening and may indicate whether or not a test should be continued.

10.3.3 Processing the Data

All measurements of the water level, time, and discharge of the pump should preferably be noted on pre-printed forms or entered in a portable or handheld computer. After 3 h of pumping, sufficient time will become available in the field to draw the time–drawdown curves of each observation point. These

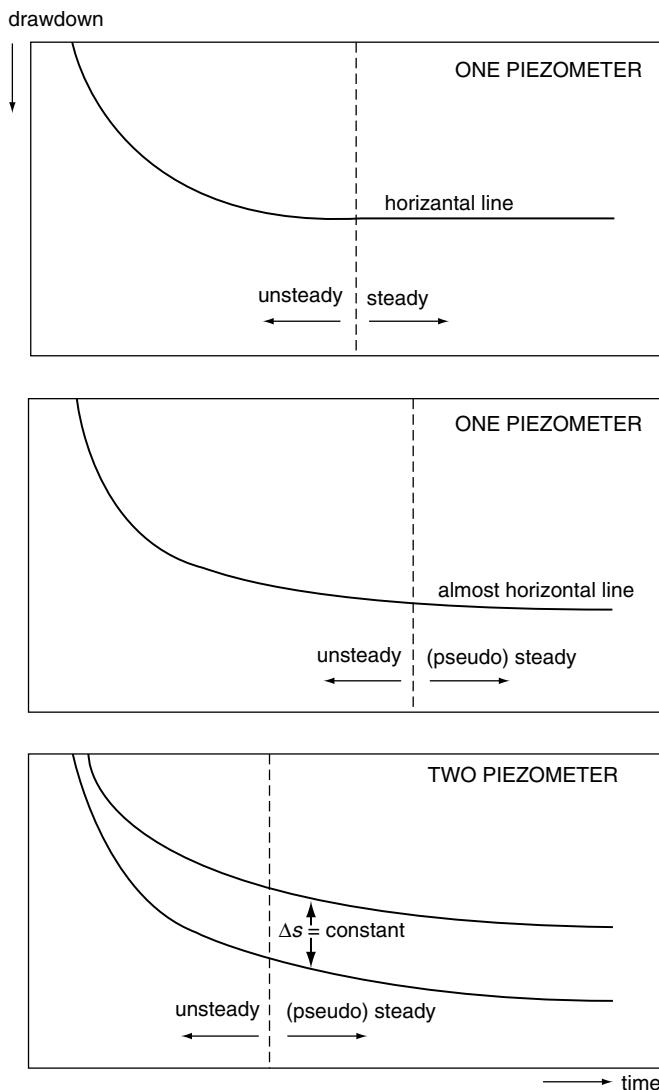


FIGURE 10.11 Time–drawdown plots showing the changes in drawdown during an aquifer test and their interpretations.

graphs will be helpful in checking whether the test is running well and in deciding on the time that the pump can be shut down when steady or pseudo-steady state flow has been reached.

As the well-flow equations require drawdown data instead of depth-to-water table data, the latter should therefore be converted to drawdown data, in other words, the initial depth to the water table must be subtracted from the depth-to-water table data during the test. It may be necessary to correct the observed drawdowns for external influences, that is, influences independent of the test (trend correction). If, during the post-recovery period, the same constant water level is observed as during the pre-testing period, it can be assumed that no external forces influenced the hydraulic head in the aquifer during the test. If, however, the water level is subject to unidirectional or rhythmic changes due to tidal influences for example, the observed drawdowns will have to be corrected before being used in the analysis. Such phenomena are likely to occur during pumping tests of long duration.

When the evaluation of the test data has been completed, a report on the results should be written. A copy of the report should be kept on file for further reference and later studies. Samples of the different

layers penetrated by the borings should be stored too, as they may be needed for other studies in a later phase of investigations. The basic field measurements of the test should be put on file as well. The conclusions drawn from the test may become obsolete in the light of new insights, but the hard facts, carefully collected in the field, remain facts and can always be reevaluated.

10.3.4 Interpretation of the Data

Calculating physical properties would be relatively easy if the aquifer system (i.e., aquifer plus well) were precisely known. This is generally not the case, so interpreting aquifer tests is primarily a matter of identifying an unknown system. System identification relies on models, the characteristics of which are assumed to represent the characteristics of the real aquifer system. In Section 10.2, well-flow equations were presented based on idealized theoretical models: homogeneous and isotropic aquifers of seemingly infinite extent, no well bore storage inside the pumped well, and the like.

To identify an aquifer system, one must compare its drawdown behavior with that of the various theoretical models. The model that compares best with the real system is then selected for the calculation of the physical properties. System identification includes the construction of diagnostics plots. In the literature, both log-log plots and semi-log plots of the drawdown vs. the time are used for this purpose. In the following sections, only the semi-log plots are used for identification and analysis purposes, as (1) they have more diagnostic value than log-log plots due to the occurrence of straight-line segments, and (2) they can also be used to analyze the field data of single-well tests. The characteristic shapes of the curves can help in selecting the appropriate model.

The choice of theoretical model is a crucial step in the interpretation of aquifer tests. If the wrong model is chosen, the physical properties for the real aquifer will not be correct. A troublesome fact is that theoretical solutions to well-flow problems are usually not unique. Some models, developed for different aquifer systems, yield similar responses to a given stress exerted on them. This makes system identification and model selection a difficult affair.

10.4 Time-Drawdown Analyses

All the analysis methods have in common that they are based on graphical solution procedures. For the pumping period, time-drawdown plots when plotted on semi-log paper will show characteristic shapes according to the governing equations as was illustrated in Figures 10.2, 10.4, 10.6, and 10.8. For the recovery period, time-ratio-residual-drawdown plots when plotted on semi-log paper will give characteristic shapes as was illustrated in Figure 10.9.

For the three main types of aquifers, the analysis procedures are now discussed together with sample calculations. These analysis procedures can be applied both to aquifer-test data as well as single-well test data. With the latter, the value of the specific yield or storativity can, in principle, not be determined.

10.4.1 Aquifer Test in a Confined Aquifer

The physical properties of a confined aquifer can be found by developing the time-drawdown relationship based on Equation 10.3. If the pumping time is long enough, a plot of the drawdown s observed at a particular distance r from the pumped well vs. the logarithm of time t , will appear as a straight line. If the slope of the straight-line segment is expressed as the drawdown difference ($\Delta s = s_1 - s_2$) per log cycle of time ($\log t_2/t_1 = 1$), rearranging Equation 10.3 gives

$$T = \frac{2.3Q}{4\pi \Delta s} \quad (10.17)$$

TABLE 10.2 Time–Drawdown Values Observed in a Piezometer at 90 m Distance from the Pumped Well of an Aquifer Test

Time (min)	Drawdown (cm)	Time (min)	Drawdown (cm)	Time (min)	Drawdown (cm)
1	2.5	30	20.0	150	35.0
2	3.9	40	22.6	360	42.6
4	6.1	50	24.7	550	44.0
6	8.0	60	26.4	720	44.5
9	10.6	90	30.4		
20	16.8	120	33.0		

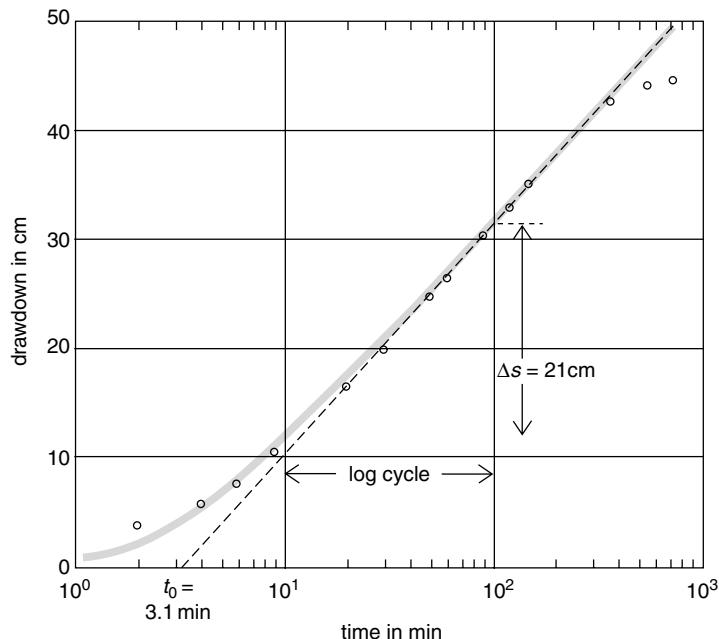


FIGURE 10.12 Time–drawdown plot of field data of an aquifer test in a confined aquifer.

If this line is extended until it intercepts the time-axis where $s = 0$, the interception point has the coordinates $s = 0$ and $t = t_0$. Substituting these values into Equation 10.3, after rearrangement, gives

$$S = \frac{2.25 T t_0}{r^2} \quad (10.18)$$

This procedure will now be illustrated with field data. An aquifer test was made in a confined aquifer. The well was pumped at a constant rate of $528 \text{ m}^3/\text{d}$ during 12 h. Table 10.2 shows the drawdowns of a piezometer at 90 m distance from the pumped well at various times. Figure 10.12 shows the time–drawdown plot on semi-log paper. From visual inspection, the data plot exhibits a straight line in the time range from 20 to 360 min. Through these points a best fitting straight line was drawn; Figure 10.12 shows that its slope Δs is $21 \text{ cm} = 0.21 \text{ m}$ per log cycle of time. Substitution of the appropriate values into Equation 10.17 gives

$$T = \frac{2.3 \times 528}{4 \times 3.14 \times 0.21} = 460 \text{ m}^2/\text{d}$$

From Figure 10.12 it can be seen that the extrapolated straight line intersects the x -axis where $s = 0$ at $t_0 = 3.1 \text{ min} = 2.2 \times 10^{-3} \text{ d}$. Substitution of the appropriate values into Equation 10.18 yields

$$S = \frac{2.25 \times 460 \times 2.2 \times 10^{-3}}{90^2} = 2.8 \times 10^{-4}$$

Finally, the selected data points to draw the best fitting straight line in Figure 10.12 should have corresponding $1/u$ values larger than 10 which is a practical condition for the applicability of the Jacob method. Rearranging Equation 10.2 and substituting $1/u > 10$, this condition gives

$$t > \frac{10r^2 S}{4T} = \frac{10 \times 90^2 \times 2.8 \times 10^{-4}}{4 \times 460} = 1.2 \times 10^{-2} \text{ d} = 18 \text{ min}$$

So, theoretically, all the observed drawdown values with t values larger than 18 min have $1/u$ values larger than 10 and can thus be expected to lie on a straight line. As the straight line was initially drawn through the plotted points in the time range from 20 to 360 min the above condition was already fulfilled. For an explanation for the fact that the last two drawdown values do not fall on the straight line, the reader is referred to Section 10.7.

The well was not observed during the recovery period, so no second estimate for the physical properties could be made based on the time–residual-drawdown relationship. Also no other piezometers were drilled, so the analysis of this aquifer test yielded only one value of the transmissivity and one value of the storativity and thus the consistency of these parameter values could not be checked.

To check the consistency of the analysis results in another way, the following procedure was followed. For all the data points, their u and corresponding $W(u)$ values were calculated and substituted into Equation 10.1 together with the appropriate values for Q and T . These theoretical drawdown values based on the Theis equation were plotted as a shaded line in Figure 10.12. It clearly shows that there is a good match between these theoretical drawdown values and those observed in the field. The above physical properties can thus be regarded to be representative for the aquifer under consideration. With the software package SATEM (Boonstra and Kselik, 2001), these theoretical drawdown values are always plotted in the data plot to check the consistency of the analysis results.

Finally, it can be noted that the more strict criterion, that is, that the data points should have $1/u$ values larger than 100, would have resulted in no analysis at all; in that case, only the last three data points in Figure 10.12 ($t > 180 \text{ min}$) would have been eligible to draw a best fitting straight line. This test clearly demonstrates that the limiting condition of large $1/u$ values to be able to apply the Jacob method, should be treated judiciously.

10.4.2 Aquifer Test in an Unconfined Aquifer

Figure 10.4 shows that even with a delayed yield effect, the time–drawdown curve will eventually exhibit a straight-line segment under a slope for late-time conditions. So, if the pumping time is sufficiently long, the physical properties of an unconfined aquifer can thus also be found by using Equation 10.17 and Equation 10.18, provided that the drawdown are corrected according to Equation 10.5. This procedure will now be illustrated with field data. An aquifer test was made in a 470-m thick, unconfined aquifer. The well only penetrating the aquifer to a depth of 60 m with a screen length of 40 m, was pumped at a constant rate of $6350 \text{ m}^3/\text{d}$ during 100 h. Table 10.3 shows the drawdowns observed in a piezometer at a distance of 91.5 m from the pumped well as function of time. The depth of this piezometer was 46 m with the last two meters screened.

The first step is to determine whether the observed drawdown values are small compared to the aquifer's thickness. The last observed drawdown value in Table 10.3 is less than 5% of the aquifer thickness. So, the uncorrected drawdown values can thus be used in the analysis. Figure 10.13 shows the time–drawdown plot on semi-log paper. From visual inspection, the data plot exhibits a straight line in the time range from

TABLE 10.3 Time–Drawdown Values Observed in a Piezometer at 91.5 m Distance from the Pumped Well of an Aquifer Test

Time (min)	Drawdown (cm)	Time (min)	Drawdown (cm)	Time (min)	Drawdown (cm)
1	1.0	40	18.0	600	25.9
2	1.0	50	18.6	750	26.5
3	2.1	60	19.2	1000	27.4
4	3.4	75	20.1	1250	28.7
6	6.1	100	20.7	1500	29.3
8	8.8	125	21.3	1750	29.9
10	11.0	150	21.6	2000	30.5
12	12.2	175	21.9	2500	32.6
15	13.4	200	22.3	3000	33.5
18	14.3	250	22.9	4000	35.7
21	15.2	300	23.8	5000	36.9
25	15.8	400	24.4	6000	36.9
30	16.8	500	25.3		

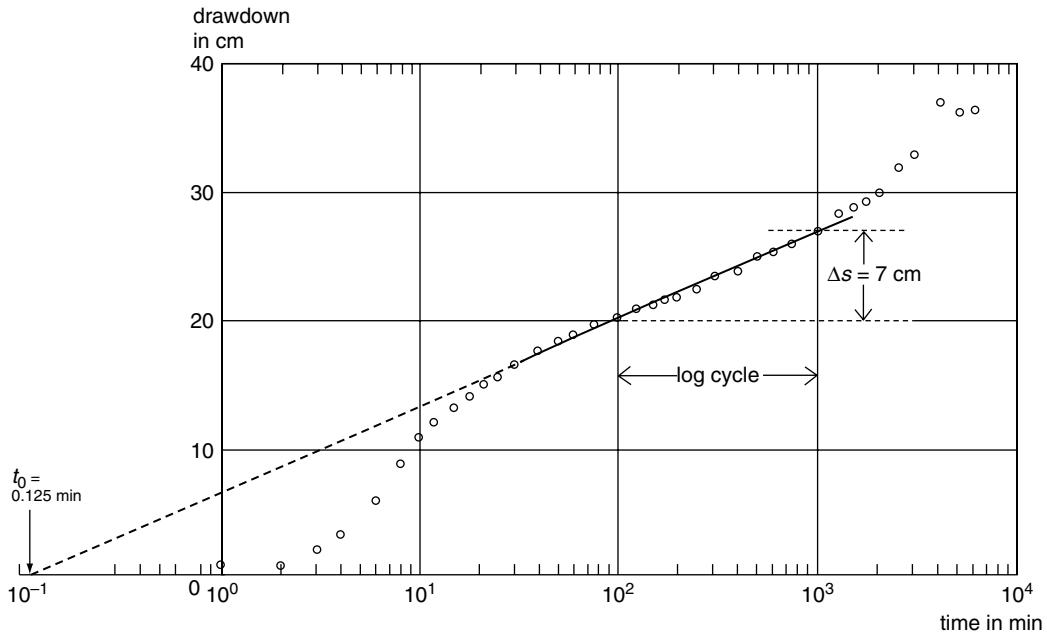


FIGURE 10.13 Time–drawdown plot of field data of an aquifer test in an unconfined aquifer with partially penetrating pumped well.

30 to 1000 min. Through these points a best fitting straight line was drawn; Figure 10.13 shows that its slope Δs is $7 \text{ cm} = 0.07 \text{ m}$ per log cycle of time. Substitution of the appropriate values into Equation 10.17 gives

$$T = \frac{2.3 \times 6350}{4 \times 3.14 \times 0.07} = 16,603 \text{ m}^2/\text{d}$$

Equation 10.18 cannot be used to estimate the specific yield due to the fact that there is a partially penetrating pumped well. Substituting the values $s = 0$ and $t = t_0$ into Equation 10.12, the specific yield

TABLE 10.4 Values of Components in Summation Series of Equation 10.11

n	K_0	sin	sin	cos	Component	Summation
1	0.7629	0.3904	0.1333	0.9551	0.1874	0.1874
2	0.3095	0.7188	0.2642	0.8244	0.0580	0.2454
3	0.1401	0.9332	0.3904	0.6197	0.0157	0.2611
4	0.0665	0.9994	0.5096	0.3594	0.0029	0.2640
5	0.0326	0.9071	0.6197	0.0668	0.0001	0.2641

value can then be obtained by

$$S_y = \frac{2.25 T t_o}{r^2} e^{f_s} \quad (10.19)$$

From Figure 10.13 it can be seen that the extrapolated straight line intersects the x -axis where $s = 0$ at $t_o = 0.125 \text{ min} = 8.7 \times 10^{-5} \text{ d}$. Before the value of S_y can be calculated, the value of f_s needs to be determined. Table 10.4 shows the different components of the first five values in the summation series of Equation 10.11 with $b = 60 \text{ m}$, $d = 20 \text{ m}$, $b' = 46 \text{ m}$, $d' = 44 \text{ m}$, and $D = 470 \text{ m}$. Substituting the appropriate values into Equation 10.11 gives

$$f_s = \frac{4 \times 470}{3.14(60 - 20)} \times 0.2641 = 14.96 \times 0.2641 = 3.95$$

Substitution of the appropriate values into Equation 10.19 then yields

$$S_y = \frac{2.25 \times 16,603 \times 8.7 \times 10^{-5}}{91.5^2} e^{3.95} = 3.9 \times 10^{-4} \times 51.9 = 0.02$$

Finally, the limiting condition to use Equation 10.12 for evaluating the physical properties of the tested unconfined aquifer should be checked:

$$t > \frac{D^2 S_y}{2T} = \frac{470^2 \times 0.02}{2 \times 16,603} = 0.133 \text{ d} = 192 \text{ min}$$

So, theoretically, all the observed drawdown values with t values larger than 192 min, can be expected to lie on a straight line. As the straight line was initially drawn through the plotted points in the time range from 30 to 1,000 min, the above condition was thus violated. When another straight line is drawn through the points in the time range from 200 to 1,000 min, its slope will be somewhat steeper. A repetition of the above calculations then yields a transmissivity value of $15,828 \text{ m}^2/\text{d}$ and a specific yield value of 0.016. These values do not differ significantly from the first ones.

For this aquifer test, two other piezometers were drilled at distances of 15.2 and 30.5 m from the pumped well (Boonstra, 1992). The time–drawdown analyses from these piezometers gave two other estimates of the physical properties: transmissivities of 16,691 and $14,104 \text{ m}^2/\text{d}$ and specific yield values of 0.011 and 0.013, respectively. This illustrates well the fact that with aquifer tests with more than one piezometer, a comparison between the various estimates of the physical properties can be made. When the various estimates are consistent in their values as with this test, more reliability can be given to the analyses results. It should be noted that the range in transmissivity values is usually smaller than the range in specific yield values due to the nature of the analysis procedure. This phenomenon also pertains to the aquifer-test results in confined and leaky aquifers.

10.4.3 Aquifer Test in a Leaky Aquifer

The physical properties of a leaky aquifer can be found by developing the time–drawdown relationship based on Equation 10.6. Figure 10.6 shows that the Hantush well function exhibits an inflection point and for large values of $1/u$ a horizontal straight-line segment indicating steady-state flow. On these phenomena the inflection-point method was based. Hantush (1956) showed that for the inflection point (s_p , t_p) the following relationships hold:

1. The drawdown value s_p is given by

$$s_p = 0.5s_m \quad (10.20)$$

2. The u_p value is given by

$$u_p = \frac{r^2 S}{4Tt_p} = \frac{r}{2L} \quad (10.21)$$

3. The slope of the curve at the inflection point Δs_p per log cycle of time is given by

$$\Delta s_p = \frac{2.3Q}{4\pi T} e^{-r/L} \quad (10.22)$$

4. At the inflection point, the relation between the drawdown and the slope of the curve is given by

$$2.3 \frac{s_p}{\Delta s_p} = e^{r/L} K_o(r/L) \quad (10.23)$$

This procedure will now be illustrated with field data. An aquifer test was made in a leaky aquifer. The well was pumped at a constant rate of $545 \text{ m}^3/\text{d}$ during 48 h. One of the piezometers was located 20 m away from the pumped well. Table 10.5 shows the observed drawdowns as function of time. Figure 10.14 shows that steady-state conditions did occur at the end of the pumping test

TABLE 10.5 Time–Drawdown Values Observed in a Piezometer at 20 m Distance from the Pumped Well of an Aquifer Test

Time (min)	Drawdown (m)	Time (min)	Drawdown (m)	Time (min)	Drawdown (m)
1	0.265	60	1.615	720	2.320
2	0.347	70	1.675	840	2.337
3	0.490	80	1.725	960	2.355
4	0.584	90	1.767	1080	2.377
5	0.635	120	1.895	1200	2.373
6	0.707	150	1.977	1320	2.385
7	0.793	180	2.020	1440	2.400
8	0.839	210	2.075	1620	2.410
9	0.882	240	2.113	1800	2.420
10	0.930	300	2.170	1980	2.440
15	1.079	360	2.210	2160	2.440
20	1.187	420	2.245	2340	2.450
25	1.275	480	2.270	2520	2.450
30	1.345	540	2.281	2700	2.440
40	1.457	600	2.298	2880	2.440
50	1.554	660	2.310		

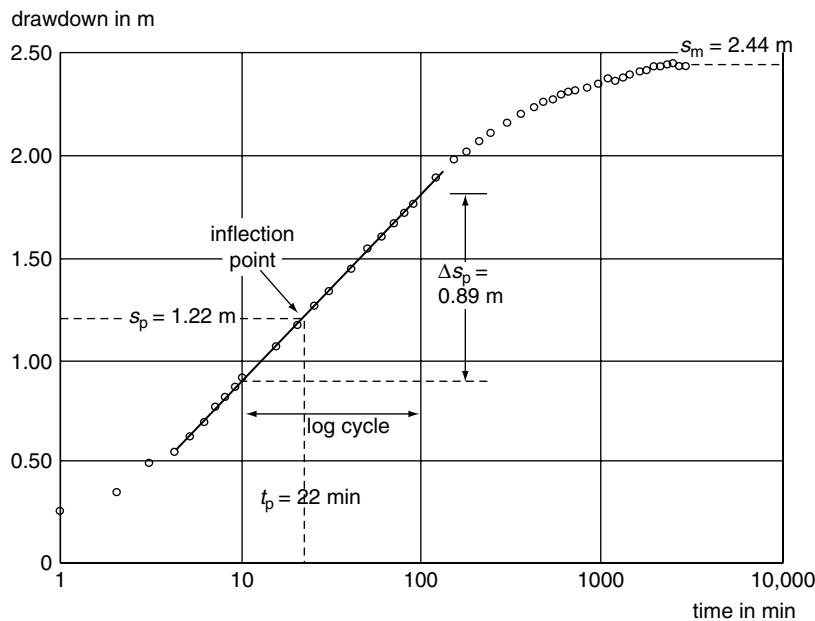


FIGURE 10.14 Time–drawdown plot of field data of an aquifer test in a leaky aquifer.

($s_m = 2.44 \text{ m}$). According to Equation 10.20 the s_p value is then 1.22 m . Locating the inflection point in Figure 10.14 using this s_p value, gives a t_p value of $22 \text{ min} = 1.5 \times 10^{-2} \text{ d}$.

The slope of the curve at this inflection point is 0.89 m per log cycle of time. Substitution of the appropriate values into Equation 10.23 yields

$$e^{r/L} K_o(r/L) = 2.30 \frac{s_p}{\Delta s_p} = 2.3 \frac{1.22}{0.89} = 3.15$$

A table with r/L , $K_o(r/L)$ and $e^{r/L} K_o(r/L)$ values is provided in Appendix 10.1. The r/L value is then 0.0575 . Substitution of the r value yields $L = 20/0.0575 = 348 \text{ m}$. Rearranging Equation 10.22 and substituting the appropriate values yields

$$T = \frac{2.3Q}{4\pi \Delta s_p} e^{-r/L} = \frac{2.3 \times 545}{4 \times 3.14 \times 0.89} e^{-0.0575} = 106 \text{ m}^2/\text{d}$$

The c value is then $L^2/T = 348^2/106 = 1142 \text{ d}$. Rearranging Equation 10.21 and substituting the appropriate values yields

$$S = \frac{4Tt_p}{r^2} \frac{r}{2L} = \frac{4 \times 106 \times 1.5 \times 10^{-2}}{20^2} \times \frac{20}{2 \times 348} = 4.6 \times 10^{-4}$$

It should be noted that the accuracy of the calculated physical properties depends to a large extent on the accuracy of the value of the (extrapolated) steady-state drawdown s_m . The calculations should therefore be checked by substituting the different values into Equation 10.6; calculations of s should be made for the observed values of t . If the values of t are not too small, the values of s should fall on the observed data curve. If the calculated data deviate from the observed data, the (extrapolated) value of s_m should be adjusted. Sometimes, the observed data curve can be drawn somewhat steeper or flatter through the plotted points, and so Δs_p can be adjusted too. With the new values of s and Δs_p , the calculation is then repeated.

TABLE 10.6 Time–Drawdown Values of a Single-Well Test

Time (min)	Drawdown (m)	Time (min)	Drawdown (m)	Time (min)	Drawdown (m)
15	4.161	50	4.486	240	4.808
20	4.283	56	4.471	300	4.776
25	4.257	60	4.474	360	4.885
30	4.357	80	4.534	420	4.960
36	4.358	105	4.618	480	4.906
40	4.399	120	4.672	540	4.972
46	4.456	180	4.748	600	5.016

With the software package SATEM (Boonstra and Kselik, 2001), the above time-consuming work can be avoided. This program follows the same procedure as described above. In addition, it displays the drawdowns calculated with Equation 10.6 on the monitor, together with the data observed in the field. This makes it easy to check whether the correct (extrapolated) steady-state drawdown s_m has been selected.

10.4.4 Single-Well Tests

With single-well tests, basically the same procedures can be applied as with aquifer tests. The r value now represents the effective radius of the single well. This is difficult to determine under field conditions; as a “best” estimate, the outer radius of the well screen is often used.

A complicating factor is the phenomenon that due to nonlinear well losses the water levels inside the well can be considerably lower than those directly outside the well screen. This implies that drawdown data from the pumped well can, in general, only be used for the analysis when corrected for these nonlinear well losses using the results of so-called step-drawdown tests (see Chapter 11).

The procedure with single-well tests will now be illustrated with field data from an unconfined aquifer. The well was pumped at a constant rate of $3853 \text{ m}^3/\text{d}$ for 10 h. The outer radius of the well screen was 0.20 m. Table 10.6 shows the observed drawdowns as a function of time. At the same site, a step-drawdown test was also made. From the analysis of this test, the nonlinear well loss was calculated as 1.93 m (see Chapter 11 under the heading “Jacob’s Method”).

The first step is to determine whether the observed drawdown values are small compared with the aquifer’s thickness. The depth of the pumped borehole was 271 m. Substituting this value and the last observed drawdown value corrected for nonlinear well loss into Equation 10.5 gives a maximum correction value of $(5.016 - 1.93)^2 / (2 \times 271) = 0.02 \text{ m}$. So, no corrections on the observed drawdowns are needed with respect to Equation 10.5.

Figure 10.15 shows the time–drawdown plot of the observed drawdowns in the pumped well corrected for nonlinear well loss on semi-log paper. From visual inspection, the data plot exhibits a straight line over the whole time range. Through all the points a best fitting straight line was drawn; Figure 10.15 shows that its slope Δs is 0.50 m per log cycle of time. Substitution of the appropriate values into Equation 10.17 gives

$$T = \frac{2.3 \times 3853}{4 \times 3.14 \times 0.50} = 1410 \text{ m}^2/\text{d}$$

From Figure 10.15 it can be seen that the intersection of the straight line with the x -axis where $s = 0$ cannot be determined directly. When such a situation occurs, the following procedure can be used. The Δs value was calculated as 0.50 m. Figure 10.15 shows that for a drawdown of 2.50 m — being 5 times the Δs value — the corresponding time is 44 min. The t_0 value is then equal to $44 \times 10^{-5} \text{ min} = 3.1 \times 10^{-7} \text{ d}$. Substitution of the appropriate values into Equation 10.18 yields

$$S_y = \frac{2.25 \times 1410 \times 3.1 \times 10^{-7}}{0.20^2} = 0.025$$

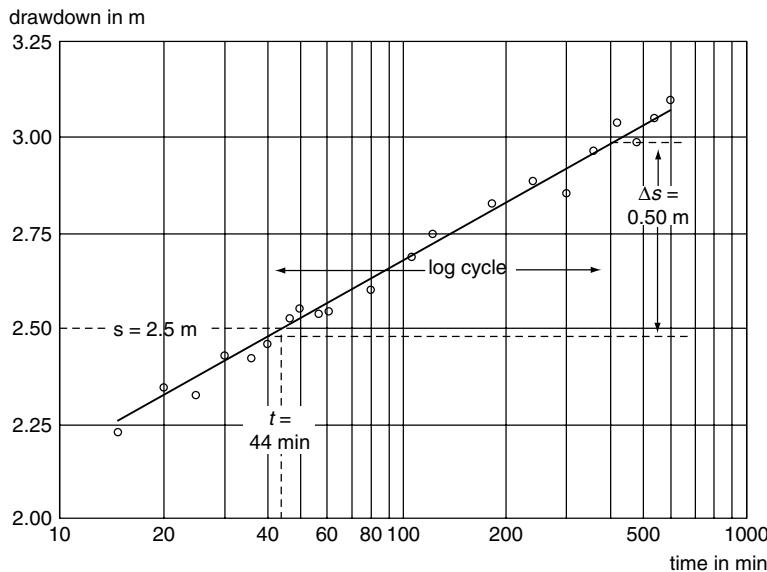


FIGURE 10.15 Time–drawdown plot of field data of a single-well test in an unconfined aquifer.

Finally, the selected data points to draw the best fitting straight line in Figure 10.15 should have corresponding $1/u$ values larger than 10 which is a practical condition for the applicability of the Jacob method. Rearranging Equation 10.2 and substituting $1/u > 10$, this condition gives

$$t > \frac{10r^2S_y}{4T} = \frac{10 \times 0.2^2 \times 0.025}{4 \times 1410} = 1.8 \times 10^{-6} d = 2.6 \times 10^{-3} \text{ min}$$

So, theoretically, all the observed drawdown values with t values larger than 2.6×10^{-3} min, can be expected to lie on a straight line. In other words, all the observed drawdown data as plotted in Figure 10.15 can be used to determine the slope of the straight line. It should be noted that the condition $1/u > 10$ is usually fulfilled in single-well tests because of the small r value.

With the above procedure, we can also use the uncorrected drawdown data and still get representative transmissivity values, as the slope of the straight-line segment in the time–drawdown plot on semi-log paper is not affected by the constant value of the nonlinear well loss. Specific yield values based on the corrected data, should, however, be treated with caution, as they are highly sensitive to the value of the effective radius of the pumped well.

For this single-well test, field measurements were also made during the recovery period. Table 10.7 shows the observed residual drawdowns as a function of time. Figure 10.16 shows the plot of residual drawdown vs. time on semi-log paper. From visual inspection, the data plot exhibits a straight line in the time-ratio range from 2 to 40. Through these points a best fitting straight line was drawn; Figure 10.16 shows that its slope $\Delta s'$ is 0.56 m per log cycle of time ratio. Rearranging Equation 10.15 and assuming that $S'_y = S_y$ gives

$$T = \frac{2.3Q}{4\pi \Delta s'} = \frac{2.3 \times 3853}{4 \times 3.14 \times 0.56} = 1260 \text{ m}^2/\text{d}$$

From Figure 10.16 it can be seen that the straight line intersects the axis where $s'(r, t/t') = 0$, at $(t/t')_0 = 1.39$. Substituting the appropriate values into Equation 10.15 yields

$$S'_y = \frac{S_y}{1.39} = 0.7S_y$$

TABLE 10.7 Time-Residual Drawdown Values of a Single-Well Test

Time (min)	Residual Drawdown (m)	Time (min)	Residual Drawdown (m)	Time min)	Residual Drawdown (m)
5	0.888	46	0.588	240	0.226
10	0.847	50	0.563	330	0.158
15	0.817	56	0.517	450	0.193
20	0.760	60	0.514	570	0.060
25	0.683	75	0.460	630	0.075
30	0.648	105	0.393	750	0.015
34	0.636	120	0.321	870	0.015
40	0.588	180	0.310	990	0.001

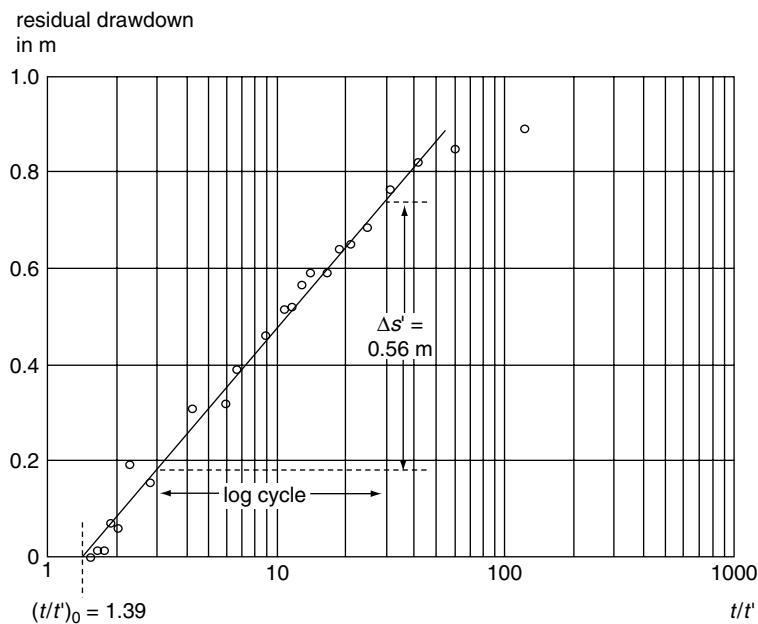


FIGURE 10.16 Time–ratio–residual-drawdown plot of field data of a single-well test in an unconfined aquifer.

The above implies that the specific yield value during recovery is less than during pumping. This phenomenon is often encountered because of air entrapment when the pores are again filled with water.

The above transmissivity value corresponds well with that found from the time–drawdown analysis. The application of time–drawdown and time–recovery analyses thus enables us to check the calculated transmissivity value. When the two values are close to each other, it implies that the data are consistent, that is, that the results of the test are reliable.

10.5 Distance–Drawdown Analyses

These analysis methods have also in common that they are based on graphical solution procedures. For all three main types of aquifers, plots of steady-state drawdowns of piezometers vs. the logarithms of the distances of these piezometers from the pumped well will show characteristic shapes; these diagnostic plots are from hereon referred to as distance–drawdown plots. Steady-state drawdowns of at least two piezometers are required for this type of analysis. The presented distance–drawdown analyses can only

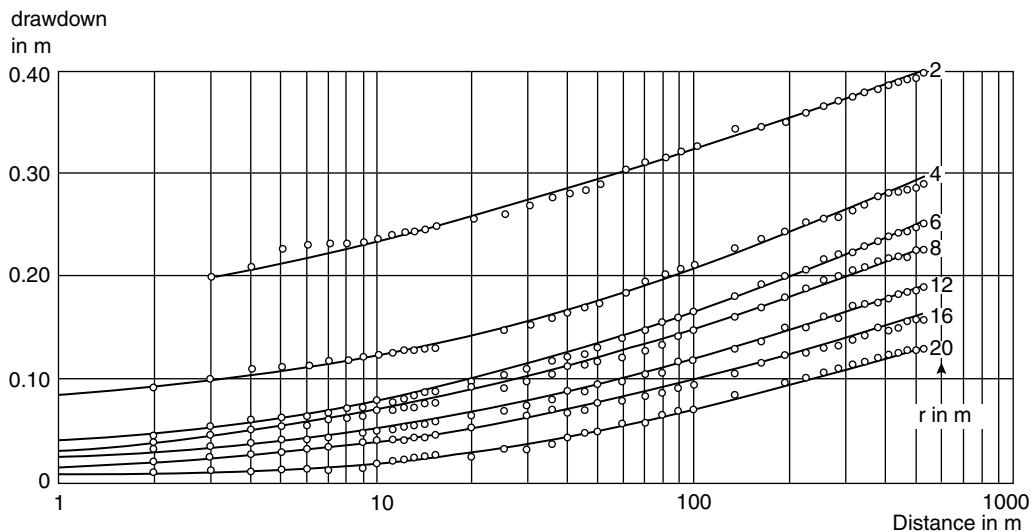


FIGURE 10.17 Time–drawdown curves of seven piezometers showing field data of an aquifer test in an unconfined aquifer.

be applied to aquifer tests with fully-penetrating pumped wells. The advantage of distance–drawdown analyses is that an independent estimate of the aquifer’s transmissivity can be made, in addition to the estimates obtained from the time–drawdown analyses of each piezometer separately. For leaky aquifers, an independent estimate of the aquitard’s hydraulic resistance can be made as well.

10.5.1 Confined and Unconfined Aquifers

Theoretically, in confined and unconfined aquifers that are pumped at a constant rate, there will always be an unsteady state, as these aquifers are not recharged by an outside source. Instead of steady state, the concept of pseudo-steady state is used for these types of aquifers. It is said that the aquifer is in pseudo-steady state when the hydraulic gradient has become constant, that is, when the time–drawdown curves are running parallel (see Figure 10.11).

The physical properties of a confined or unconfined aquifer can be found by developing the distance–drawdown relationship based on Equation 10.4. From this equation it can be seen that the corresponding distance–drawdown plot will show a straight line. If the slope of this straight line is expressed as the drawdown difference ($\Delta s = s_1 - s_2$) per log cycle of distance ($\log r_2/r_1 = 1$), rearranging Equation 10.4 gives

$$T = \frac{2.3Q}{2\pi \Delta s} \quad (10.24)$$

This procedure will now be illustrated with field data. An aquifer test was made in a shallow unconfined aquifer; prior to pumping, the aquifer’s saturated thickness was only 6.5 m. The well was pumped at a constant rate of $167 \text{ m}^3/\text{d}$ for 520 min. The water table was observed in seven piezometers. Figure 10.17 shows the time–drawdown graphs of these seven wells; the plotted drawdowns were already corrected using Equation 10.5. As can be seen from this figure, the curves run almost parallel in the last few hours of the test. Table 10.8 shows the drawdown values in the seven piezometers at the end of the test; these were assumed to represent the pseudo-steady-state drawdowns. Figure 10.18 shows the distance–drawdown plot on semi-log paper. From visual inspection, the data plot exhibits a straight line in the distance range from 4 to 20 m. Through these points a best fitting straight line was drawn; Figure 10.18 shows that its slope Δs is 0.22 m per log cycle of distance. Substitution of the appropriate values into Equation 10.24

TABLE 10.8 Pseudo-Steady-State Drawdown
Values of an Aquifer Test

Distance to Pumped Well (m)	Pseudo-Steady-State Drawdown	
	Uncorrected (m)	Corrected (m)
2	0.407	0.394
4	0.294	0.287
6	0.252	0.247
8	0.228	0.224
12	0.193	0.190
16	0.161	0.159
20	0.131	0.130

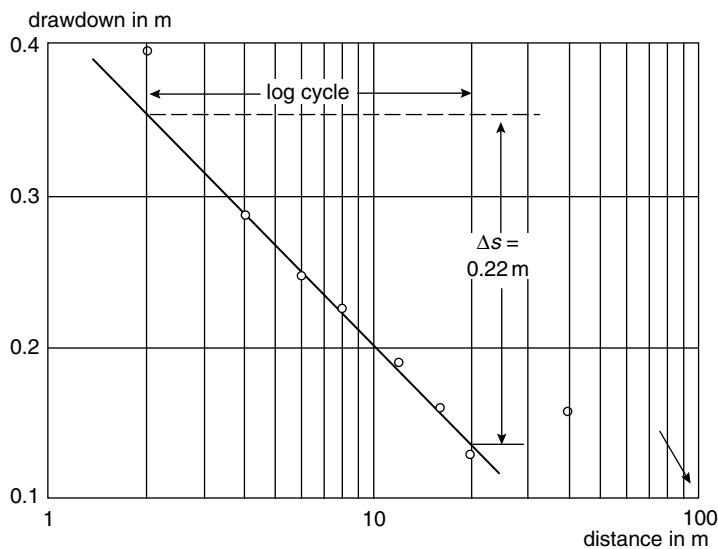


FIGURE 10.18 Distance-drawdown plot of field data of an aquifer test in an unconfined aquifer.

gives

$$T = \frac{2.3 \times 167}{2 \times 3.14 \times 0.22} = 278 \text{ m}^2/\text{d}$$

The fact that the drawdown of the piezometer at 2 m distance is not lying on the straight line, can be explained by an additional head loss that usually occurs near the well because of the relatively strong curvature of the water table. It should be noted that, as the aquifer is very thin, the piezometers were drilled at relatively short distances from the pumped well. The above transmissivity value corresponds well with those found from the time–drawdown analyses based on the individual piezometers; the latter ranged from 210 to 260 m^2/d .

10.5.2 Leaky Aquifers

The physical properties of a leaky aquifer can be found by developing the distance–drawdown relationship based on Equation 10.8. The transmissivity of a leaky aquifer can also be found from Equation 10.24 (Boonstra and de Ridder, 1994). If the straight line is extended until it intercepts the distance-axis where

TABLE 10.9 Extrapolated Steady-State Drawdown Values of an Aquifer Test

Distance from Pumped Well (m)	Steady-State Drawdown (m)
10	0.282
30	0.235
60	0.170
90	0.147
120	0.132
400	0.059

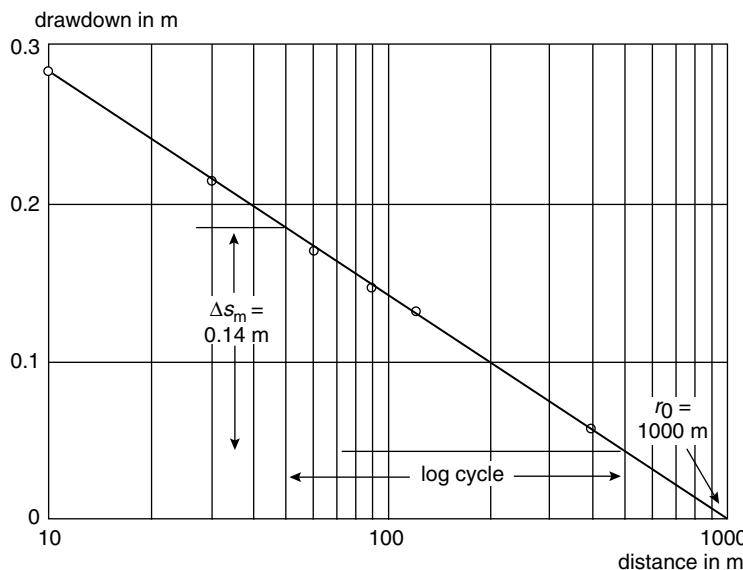


FIGURE 10.19 Distance-drawdown plot of field data of an aquifer test in a leaky aquifer.

$s_m = 0$, the interception point has the coordinates $s_m = 0$ and $r = r_0$. Substituting these values into Equation 10.8 gives

$$\log \frac{1.12L}{r_0} = 0 \quad \text{or} \quad \frac{1.12L}{r_0} = 1 \quad \text{or} \quad L = \frac{r_0}{1.12} \quad (10.25)$$

This procedure will now be illustrated with field data. An aquifer test was made in a leaky aquifer with a thickness of 10 m. The well was pumped at a constant rate of $761 \text{ m}^3/\text{d}$ for 480 min. Table 10.9 shows the extrapolated steady-state drawdowns in the six piezometers. Figure 10.19 shows the distance-drawdown plot on semi-log paper. From visual inspection, the data plot exhibits a straight line over the whole distance range. Through all the points a best fitting straight line was drawn; Figure 10.19 shows that its slope Δs_m is 0.14 m per log cycle of distance. Substitution of the appropriate values into Equation 10.24 gives

$$T = \frac{2.3 \times 761}{2 \times 3.14 \times 0.14} = 1990 \text{ m}^2/\text{d}$$

From Figure 10.19 it can be seen that the extrapolated straight line intersects the x -axis where $s_m = 0$ at $r_o = 1000$ m. Substitution of this value into Equation 10.25 gives

$$L = \frac{r_0}{1.12} = \frac{1000}{1.12} = 890 \text{ m}$$

The c value can now be calculated as $c = L^2/T = 890^2/1990 = 398$ d. Finally, the limiting conditions should be checked. Substituting the above L value into the condition $L > 3D$ gives $D < 296$ m, so this condition is fulfilled. Substituting the appropriate values into the condition $r/L < 0.2$ gives $r < 178$ m. According to this condition, the drawdown value of the piezometer at a distance of 400 m should be eliminated from the analysis. Figure 10.19, however, shows that this point, too, lies on the straight line, so in this case this condition is not a limiting factor. The above transmissivity and hydraulic resistance values correspond well with those found from the time–drawdown analyses based on the individual piezometers; the latter ranged from 1730 to 1980 m^2/d for the aquifer's transmissivity and from 350 to 510 d for the aquitard's hydraulic resistance.

10.6 Interpretation of Heterogeneous Aquifers

The assumptions shown in Section 10.2 are not likely to be found in a real field situation. Especially the assumption that the aquifer is an isotropic homogeneous medium is often far from the truth. This may not always result in “bad” aquifer properties, because the well tests described before will integrate the heterogeneity of the aquifer over large areas. The heterogeneity of an aquifer is thus related to scale, and water flow may still be described well with the obtained aquifer properties. However, for some applications, it is more important to pay attention to the heterogeneity of aquifers, especially with respect to describing the behavior of contaminants.

As it is currently not possible to describe aquifer heterogeneity with certainty, the art of groundwater hydrology can be combined with statistical methods based on geological observations at the groundwater wells that have been installed in the past. This method has been described in detail by Carle et al. (1998). Using geostatistics, one can develop a set of scenarios that is likely to describe the aquifer geology. Through groundwater modelling (described in Chapter 23) and comparison with the observed hydrological data, separate aquifer properties for the defined hydrogeological units can then be estimated. It is important to note that the resulting properties are estimates, and that the description of the geological units are a representation of the statistical likeliness of occurrence that equal the statistical occurrence of the hydrogeological properties at the locations of the installed wells.

Methods exist to improve on the estimation of the hydrological properties. A variety of aquifer representations based on the same statistical inputs can be developed. The groundwater model can then be run multiple times, and a statistical distribution of the results can be created (Monte-Carlo analysis). This will result in an average value of the properties for each hydrogeological unit, and include a standard deviation for each property. As the hydrological properties of the actual aquifer system will vary spatially, a distributed value for the hydrological properties is likely to describe the actual situation better than a single value.

The equations shown in this chapter can all be run using a distribution of input parameters and will result in a distribution of the output. However, this is time intensive and may not necessarily result in a better understanding of the aquifer system. For groundwater resources development, sufficient other properties have uncertainties (as in thickness or extent of the pumped aquifer) that a closer representation of the actual aquifer properties will not improve groundwater management practices. However, for certain applications, especially contaminant transport a better understanding of the distribution of geological material and its property distribution may be needed.

Actual contaminant transport will be different in an aquifer when it is considered heterogeneous than when it is considered homogeneous. This is related to the fact that the bulk water flow, as used in theories assuming a homogenous aquifer system, are different than the aquifer properties that an actual water

or contaminant particle will encounter. Contaminant flow in a heterogeneous aquifer system will have a different flow path than the assumed bulk water flow (LaBolle et al., 2000). When the actual behavior of contaminants in a heterogeneous aquifer system can be estimated, groundwater remediation activities can be more optimally implemented.

10.7 Concluding Remarks

The diagnostic plots of time–drawdown data based on the well-flow equations are theoretical curves. The time–drawdown curves based on field data will often deviate from these theoretical shapes. These deviations can stem from the fact that one or more of the general assumptions and conditions listed in Section 10.2 are not met in the field, or that the method selected is not the correct one for the test site. Some of the common departures from the theoretical curves will now be discussed.

All real aquifers are limited by geological or hydrological boundaries. If, however, at the end of the pumping period, no such boundaries have been met within the cone of depression, it is said that the aquifer has a seemingly infinite areal extent. When the cone of depression intersects an impervious boundary (e.g., a fault or an impermeable valley wall), it can expand no farther in that direction. The cone must expand and deepen more rapidly at the fault or valley wall to maintain the yield of the well. This will result in time–drawdown plots in which the last part of the late-time drawdown data will deviate from a straight line under a slope. This part of the plot should be disregarded when the slope of the straight-line segment is being determined. The same phenomenon can occur when, in one of the directions, the sediments become finer and the hydraulic conductivity decreases. Well interference where two wells are close to each other, will also result in a similar phenomenon (see Chapter 11).

An opposite effect is encountered when the cone of depression intersects an open water body. If the open water body is hydraulically connected with the aquifer, the aquifer is recharged at an increasing rate as the cone of depression spreads with time. This results in a flattening of the slope of the time–drawdown curve at later times, which resembles the recharge that occurs in a leaky aquifer. The same phenomenon occurs when, in one of the directions, the hydraulic conductivity or the aquifer thickness increases. The above cases will result in time–drawdown plots in which the last part of the late-time drawdown data will deviate from a straight line under a slope. This part of the plot should be disregarded when the slope of the straight-line segment is being determined.

It will be clear that there are various reasons why time–drawdown data depart from the theoretical curves. It will also be clear that different phenomena can cause approximately the same anomalies. So, if one is to make a correct analysis, one must have a proper knowledge of the geology of the test site. As, unfortunately, this knowledge is often fragmentary, determining physical properties is more an art than science. This is one of the main reasons why it is strongly recommended to continue to monitor the water table behavior during the recovery period. This allows a second estimate of the aquifer’s transmissivity to be made, which can then be compared with the one found during the pumping period. Even with single-well tests, this second estimate is possible. With computer software, even a third estimate of the transmissivity and a second estimate for the specific yield or storativity are then possible (Boonstra and Kselik, 2001).

Finally, a few remarks on the difference between aquifer tests and single-well tests. The results of aquifer tests are more reliable and more accurate than those of single-well tests. Another advantage is that aquifer tests allow estimates to be made of both the aquifer’s transmissivity and its specific yield or storativity, which is not possible with single-well tests. Further, if an aquifer test uses more than one piezometer, separate estimates of the physical properties can be made for each well, allowing the various values to be compared. Moreover, one can make yet another estimate of the physical properties by using not only the time–drawdown relationship, but also the distance–drawdown relationships.

Glossary

Aquiclude: water-bearing layer in which both the horizontal and vertical flow components are so small that they can be neglected. In an aquiclude the groundwater flow is assumed to be zero.

Aquifer: water-bearing layer in which the vertical flow component with respect to the horizontal-flow component is so small that it can be neglected. In an aquifer the groundwater flow is assumed to be predominantly horizontal.

Aquitard: water-bearing layer in which the horizontal-flow component with respect to the vertical flow component is so small that it can be neglected. In an aquitard the groundwater flow is assumed to be predominantly vertical.

Drawdown: change in observed water level as function of pumping time with respect to original water level prior to pumping.

Homogeneous: an aquifer or aquitard is called homogeneous when its hydraulic conductivity is the same throughout such a water-bearing layer; its value is thus independent of the location.

Isotropic: an aquifer or aquitard is called isotropic when its hydraulic conductivity is the same in all directions at a particular location in such a water-bearing layer; its value is thus independent of the orientation.

Piezometer: small-diameter pipe placed in the subsoil so that there is no leakage around the pipe and all water enters the pipe through its open bottom. A piezometer indicates only the hydrostatic pressure of the groundwater at the specific point in the subsoil at its open lower end. To prevent clogging, piezometers are sometimes equipped with a short screen of a few decimeters length.

Residual drawdown: change in observed water level as function of recovery time with respect to original water level prior to pumping.

References

- Boonstra, J. 1992. Aquifer tests with partially penetrating wells: theory and practice. *J. Hydrol.*, 137, 165–179.
- Boonstra, J. and de Ridder, N.A. 1994. Single well and aquifer tests. In: *Drainage Principles and Applications*. 2nd ed. H.P. Ritzema (ed.). ILRI Publication 16, Wageningen, pp. 341–381.
- Boonstra, J. and Kselik, R.A.L. 2001. *SATEM 2002: Software for Aquifer Test Evaluation*. ILRI Publication 57, Wageningen, 148 p.
- Bos, M.G. (ed.) 1989. *Discharge Measurement Structures*. 3rd rev. ed. ILRI Publication 20, Wageningen, 401 p.
- Carle, S.F., Labolle, E.M., Weissmann, G.S., VanBrocklin, D., and Fogg, G.E. 1998. Geostatistical simulation of hydrofacies architecture: a transition probability/Markov approach. In: *Concepts in Hydrogeology and Environmental Geology No. 1, Hydrogeological Models of Sedimentary Aquifers*. G.S. Fraser and J.M. Davis (eds.). Society for Sedimentary Geology, Special Publication, Tulsa, Oklahoma, pp. 147–170.
- Cooper, H.H. and Jacob, C.E. 1946. A generalized graphical method for evaluating formation constants and summarizing well field history. *Am. Geophys. Union Trans.*, 27, 526–534.
- De Glee, G.J. 1930. *Over grondwaterstromingen bij wateronttrekking door middel van putten*. Waltman, Delft, 175 p.
- De Glee, G.J. 1951. Berekeningsmethoden voor de winning van grondwater. In: *Winning van grondwater : derde vacanciecursus in drinkwatervoorziening*. Delft University of Technology, pp. 38–80.
- Driscoll, F.G. 1986. *Groundwater and Wells*. 2nd ed. Johnson Division, St. Paul, 1089 p.
- Hantush, M.S. 1956. Analysis of data from pumping tests in leaky aquifers. *Am. Geophys. Union Trans.*, 37, 702–714.
- Hantush, M.S. 1962. Aquifer tests on partially-penetrating wells. *Am. Soc. Civ. Eng. Trans.*, 127, Part I, 284–308.
- Hantush, M.S. 1964. Hydraulics of wells. In: *Advanced Hydroscience 1*. V.T. Chow (ed.). Academic Press, New York, pp. 281–432.
- Hantush, M.S. and Jacob, C.E. 1955. Non-steady radial flow in an infinite leaky aquifer. *Am. Geophys. Union Trans.*, 36, 95–100.

- Jacob, C.E. 1944. Notes on determining permeability by pumping tests under watertable conditions. *U.S. Geol. Surv. Open File Report*.
- Jacob, C.E. 1950. Flow of groundwater. In: *Engineering Hydraulics*, H. Rouse (ed.). Wiley, New York, pp. 321–386.
- Kruseman, G.P. and de Ridder, N.A. 1994. *Analysis and Evaluation of Pumping Test Data*. ILRI Publication 47, Wageningen, 377 p.
- LaBolle, E.M., Quastel, J., Fogg, G.E., and Gravner, J. 2000. Diffusion processes in composite porous media and their numerical integration by random walks: generalized stochastic differential equations with discontinuous coefficients. *Water Resour. Res.*, 36, 651–662.
- Neuman, S.P. 1972. Theory of flow in unconfined aquifers considering delayed response of the watertable. *Water Resour. Res.*, 8, 1031–1045.
- Neuman, S.P. 1975. Analysis of pumping test data from anisotropic unconfined aquifers considering delayed gravity response. *Water Resour. Res.*, 11, 329–342.
- Theis, C.V. 1935. The relation between the lowering of the piezometric surface and the rate and duration of discharge of a well using groundwater storage. *Am. Geophys. Union Trans.*, 16, 519–524.
- Thiem, G. 1906. *Hydrologische Methoden*. Gebhardt, Leipzig, 56 p.
- Weeks, E.P. 1969. Determining the ratio of horizontal to vertical permeability by aquifer-test analysis. *Water Resources Res.*, Vol. 5, 196–214.

Further Information

Kruseman and de Ridder (1994) provide an excellent, well-illustrated handbook to well-flow equations that cover a wider range of conditions than here presented, that is, bounded, sloping, and consolidated fractured aquifers and their application to field data.

Boonstra and Kselik (2001) present a software package that facilitates the analysis of aquifer-test data of (un)confined and leaky aquifers, including partially penetrating wells, based on time–drawdown, distance–drawdown, and time–ratio–residual–drawdown relationships.

Appendix 10.1

Values of $K_0(r/L)$ and $e^{r/L}K_0(r/L)$ as Function of r/L

r/L	$K_0(r/L)$	$e^{r/L}K_0(r/L)$									
1.0(-2)	4.72	4.77	3.8(-2)	3.39	3.52	6.6(-2)	2.84	3.03	9.4(-2)	2.49	2.73
1.1(-2)	4.63	4.68	3.9(-2)	3.36	3.50	6.7(-2)	2.82	3.02	9.5(-2)	2.48	2.72
1.2(-2)	4.54	4.59	4.0(-2)	3.34	3.47	6.8(-2)	2.81	3.01	9.6(-2)	2.47	2.72
1.3(-2)	4.46	4.52	4.1(-2)	3.31	3.45	6.9(-2)	2.79	2.99	9.7(-2)	2.46	2.71
1.4(-2)	4.38	4.45	4.2(-2)	3.29	3.43	7.0(-2)	2.78	2.98	9.8(-2)	2.45	2.70
1.5(-2)	4.32	4.38	4.3(-2)	3.26	3.41	7.1(-2)	2.77	2.97	9.9(-2)	2.44	2.69
1.6(-2)	4.25	4.32	4.4(-2)	3.24	3.39	7.2(-2)	2.75	2.96	1.0(-1)	2.43	2.68
1.7(-2)	4.19	4.26	4.5(-2)	3.22	3.37	7.3(-2)	2.74	2.95	1.1(-1)	2.33	2.60
1.8(-2)	4.13	4.21	4.6(-2)	3.20	3.35	7.4(-2)	2.72	2.93	1.2(-1)	2.25	2.53
1.9(-2)	4.08	4.16	4.7(-2)	3.18	3.33	7.5(-2)	2.71	2.92	1.3(-1)	2.17	2.47
2.0(-2)	4.03	4.11	4.8(-2)	3.15	3.31	7.6(-2)	2.70	2.91	1.4(-1)	2.10	2.41
2.1(-2)	3.98	4.06	4.9(-2)	3.13	3.29	7.7(-2)	2.69	2.90	1.5(-1)	2.03	2.36
2.2(-2)	3.93	4.02	5.0(-2)	3.11	3.27	7.8(-2)	2.67	2.89	1.6(-1)	1.97	2.31
2.3(-2)	3.89	3.98	5.1(-2)	3.09	3.26	7.9(-2)	2.66	2.88	1.7(-1)	1.91	2.26
2.4(-2)	3.85	3.94	5.2(-2)	3.08	3.24	8.0(-2)	2.65	2.87	1.8(-1)	1.85	2.22
2.5(-2)	3.81	3.90	5.3(-2)	3.06	3.22	8.1(-2)	2.64	2.86	1.9(-1)	1.80	2.18
2.6(-2)	3.77	3.87	5.4(-2)	3.04	3.21	8.2(-2)	2.62	2.85	2.0(-1)	1.75	2.14
2.7(-2)	3.73	3.83	5.5(-2)	3.02	3.19	8.3(-2)	2.61	2.84	2.1(-1)	1.71	2.10
2.8(-2)	3.69	3.80	5.6(-2)	3.00	3.17	8.4(-2)	2.60	2.83	2.2(-1)	1.66	2.07
2.9(-2)	3.66	3.76	5.7(-2)	2.98	3.16	8.5(-2)	2.59	2.82	2.3(-1)	1.62	2.04
3.0(-2)	3.62	3.73	5.8(-2)	2.97	3.14	8.6(-2)	2.58	2.81	2.4(-1)	1.58	2.01
3.1(-2)	3.59	3.70	5.9(-2)	2.95	3.13	8.7(-2)	2.56	2.80	2.5(-1)	1.54	1.98
3.2(-2)	3.56	3.67	6.0(-2)	2.93	3.11	8.8(-2)	2.55	2.79	2.6(-1)	1.50	1.95
3.3(-2)	3.53	3.65	6.1(-2)	2.92	3.10	8.9(-2)	2.54	2.78	2.7(-1)	1.47	1.93
3.4(-2)	3.50	3.62	6.2(-2)	2.90	3.09	9.0(-2)	2.53	2.77	2.8(-1)	1.44	1.90
3.5(-2)	3.47	3.59	6.3(-2)	2.88	3.07	9.1(-2)	2.52	2.76	2.9(-1)	1.40	1.88
3.6(-2)	3.44	3.57	6.4(-2)	2.87	3.06	9.2(-2)	2.51	2.75	3.0(-1)	1.37	1.85
3.7(-2)	3.41	3.54	6.5(-2)	2.85	3.04	9.3(-2)	2.50	2.74	3.1(-1)	1.34	1.83

3.2(-1)	1.31	6.0(-1)	7.78(-1)	1.42	8.8(-1)	5.01(-1)	1.21	2.6
3.3(-1)	1.29	6.1(-1)	7.65(-1)	1.41	8.9(-1)	4.94(-1)	1.20	2.7
3.4(-1)	1.26	6.2(-1)	7.52(-1)	1.40	9.0(-1)	4.87(-1)	1.20	2.8
3.5(-1)	1.23	6.3(-1)	7.40(-1)	1.39	9.1(-1)	4.80(-1)	1.19	2.9
3.6(-1)	1.21	6.4(-1)	7.28(-1)	1.38	9.2(-1)	4.73(-1)	1.19	3.0
3.7(-1)	1.18	6.5(-1)	7.16(-1)	1.37	9.3(-1)	4.66(-1)	1.18	3.1
3.8(-1)	1.16	6.6(-1)	7.04(-1)	1.36	9.4(-1)	4.59(-1)	1.18	3.2
3.9(-1)	1.14	6.7(-1)	6.93(-1)	1.35	9.5(-1)	4.52(-1)	1.17	3.3
4.0(-1)	1.11	6.8(-1)	6.82(-1)	1.35	9.6(-1)	4.46(-1)	1.16	3.4
4.1(-1)	1.09	6.9(-1)	6.71(-1)	1.34	9.7(-1)	4.40(-1)	1.16	3.5
4.2(-1)	1.07	7.0(-1)	6.61(-1)	1.33	9.8(-1)	4.33(-1)	1.15	3.6
4.3(-1)	1.05	7.1(-1)	6.50(-1)	1.32	9.9(-1)	4.27(-1)	1.15	3.7
4.4(-1)	1.03	7.2(-1)	6.40(-1)	1.31	1.0	4.21(-1)	1.14	3.8
4.5(-1)	1.01	7.3(-1)	6.30(-1)	1.31	1.1	3.66(-1)	1.10	3.9
4.6(-1)	9.94(-1)	7.4(-1)	6.20(-1)	1.30	1.2	3.19(-1)	1.06	4.0
4.7(-1)	9.76(-1)	7.5(-1)	6.11(-1)	1.29	1.3	2.78(-1)	1.02	4.1
4.8(-1)	9.58(-1)	7.6(-1)	6.01(-1)	1.29	1.4	2.44(-1)	9.88(-1)	4.2
4.9(-1)	9.41(-1)	7.7(-1)	5.92(-1)	1.28	1.5	2.14(-1)	9.58(-1)	4.3
5.0(-1)	9.24(-1)	7.8(-1)	5.83(-1)	1.27	1.6	1.88(-1)	9.31(-1)	4.4
5.1(-1)	9.08(-1)	7.9(-1)	5.74(-1)	1.26	1.7	1.65(-1)	9.06(-1)	4.5
5.2(-1)	8.92(-1)	8.0(-1)	5.65(-1)	1.26	1.8	1.46(-1)	8.83(-1)	4.6
5.3(-1)	8.77(-1)	8.1(-1)	5.57(-1)	1.25	1.9	1.29(-1)	8.61(-1)	4.7
5.4(-1)	8.61(-1)	8.2(-1)	5.48(-1)	1.25	2.0	1.14(-1)	8.42(-1)	4.8
5.5(-1)	8.47(-1)	8.3(-1)	5.40(-1)	1.24	2.1	1.01(-1)	8.25(-1)	4.9
5.6(-1)	8.32(-1)	8.4(-1)	5.32(-1)	1.23	2.2	8.93(-2)	8.06(-1)	5.0
5.7(-1)	8.18(-1)	8.5(-1)	5.24(-1)	1.23	2.3	7.91(-2)	7.89(-1)	5.48(-1)
5.8(-1)	8.04(-1)	8.6(-1)	5.16(-1)	1.22	2.4	7.02(-2)	7.74(-1)	
5.9(-1)	7.91(-1)	8.7(-1)	5.09(-1)	1.21	2.5	6.23(-2)	7.60(-1)	

Note: 3.7(-2) means 3.7×10^{-2} or 0.037.

Example: $r/L = 5.0(-1) = 5 \times 10^{-1} = 0.5$ $K_0(0.5) = 0.924$ $e^{0.5} K_0(0.5) = 1.52$.

11

Well Design and Construction

11.1	Introduction.....	11-1
11.2	Well-Field Equations	11-2
	Partial Penetration	
11.3	Aquifer and Well Losses	11-5
	Well Efficiency	
11.4	Step-Drawdown Tests	11-7
	Jacob's Method • Rorabaugh's Method	
11.5	Well Design	11-13
	Casing Section • Calculation Example Length of Pump	
	Housing • Screen Section • Calculation Example Well	
	Screen • Gravel Pack • Sand Trap • Pump • Design	
	Optimization • Groundwater Remediation Wells •	
	Monitoring Wells	
11.6	Well Construction and Maintenance	11-24
	Construction Methods • Well Development • Maintenance	
	Glossary.....	11-30
	References	11-30
	Further Information	11-30

Hans Boonstra and
Richard Soppe
Alterra-ILRI, The Netherlands

11.1 Introduction

Wells are drilled either for exploration or exploitation purposes. With exploration wells, the objective is to collect information on the geology of the underlying aquifer. Apart from describing the various geological layers encountered (well logs), the physical properties are commonly assessed by conducting aquifer tests. These aquifer tests are commonly preceded by step-drawdown tests to determine the proper discharge rate. Proper in this respect means that good measurable drawdowns are produced in the piezometers, and that the drawdown in the pumped well does not reach the top of the well screen during the pumping period.

For exploitation or production wells, the objective can be summarized as pumping the required amount of water at the lowest cost considering investment, operation, and maintenance. Exploitation wells are drilled for water supply for municipal, industrial, and irrigation purposes, and for watertable control for drainage purposes. In the latter two cases, the exploitation wells are usually placed in the so-called well fields; well fields are also required when the required pumping capacity for municipal and industrial purposes cannot be met by a single exploitation well.

In the design of exploitation wells, two main aspects can be distinguished: the maximum expected drawdown in the well and the production capacity of the underlying aquifer system. The maximum expected drawdown determines the length of the pump casing section and is mainly composed of aquifer losses and well losses. To these aquifer losses, interference effects and partially penetrating effects should be added, when applicable. The production capacity of the underlying aquifer system determines the minimum length of the screen section that mainly depends on maximum screen velocities and screen type. To this end, well-field equations and step-drawdown tests are presented to relate aquifer losses to discharge rate and recharge rate of the aquifer, and to determine well losses and production capacity of exploration wells that will be used later on as exploitation wells. Step-drawdown tests are also required to analyze well performance of exploitation wells with time for maintenance and rehabilitation purposes.

11.2 Well-Field Equations

Chapter 10 described the flow to single wells pumping in extensive aquifers. It is usually assumed that the aquifer is not recharged during the relatively short pumping period of aquifer tests. For exploitation wells this assumption is not realistic as they are operated on a more or less continuous basis over a long time span. Most aquifer systems receive continuous or intermittent recharge from percolating rainfall and losses from surface water systems. When recharge is intermittent because of seasonal effects, the aquifer may perform without recharge for periods of several months or even longer. There is clear evidence that after such a period watertables will completely recover once the aquifer is recharged again by rainfall.

In this section, we assume that the aquifer is replenished at a constant rate, R , expressed as a volume per unit surface per unit of time ($\text{m}^3/\text{m}^2 \text{ d}$). The well-flow equations that will be presented are based on a steady-state situation. The flow is said to be in a steady state as soon as the recharge and the discharge balance each other. In such a situation, beyond a certain distance from the well, there will be no drawdown induced by pumping. This distance is called the radius of influence of the well, r_e . When the wells in a well field are placed in a triangular pattern, their individual radii of influence hardly overlap (Figure 11.1). The simplifying assumption is then made that the discharge and the drawdown of each well will not be affected by those of neighboring wells. In other words, the theory of a single well can be used.

In a well field, there is a direct relationship between the discharge rate of the well, the recharge rate of the aquifer by percolation, and the area affected by pumping. The decline of the water level due to

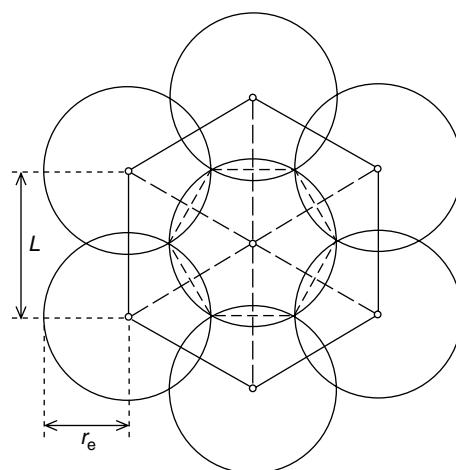


FIGURE 11.1 Layout of a well field where the wells are located in a pattern of equilateral triangles ($L = r_e\sqrt{3}$).

pumping is determined by the discharge rate of the well and the permeability and thickness of the aquifer. The discharge rate and the drawdown in the well are important factors in calculating the pumping costs of well fields. In an unconfined aquifer, the steady-state flow through an arbitrary cylinder at a distance r from the well is given by

$$Q_r = \pi(r_e^2 - r^2)R \quad (11.1)$$

where Q_r is the volumetric flow rate in m^3/d at a distance r from the well, r_e is the radius of influence of the well in meters, and R is the recharge rate of the aquifer per unit surface area in m/d . According to Darcy's law, Q_r equals the product of the cylindrical area of flow and the flow velocity. Hence, the discharge at distance r from the well can also be expressed by

$$Q_r = 2\pi rhK \frac{\partial h}{\partial r} \quad (11.2)$$

where h is the hydraulic head in meters, K is the hydraulic conductivity of the aquifer in m/d , and $\partial h/\partial r$ is the hydraulic gradient in the aquifer at distance r . As, in steady state, the discharge of the well, Q , equals the vertical recharge of the area within the radius of influence, the following relationship can be used

$$Q = \pi r_e^2 R \quad (11.3)$$

Combining Equation 11.1 and Equation 11.3 yields

$$Q_r = Q - \pi r^2 R \quad (11.4)$$

or, combining Equation 11.2 and Equation 11.4 and separating r and h gives

$$\left(\frac{Q}{r} - \pi r R \right) \partial r = 2\pi K h \partial h \quad (11.5)$$

Integration between the limits $r = r_w$, $h = h_w$, and $r = r_e$, $h = h_e$ yields

$$Q \ln \left(\frac{r_e}{r_w} \right) - \frac{1}{2} \pi R (r_e^2 - r_w^2) = \pi K (h_e^2 - h_w^2) \quad (11.6)$$

The quantity $1/2\pi R r_w^2$ is very small in comparison with $1/2\pi R r_e^2$ and can be neglected. If, moreover, the drawdown in the well is small in comparison with the original hydraulic head, the right-hand side of Equation 11.6 can be expressed as (Peterson et al., 1952)

$$\pi K (h_e + h_w)(h_e - h_w) \approx \pi K 2D(h_e - h_w) = 2\pi K D \Delta h_r \quad (11.7)$$

where D is the saturated thickness of the aquifer before pumping in meters and Δh_r is the drawdown due to radial flow toward the pumped well in meters. As, according to Equation 11.3,

$$r_e^2 = \frac{Q}{\pi R} \quad (11.8)$$

Equation 11.6 can be written as

$$\Delta h_r = \frac{Q}{2\pi K D} \left[2.3 \log \left(\frac{r_e}{r_w} \right) - \frac{1}{2} \right] \quad (11.9)$$

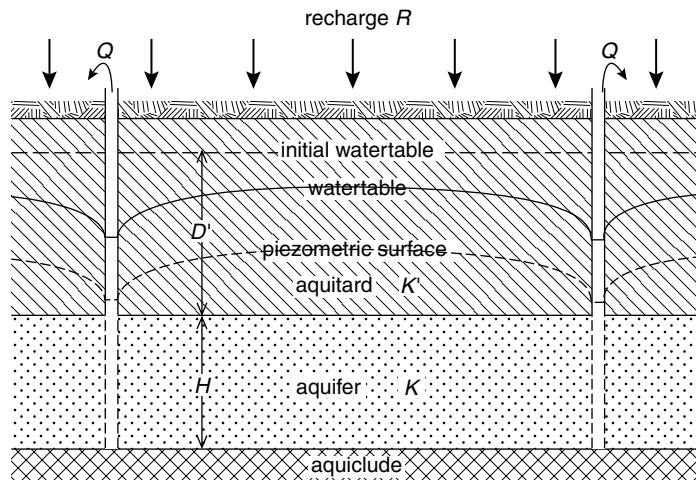


FIGURE 11.2 Schematic cross-section of wells located in a leaky aquifer.

If $r_e/r_w > 100$, and if we accept an error of 10%, the term $-1/2$ can be neglected and Equation 11.9 reduces to

$$\Delta h_r = \frac{2.3Q}{2\pi T} \log \frac{r_e}{r_w} \quad (11.10)$$

where $T = KD$ is the transmissivity of the aquifer in m^2/d . Equation 11.10 can be used to calculate the drawdown in a single well or, when no interference between the wells is assumed, the drawdown in a well field. From Figure 11.1, it can be seen that the distance between the wells (L) should then be equal to $r_e\sqrt{3}$.

Figure 11.2 shows a leaky aquifer whose overlying layer, the aquitard, is replenished by percolating rain or excess irrigation water at a constant rate R . Depending on the recharge rate and the hydraulic resistance of the aquitard, a difference in head between the free watertable in the aquitard and the piezometric level of the aquifer will develop. Under steady-state conditions, the same recharge rate will replenish the underlying aquifer. So, Equation 11.10 can also be used to calculate the head loss in a well or well field for leaky aquifers.

The application of Equation 11.10 will now be illustrated with an example calculation. In Chapter 10, data were presented for a single-well test. The analysis of this test yielded a transmissivity value of some $1400 \text{ m}^2/\text{d}$. Suppose that from a water-balance analysis, the long-term average net recharge to the aquifer was estimated as $550 \text{ mm/yr} (=1.5 \text{ mm/d})$. To calculate the drawdown in a well with a constant discharge of $200 \text{ m}^3/\text{h}$, the radius of influence needs to be calculated. Using Equation 11.3, the radius of influence will be

$$r_e = \sqrt{\frac{200 \times 24}{3.14 \times 0.0015}} = 1010 \text{ m}$$

Substituting this value into Equation 11.10 with an outer well screen radius (r_w) of 0.2 m gives

$$\Delta h_r = \frac{2.3 \times 200 \times 24}{2 \times 3.14 \times 1400} \log \frac{1010}{0.2} = 4.65 \text{ m}$$

So, the steady-state drawdown in a single well due to radial flow amounts to 4.65 m . A similar drawdown can be expected when wells are placed in a well field with a triangular pattern. The distance between the wells in Figure 11.1 should then be $L = 1010 \times 1.73 = 1747 \text{ m}$; it is then assumed that there will be no mutual interference between the various wells. If the wells are located at shorter distances or in a

TABLE 11.1 Comparison between Pumping Capacities, Drawdown due to Aquifer Losses, and Distances between Wells (Triangular Pattern) for a Net Recharge of 1.5 mm/d

Pumping capacity (m ³ /h)	Radius of Influence (m)	Aquifer Losses (m)	Distance Between Wells (m)	Area Per Well (ha)
100	714	2.23	1236	160
200	1010	4.65	1747	320
300	1236	7.14	2141	480

rectangular pattern, the drawdown in the wells will increase due to well interference; this phenomenon will be numerically illustrated in Section 11.5.

In well-field design for drainage purposes, there is a trade-off between the capacity and the number of wells required as is shown in Table 11.1. If larger pumps are installed, fewer pumps will be required that generally results in lower investment costs. On the other hand, larger capacity pumps may result in considerably higher drawdowns and thus higher energy costs. Determining pump capacities on a purely economic basis could lead to very high pumping rates. There are, however, several practical constraints to these high pump capacities for drainage purposes. A well with a very high pump capacity may serve a very large area that exceeds the spacing determined by other factors. If such a well were to be out of order for a prolonged period, the neighboring wells would be overburdened, and proper drainage of the area would be impossible. Moreover, if the water is also used for irrigation, pump capacities are often limited by the requirements of the farmers.

11.2.1 Partial Penetration

Equation 11.10 was derived under the assumption that the well fully penetrates the pumped aquifer. Some aquifers are so thick, however, that installing a fully penetrating well would not be justified. In these cases, the aquifer has to be pumped by a partially penetrating well. This partial penetration causes the flow velocity in the immediate vicinity of the well to be higher than it would otherwise be, leading to an extra loss of head. According to Hantush (1964), the effect of partial penetration in an unconfined aquifer is similar to that in a confined aquifer, as the drawdown is usually small in relation to the saturated thickness of the aquifer. To calculate the effect of partial penetration, the following equation can be used

$$\Delta h_p = \frac{Q}{4\pi T} F \quad (11.11)$$

and

$$F = 2 \frac{D}{p} \left[\left(1 - \frac{p}{D} \right) \ln \left(\frac{2p}{r_w} \right) - \frac{p}{D} \ln \frac{2D}{p} - 0.423 \frac{p}{D} + \ln \frac{2D+p}{2D-p} \right] \quad (11.12)$$

where Δh_p is the additional head loss due to partial penetration effects in meters, Q is the well discharge in m³/d, T is the transmissivity of the aquifer in m²/d, D is the thickness of the aquifer in meters, and p is penetration depth of the well into the aquifer in meters, assuming screening over the full depth of the well. So, when the well only partially penetrates the aquifer, the additional head loss calculated from Equation 11.11 should be added to the drawdown calculated by Equation 11.10.

11.3 Aquifer and Well Losses

The drawdown in a pumped well consists of two components: the aquifer losses and the well losses (Figure 11.3). Aquifer losses are the head losses that occur in the aquifer where the flow is laminar. They

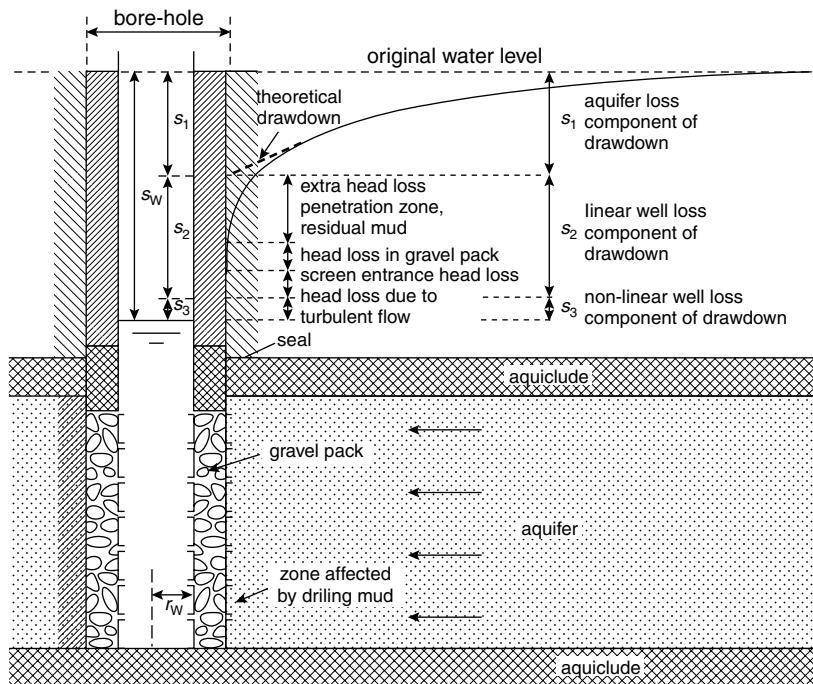


FIGURE 11.3 Various components of head losses in a pumped well.

are time dependent and vary linearly with the well discharge. The drawdown s_1 corresponding to this linear aquifer loss can be expressed as

$$s_1 = B_1(r_w, t) Q \quad (11.13)$$

where B_1 is the linear aquifer loss coefficient in d/m^2 . This coefficient can be calculated from the well-flow equations presented in Chapter 10. For confined aquifers for example, it can be expressed using Equation 10.1 and Equation 10.2 in Chapter 10 as

$$B_1(r_w, t) = \frac{W(u)}{4\pi T}$$

where $u = (r_w^2 S)/(4Tt)$. From the results of aquifer-test analyses, the values for transmissivity T and storativity S can be used to calculate B_1 values as function of r_w and t .

Well losses are divided into linear and nonlinear head losses. Linear well losses are caused by the damaging of the aquifer during drilling and completion of the well. They comprise, for example, head losses due to the compaction of the aquifer material during drilling; head losses due to plugging of the aquifer with drilling mud, that reduce the permeability near the bore hole; head losses in the gravel pack; and head losses in the screen. The drawdown s_2 corresponding to this linear well loss can be expressed as

$$s_2 = B_2 Q \quad (11.14)$$

where B_2 is the linear well loss coefficient in d/m^2 .

Among the nonlinear well losses are the friction losses that occur inside the well screen and in the suction pipe where the flow is turbulent, and head losses that occur in the zone adjacent to the well where the flow is usually also turbulent. All these losses responsible for the drawdown inside the well being much

greater than one would expect on theoretical grounds. The drawdown s_3 corresponding to this nonlinear well loss can be expressed as

$$s_3 = CQ^P \quad (11.15)$$

where C is the nonlinear well loss coefficient in $\text{d}^P/\text{m}^{3P-1}$, and P is an exponent. The general equation describing the drawdown in a pumped well as function of aquifer/well losses and discharge rate thus reads

$$s_w = (B_1 + B_2)Q + CQ^P = BQ + CQ^P \quad (11.16)$$

where $s_w = s_1 + s_2 + s_3$. Jacob (1947) used a constant value of 2 for the exponent P . According to Lennox (1966) the value of P can vary between 1.5 and 3.5; from personal experience, its value may be even higher in fractured rock aquifers. The value of $P = 2$ as proposed by Jacob, is however still widely accepted. Values of the three parameters B , C , and P in Equation 11.16 can be found from the analysis of so-called step-drawdown tests.

11.3.1 Well Efficiency

The relationship between drawdown and discharge can be expressed as the *specific capacity* of a well, Q/s_w , which describes the productivity of both the aquifer and the well. The specific capacity is not a constant but decreases as pumping continues and also decreases with increasing Q . The well efficiency, E_w , is defined as the ratio of the aquifer head loss to the total head losses; it reads when expressed as a percentage,

$$E_w = \left\{ \frac{B_1 Q}{BQ + CQ^P} \right\} \times 100\% \quad (11.17)$$

The well efficiency according to Equation 11.17 can be assessed when both the results of a step-drawdown and those of an aquifer test are available. The former are needed for the values of B , C , and P and the latter for the value of B_1 .

In practice, only the results of a step-drawdown tests are usually available. The substitution of the B , C , and P values into Equation 11.17 would overestimate the well efficiency, as $B > B_1$. For these cases, Driscoll (1986) introduced another parameter, L_p , being the ratio of the laminar head loss to the total head losses; it reads when expressed as a percentage,

$$L_p = \left\{ \frac{BQ}{BQ + CQ^P} \right\} \times 100\% \quad (11.18)$$

It should be noted that Equation 11.18 is sometimes erroneously used to calculate the well efficiency.

11.4 Step-Drawdown Tests

A step-drawdown test is a single-well test in which the well is pumped at a low constant-discharge rate until the drawdown within the well stabilizes. The pumping rate is then increased to a higher constant-discharge rate and the well is pumped until the drawdown stabilizes once more (Figure 11.4). This process is repeated through at least three steps, which should all be of equal duration, say a few hours each.

In step-drawdown analyses, the so-called diagnostic plots are used. Values of s_w/Q vs. Q are therefore plotted on arithmetic paper where s_w represents the drawdown at the end of each step. Various

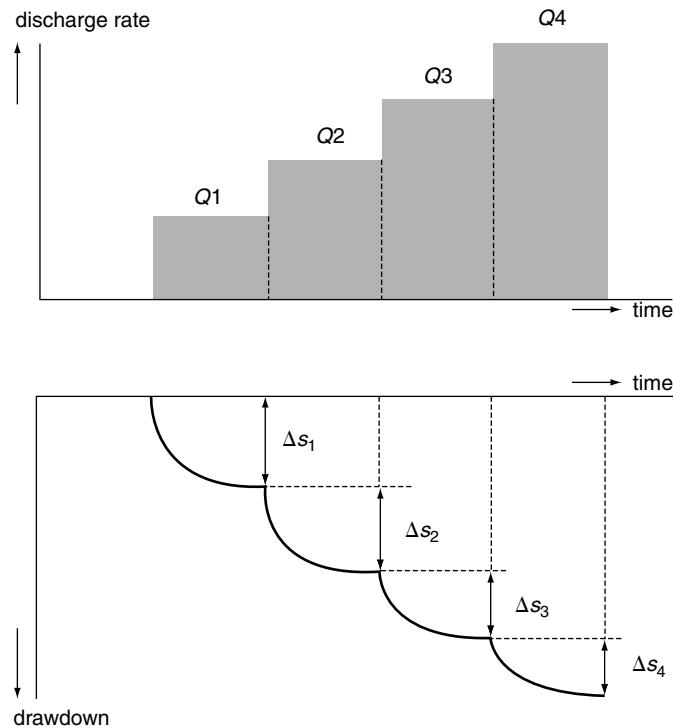


FIGURE 11.4 Principles of a step–drawdown test.

configurations of diagnostic plots are then possible:

1. The points fall on a horizontal line. This implies that $s_w/Q = B$. Equation 11.16 reduces to

$$s_w = BQ \quad (11.19)$$

Hence, there are no nonlinear well losses. This situation is only encountered with very low pumping rates. The well will act differently if the pumping rates are increased.

2. The points fall on a straight line under a slope. This means that $s_w/Q = B + CQ$. Equation 11.16 then reduces to

$$s_w = BQ + CQ^2 \quad (11.20)$$

Equation 11.20 is known as the Jacob's equation. Based on this equation, Jacob (1947) developed an analysis method to calculate the values of B and C .

3. The points fall on a curved line, that is, $P \neq 2$ in Equation 11.16. When a concave curve can be drawn through the points, it implies that $P > 2$ and for a convex curve that $P < 2$. For these cases, Rorabaugh (1953) developed an analysis method to calculate the values of B , C , and P . Both analysis methods may be applied to confined, unconfined, and leaky aquifers.

11.4.1 Jacob's Method

The values of B and C can be directly found from the diagnostic plot of s_w/Q vs. Q itself; it will yield a straight line whose slope is equal to C ; the value of B can be found by extending the straight line till it intercepts the $Q = 0$ axis.

TABLE 11.2 Time–Drawdown Values Observed in the Pumped Well of a Step–Drawdown Test

Time in Each Step (min)	Step 1 $Q: 2765$ m^3/d (m)	Step 2 $Q: 3364$ m^3/d (m)	Step 3 $Q: 3853$ m^3/d (m)	Step 4 $Q: 4306$ m^3/d (m)
2	2.476	3.614	4.541	5.301
3	2.601	3.611	4.491	5.304
4	2.657	3.629	4.511	5.299
5	2.636	3.650	4.527	5.317
6	2.669	3.641	4.517	5.339
7	2.624	3.630	4.546	5.371
8	2.677	3.683	4.509	5.391
9	2.684	3.709	4.554	5.377
10	2.691	3.714	4.539	5.300
12	2.699	3.691	4.553	5.319
14	2.704	3.710	4.559	5.289
20	2.769	3.721	4.574	5.369
30	2.814	3.762	4.613	5.404
40	2.841	3.786	4.609	5.404
50	2.850	3.794	4.637	5.440
60	2.893	3.822	4.639	5.464

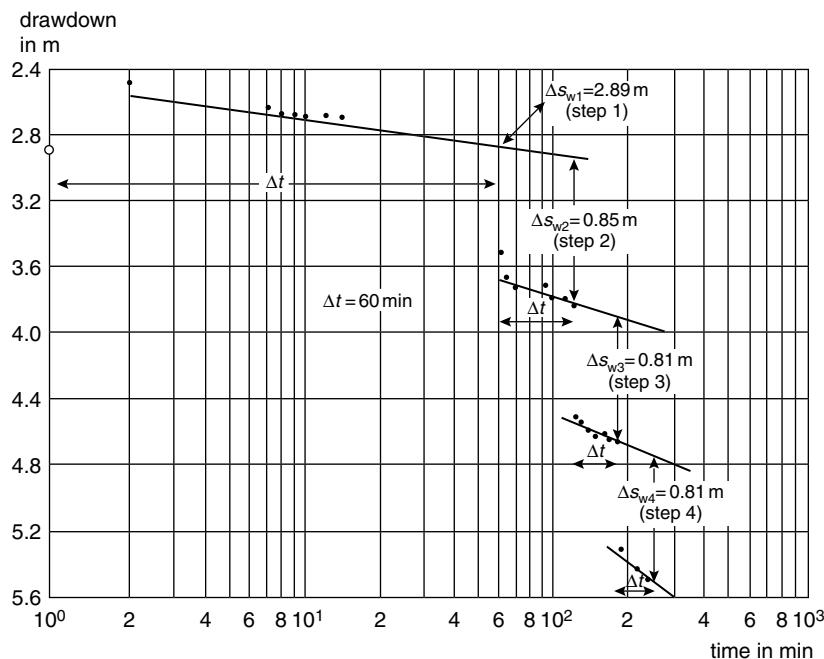


FIGURE 11.5 Time–drawdown plot of field data of a step–drawdown test in an unconfined aquifer.

This procedure will now be illustrated with field data. A step–drawdown test was carried out in an unconfined aquifer. The well was pumped with a step-wise increased discharge rate with a duration of each step of 60 min. Table 11.2 shows the observed drawdowns as function of time. Figure 11.5 shows the time–drawdown plot on semi-log paper; from this figure it can be observed that the drawdowns were not yet stabilized at the end of each step. This implies that the observed drawdown values at the

TABLE 11.3 Specific Drawdown Determined from Semi-Log Plot

Step	Δs_w (m)	s_w (m)	Q (m^3/d)	s_w/Q (d/m^2)
1	2.89	2.89	2765	1.05×10^{-3}
2	0.85	3.74	3364	1.11×10^{-3}
3	0.81	4.55	3853	1.18×10^{-3}
4	0.81	5.36	4306	1.24×10^{-3}

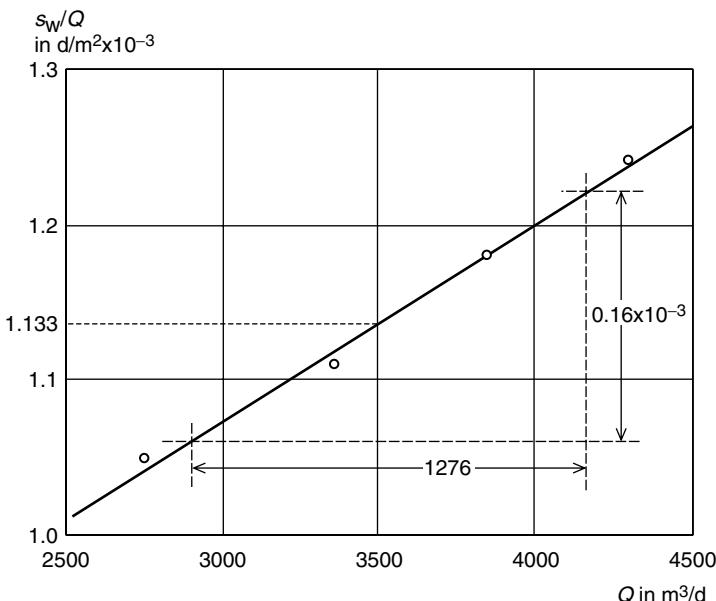


FIGURE 11.6 Diagnostic plot of s_w/Q vs. Q of field data of a step-drawdown test with the Jacob's analysis method.

end of each step need to be corrected. This can be done using the following procedure as developed by Hantush-Bierschenk (1964):

1. Extrapolate the curve through the plotted data of each step to the end of the next step.
2. Determine the increments of drawdowns $\Delta s_{w(i)}$ for each step by taking the difference between the observed drawdown at a fixed interval Δt , taken from the beginning of each step, and the corresponding drawdown on the extrapolated curve of the preceding step.
3. Determine the values of $s_{w(n)}$ corresponding to the discharge Q_n from $s_{w(n)} = \Delta s_{w(1)} + \Delta s_{w(2)} + \dots + \Delta s_{w(n)}$.

From Figure 11.5 the drawdown differences for each step were determined using a time interval $\Delta t = 60$ min. Based on these values, the specific drawdown values s_w/Q at the end of each step were calculated (see Table 11.3). Figure 11.6 shows the resulting diagnostic plot of s_w/Q vs. Q ; the data plot clearly exhibits a straight line. This implies that $P = 2$ (Equation 11.20), so the Jacob's method can be applied. Through the points a best fitting straight line was drawn; its slope was calculated as $0.16 \times 10^{-3}/1276 = 1.25 \times 10^{-7} d^2/m^5$. The interception point of the straight line with the $Q = 0$ axis can be calculated as follows: $1.133 \times 10^{-3} - 3500 \times 1.25 \times 10^{-7} = 7.0 \times 10^{-4} d/m^2$. Hence, the drawdown equation of this well can be written as

$$s_w = 7.0 \times 10^{-4} Q + 1.3 \times 10^{-7} \times Q^2 \quad (11.21)$$

This step-drawdown test was followed by a single-well test. For that test, the discharge rate of the third step was adopted, being $3853 \text{ m}^3/\text{d}$. According to Equation 11.15 the nonlinear well loss s_3 can then be calculated as

$$s_3 = CQ^P = 1.3 \times 10^{-7} \times 3853^2 = 1.93 \text{ m}$$

This non-linear well loss was subtracted from the drawdowns observed in the pumped well (see Chapter 10 under the heading “Single-Well Tests”). The analysis of these corrected drawdowns resulted in a transmissivity of $1410 \text{ m}^2/\text{d}$ and a specific yield of 0.025 for the tested aquifer. Based on these values, the theoretical drawdown according to the Theis equation (Equation [10.1] in Chapter 10) at $t = 60 \text{ min}$ can be calculated being 2.06 m. The linear aquifer loss coefficient B_1 is then according to Equation 11.13

$$B_1 = \frac{s_1}{Q} = \frac{2.06}{3853} = 5.35 \times 10^{-4} \text{ d/m}^2$$

and the linear well loss coefficient B_2 is

$$B_2 = B - B_1 = 7.0 \times 10^{-4} - 5.35 \times 10^{-4} = 1.65 \times 10^{-4} \text{ d/m}^2$$

The well efficiency for the discharge rate of $3853 \text{ m}^3/\text{d}$ can then be calculated as 45% according to Equation 11.17. If Equation 11.18 would have been used as being representative of the well efficiency, it would have resulted in a value of 58%, that is, in an overestimation of the true well efficiency.

11.4.2 Rorabaugh's Method

The values of B , C , and P cannot be directly found from the diagnostic plot of s_w/Q vs. Q itself, when $P \neq 2$. Equation 11.16 then reads

$$\frac{s_w}{Q} = B + CQ^{P-1} \quad (11.22)$$

Rearranging Equation 11.22 and taking the logarithms, Equation 11.22 can also be written as

$$\log \left[\frac{s_w}{Q} - B \right] = \log C + (P - 1) \log Q \quad (11.23)$$

Equation 11.23 implies that a plot of $(s_w/Q - B)$ vs. Q on log-log paper would yield a straight line under a slope. The slope of this straight line is equal to $P - 1$, while the value of C can be found by extending the straight line till it intercepts the $Q = 1$ axis.

This procedure will now be illustrated with field data from a step-drawdown test. Table 11.4 shows the drawdowns observed at the end of each step (Sheahan, 1971). Figure 11.7 shows the resulting diagnostic

TABLE 11.4 Step-Drawdown Data

Step	s_w (m)	Q (m^3/d)	s_w/Q (d/m^2)
1	2.62	2180	1.2×10^{-3}
2	6.10	3815	1.6×10^{-3}
3	17.22	6540	2.6×10^{-3}
4	42.98	9811	4.4×10^{-3}

Source: From Sheahan, N.T. 1971. *Ground Water*, Vol. 9, pp. 25–29.

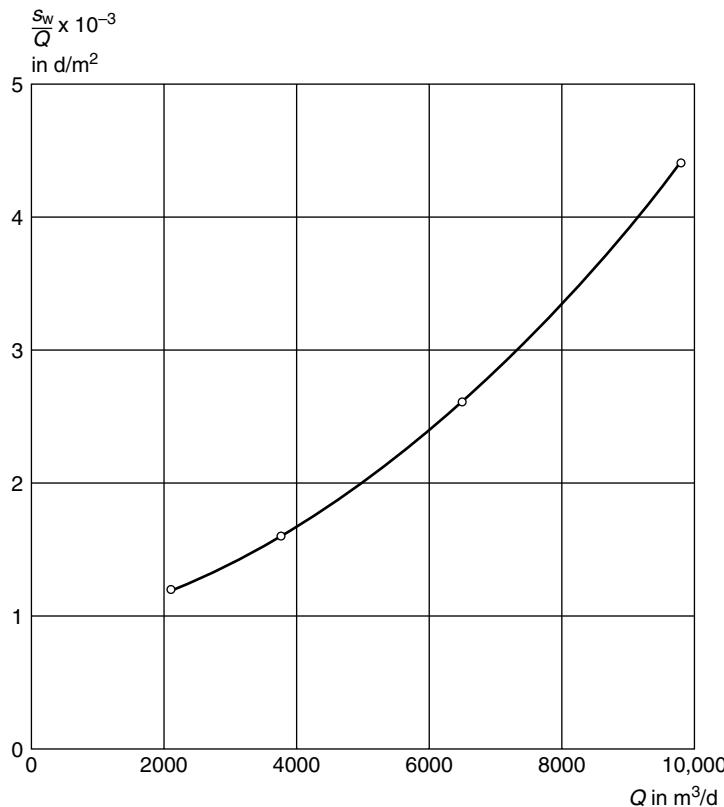


FIGURE 11.7 Diagnostic plot of s_w/Q vs. Q of field data of a step-drawdown test.

TABLE 11.5 Values of $(s_w/Q - B_i)$ as Function of B_i as used with Rorabaugh's Method

Step	$B_1 = 0$ (d/m ²)	$B_2 = 0.8 \times 10^{-3}$ (d/m ²)	$B_3 = 1 \times 10^{-3}$ (d/m ²)	$B_4 = 1.1 \times 10^{-3}$ (d/m ²)
1	1.2×10^{-3}	0.4×10^{-3}	0.2×10^{-3}	0.1×10^{-3}
2	1.6×10^{-3}	0.8×10^{-3}	0.6×10^{-3}	0.5×10^{-3}
3	2.6×10^{-3}	1.8×10^{-3}	1.6×10^{-3}	1.5×10^{-3}
4	4.4×10^{-3}	3.6×10^{-3}	3.4×10^{-3}	3.3×10^{-3}

plot of s_w/Q vs. Q ; this plot clearly shows that the points do not lie on a straight line. This implies that $P \neq 2$, so the Jacob's method cannot be applied. For different values of B (see Table 11.5), the values of $(s_w/Q - B)$ vs. Q were plotted on log-log paper. Figure 11.8 shows that for a B value of 1×10^{-3} d/m² the plotted points, almost exhibit a straight line. Through these points a best fitting straight line was drawn; its slope was calculated as

$$\frac{\Delta[s_w/Q - B_3]}{\Delta Q} = \frac{\log(0.2 \times 10^{-3}/3.4 \times 10^{-3})}{\log(2180/9811)} = 1.88$$

As the slope of this line equals $(P-1)$, it follows that $P = 2.88$. The interception point of the straight line with the $Q = 1$ axis can be calculated as follows. Figure 11.8 shows that the value of $(s_w/Q - B)$ for

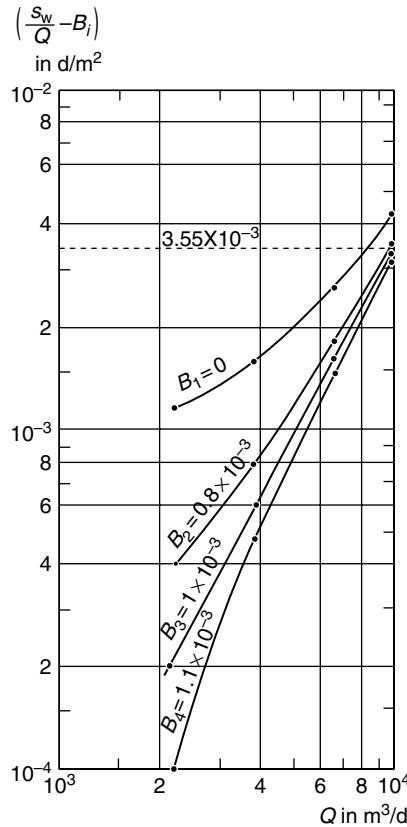


FIGURE 11.8 Log-log plot of $(s_w/Q - B)$ vs. Q to determine the parameters B , C , and P according to the Rorabaugh's method.

$Q = 10^4 \text{ m}^3/\text{d}$ is $3.55 \times 10^{-3} \text{ d/m}^2$. Hence, the intersection of the line with the $Q = 1$ axis is four log cycles to the left. This corresponds with $4 \times 1.88 = 7.52$ log cycles below the point $(s_w/Q - B) = 3.55 \times 10^{-3} \text{ d/m}^2$.

The intersection point $(s_w/Q - B)$ is then calculated as follows: $\log(s_w/Q - B) = \log(3.55 \times 10^{-3}) - 7.52 = -2.45 - 7.52 = -9.97$, thus $(s_w/Q - B) = 1.1 \times 10^{-10} = C$. Hence, the drawdown equation of this well can be written as

$$s_w = 1.0 \times 10^{-3}Q + 1.1 \times 10^{-10}Q^{2.88}$$

Based on the above relationship, the optimum discharge rate can be determined, when this well would be converted to an exploitation well.

11.5 Well Design

Exploitation wells are drilled either for water supply for municipal, industrial, and irrigation purposes or for water table control for drainage purposes. The principle objectives of a properly designed well are (1) pumping of water at the lowest cost; (2) pumping of water that is free of sand and silt; (3) minimum operation and maintenance costs; and (4) a long and economic lifetime. For water supply wells, good quality water with proper protection from contamination is an important additional objective.

Hydrogeologic information required for a proper design includes: stratigraphical information concerning the aquifer and overlying sediments, aquifer-test analyses of the physical properties of the

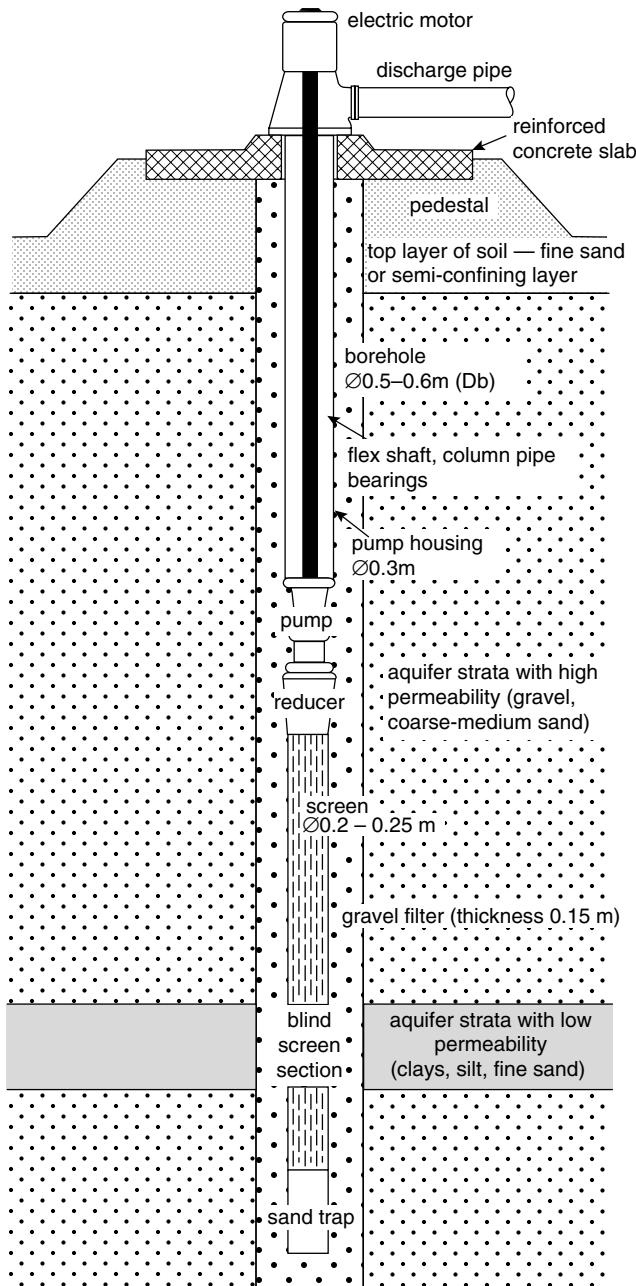


FIGURE 11.9 Typical well design.

aquifer, water-balance analyses of the sustainable yield, grain-size analyses of unconsolidated aquifer materials, and groundwater quality.

Well design is the process of specifying the physical materials and dimensions for a well. A good well design depends on many factors, some of which are discussed below. More detailed information on technical well design can be found in reference books such as those of Driscoll (1986) and Huisman (1975). Figure 11.9 shows the various well sections of a typical well design. The purpose and design of these well sections, and their position in the well, will be discussed below.

TABLE 11.6 Maximum Pumping Rates for Certain Diameters of Standard-Weight Casing, Based on an Upward Velocity of 1.5 m/sec

Casing Size (in.)	Maximum Pumping Rate (mm) ^a	(m ³ /d)
4	102	1,090
5	127	1,690
6	152	2,450
8	203	4,250
10	254	6,700
12	305	9,590
14	337	11,700
16	387	15,500
18	438	19,800
20	489	24,700
25	591	36,100

^a Actual inside diameter.

Source: From Driscoll, F.G. 1986. *Groundwater and Wells*. St. Paul, Johnson Division, Minnesota, 1089 p.

11.5.1 Casing Section

The pump housing is the upper section of blind casing that supports the well against collapse, and in which the pump is installed. The length of the pump housing should be chosen so that the pump remains below the water level in the well, for the selected discharge rate, under all conditions, and over the total lifetime of the well. A pump housing is always required when submersible pumps are used. No special pump housing is required in the case of a shallow water table with little drawdown where suction pumps can be used; both pump and engine are then installed at the surface beside the well. The diameter of the pump housing should be large enough to accommodate the pump with enough clearance for installation and efficient operation. It is recommended that the pump housing be two pipe sizes larger than the nominal diameter of the pump; the diameter of the pump depends on the selected discharge rate and the pump type.

The production casing is the lower section of the blind pipe between the bottom of the pump housing and the top of the aquifer. The production casing is not required in unconfined aquifers at shallow depth where the pump housing reaches sufficiently deep into the top section of the aquifer. The length of the production casing depends on the thickness of the aquitard overlying the pumped aquifer. To minimize the head losses in the production casing itself, the upward velocity of the pumped water should be less than 1.5 m/sec. Based on this criterion, Table 11.6 shows casing sizes recommended for various pumping rates; for the pipe sizes and pumping rates shown in this table, the head losses will be small. Moreover, the diameter of the production casing should be smaller than the diameter of the pump housing and should be larger or equal to the diameter of the underlying screen section.

11.5.2 Calculation Example Length of Pump Housing

In this calculation example, the results of the calculations presented under the heading “Well-Field Equations” and the results of the step-drawdown test presented under the heading “Jacob’s Method” will be combined with the results of the aquifer test presented under the heading “Single-Well Tests” in Chapter 10.

The actual length of the pump housing is primarily determined by the required depth of the pump. The location of the pump depends on the expected depth to which the water level inside the well will drop for

the selected design discharge rate. The procedure to determine the maximum expected water-level depth inside the pumped well is as follows.

The first step is to determine the aquifer losses. The aquifer losses are the head losses due to the laminar flow of water to the well and are determined by the hydraulic conductivity and thickness of the aquifer. Equation 11.10 can be used to estimate these losses when the wells are located in a triangular well field. Based on a constant recharge of 1.5 mm/d, a transmissivity of the aquifer of $1400 \text{ m}^2/\text{d}$ and an outer radius of the well screen of 0.2 m, the resulting aquifer losses were presented in Table 11.1 assuming that there will be no interference effects of neighboring wells. When the recharge is intermittent rather than continuous because of seasonal effects, the aquifer may perform without recharge for periods of several months or even longer. The aquifer losses should then be calculated using the equations presented in Chapter 10. Based on a specific yield of the aquifer of 0.025 together with the above transmissivity and well screen data, Equation 10.3 of Chapter 10 yields the following aquifer loss when it is assumed that the aquifer is not replenished during half a year

$$\Delta h_r = \frac{2.3 \times 200 \times 24}{4 \times 3.14 \times 1400} \log \frac{2.25 \times 1400 \times 182}{0.20^2 \times 0.025} = 5.50 \text{ m}$$

The effect of well interference should now be considered. For a discharge rate of $200 \text{ m}^3/\text{h}$, the well spacing was calculated as 1747 m (Table 11.1). The effect of well interference from one neighboring well can then be calculated as

$$\Delta h_r = \frac{2.3 \times 200 \times 24}{4 \times 3.14 \times 1400} \log \frac{2.25 \times 1400 \times 182}{1747^2 \times 0.025} = 0.55 \text{ m}$$

When the well field consists of seven wells as indicated in Figure 11.1, the aquifer loss of the center well can be calculated as $5.50 + 6 \times 0.55 = 8.8 \text{ m}$. Table 11.7 shows these aquifer losses for discharge rates and distances between the wells according to Table 11.1. With partially penetrating tubewells, the additional head loss according to Equation 11.11 should be added to these aquifer losses.

The second step is to determine the well losses from the results of the step-drawdown test. Based on the values of C and P in Equation 11.21, Equation 11.15 gives the following nonlinear loss component

$$s_3 = CQ^P = 1.3 \times 10^{-7} \times (200 \times 24)^2 = 3.00 \text{ m}$$

Table 11.7 shows the contribution of the nonlinear well losses to the total head losses as function of the discharge rate. Such calculations of the nonlinear well loss based on the results of a step-drawdown test can only be applied to wells with designs similar to that of the tested well. For wells with different designs, Huisman (1975) presented methods involving rather complicated calculations to estimate their expected well losses.

The third step is to assess the maximum depth to the water table; this information can be found from analyzing groundwater hydrographs showing the seasonal fluctuations of the water table throughout the years. Suppose that this maximum water table depth is 8 m below land surface. The required minimum length of the pump housing for the center well of the well field as shown in Figure 11.1 would be

TABLE 11.7 Various Components of the Total Head Loss Based on Data of Table 11.1 After Half a Year without Recharge

Pumping Rate (m^3/h)	Aquifer Loss Components (m)		Non-linear Well Loss (m)	Total Head Loss (m)
	Single Well	Well Interference		
100	2.75	2.22	0.75	5.72
200	5.50	3.30	3.00	11.80
300	8.24	3.96	6.74	18.94

$11.80 + 8.00 = 19.80$ m for a discharge rate of $200 \text{ m}^3/\text{h}$. Usually an additional length of several meters is added for safety. So, the total length of the pump housing can be taken as 23 m for this well.

When wells will be used for drainage purposes, the required design depth to the water table should be used instead of the maximum expected depth to the water table. In leaky aquifers, the difference in head between the free water table in the overlying aquitard and the piezometric head of the pump should then also be added; this head loss depends on the hydraulic resistance of the aquitard and the *drainable surplus*. Representative values of the former can be found from aquifer-test analyses, while values of the latter can be found from water-balance analyses.

11.5.3 Screen Section

Important properties of the screen are that it prevents sand and fine material from entering the well during pumping, has a large percentage of open area to minimize the head loss and entrance velocity, supports the wall of the well against collapse, and is resistant to chemical and physical corrosion by the pumped water.

Polyvinyl chloride (PVC) and fiberglass screens are lighter and more resistant to corrosion by chemically aggressive water, but have a lower collapse strength than steel screens and casings. In practice, PVC and fiberglass-reinforced screens and casings will be technically and economically attractive for wells in alluvial aquifers, where wells are placed at moderate depths of up to 400 m. Steel screens are required in deep wells drilled in hard rock aquifers. Stainless steel screens combine both strength and resistance to corrosion and chemically aggressive water, but are more expensive.

The selection of the screen slot size depends on the type of aquifer and the use of a gravel pack. The screen slot size must be selected so as to ensure that most of the finer materials in the formation around the borehole are transported to the screen and removed from the well by bailing and pumping during the well-development period immediately after the borehole has been constructed and the screen and casing have been installed.

In wells without an artificial gravel pack, well development creates a zone of graded formation materials extending about 0.5 m outward from the screen. Driscoll (1986) and Huisman (1975), among others, give detailed procedures for selecting the correct slot size. They report that with good quality water and the correct slot opening, 60% of the material will pass through the screen and 40% will be retained. With corrosive water the 50%-retained size should be chosen, as even a small enlargement of the slot openings due to corrosion could cause sand to be pumped.

The screen length should be chosen so as to ensure that the actual screen entrance velocity is in accordance with the prescribed entrance velocities as listed in Table 11.8. From these screen entrance velocities, the minimum length of the well screen can be calculated from

$$Q = 86400 v_e l_{\min} A_0 \quad (11.24)$$

where Q is the discharge rate of the well in m^3/d , v_e is the screen entrance velocity in m/sec , l_{\min} is the minimum screen length in meters, and A_0 is the effective open area per meter screen length in m^2/m . In determining the effective open area per meter screen length, it is often assumed that 50% of the actual open area is clogged by gravel particles (Huisman, 1975). The actual open area per meter screen length depends on the type and diameter of the selected screen type. Conventional slotted screens have open areas not exceeding 10% in order not to weaken the column strength, whereas more expensive continuous slot screens of stainless steel or modern PVC screens have an open area of 30 to 50%. So, the minimum total screen length is determined by the maximum screen entrance velocity and the actual screen type.

The optimum length of the screen may differ from its minimum length. Determining the optimum screen length is rather complex; it depends on (1) all the cost factors that determine the costs of pumping the required discharge or draining 1 ha; (2) the total thickness of the aquifer. In very thick aquifers, the deeper penetration of the well will result in a smaller drawdown that reduces the pumping costs but increases the investment costs in the borehole; and (3) the selected pumping rate.

TABLE 11.8 Recommended Screen Entrance Velocities

Hydraulic Conductivity of Aquifer (m/d)	Screen Entrance Velocities (m/sec)
>250	>0.03
250–120	0.03
120–100	0.025
100–40	0.02
40–20	0.015
<20	<0.01

Source: From U.S. Environmental Protection Agency, 1975. *Manual of Water Well Construction Practices*. EPA-570/9-75-001. Office of water supply, Washington, DC. 156 p. With permission.

The total length of the required screen section is found by adding to the actual screen length, as outlined above, the total length of sections of blind (imperforated) pipe used to case off unproductive layers in the aquifer. The total length of blind pipe depends on the distribution of hydraulic conductivity in the aquifer (i.e., the distribution of layers of higher and lower hydraulic conductivity). This stratification can be determined from the driller's log, geophysical log, and sieve analysis.

11.5.4 Calculation Example Well Screen

In this calculation example, the same data will be used as for the calculation of the length of the pump housing. The productive layers as delineated from the well logs had a total thickness of 56 m. The corresponding hydraulic conductivity of the aquifer is then 25 m/d. From Table 11.8, the screen entrance velocity should then not exceed 0.015 m/sec. For a well screen with an open area of 20% and a diameter of 0.25 m, the effective open area per meter screen length can be calculated as, bearing in mind a clogging percentage of 50%

$$A_0 = 3.14 \times 0.25 \times 0.5 \times 0.20 = 0.08 \text{ m}^2/\text{m}$$

Substituting the above values into Equation 11.24 for a pumping capacity of $200 \text{ m}^3/\text{h} = 4800 \text{ m}^3/\text{d}$ yields the following minimum screen length

$$l_{\min} = \frac{4800}{86400 \times 0.015 \times 0.08} = 47 \text{ m}$$

Table 11.9 shows how this minimum screen length varies for different screen diameters (0.15, 0.20, and 0.25 m), and for different types of screens: cheap well screens with an open area of 10%, medium-priced well screens with an open area of 20%, and expensive, modern, continuously slotted well screens with an open area of up to 40%. The total available thickness of the productive layers was 56 m, so according to Table 11.9, only the expensive well screens for all the listed diameters and the medium-priced well screen with a diameter of 0.25 m can be used for a well with a discharge rate of $200 \text{ m}^3/\text{h}$.

For this discharge rate ($=4800 \text{ m}^3/\text{d}$), the diameter of the production casing connecting the pump housing with the well screen, should be 0.25 m according to Table 11.6. As its diameter should be larger or equal to the diameter of the underlying screen section, the above selected well screens can all be applied.

TABLE 11.9 Minimum Screen Lengths for Different Types of Screens and Pump Discharges

Diameter (m)	Open Area		Minimum Screen Length (m)		
	(%)	(m ² /m)	$Q = 100 \text{ m}^3/\text{h}$	$Q = 200 \text{ m}^3/\text{h}$	$Q = 300 \text{ m}^3/\text{h}$
0.15	10	0.024	77	154	231
0.15	20	0.047	39	79	118
0.15	40	0.094	20	39	59
0.20	10	0.031	60	119	179
0.20	20	0.063	29	59	88
0.20	40	0.126	15	29	44
0.25	10	0.039	47	95	142
0.25	20	0.079	23	47	70
0.25	40	0.157	12	24	35

11.5.5 Gravel Pack

The effect of gravel-packed wells is to ensure that the zone around the well screen is made more permeable by removing some formation material and replacing it with specially graded material. This relatively narrow zone separates the screen from the formation material and increases the effective hydraulic diameter of the well. A gravel pack is chosen to retain most of the formation material; a well screen opening is then selected to retain about 90% of the gravel pack after development. Gravel-pack material should ideally be clean, rounded, siliceous sands or gravels; carbonate material, shale particles, or soluble material such as gypsum should not exceed 5% of the total. Gravel-pack material should be well sorted to assure good porosity and hydraulic conductivity of these materials around the screen. The application of a gravel pack is recommended in the following formations:

1. Fine sandy alluvium and aeolian sand aquifers. In these formations, gravel packing should be considered so that larger slot openings can be used to increase the hydraulic efficiency of the well. In general, if a slot opening based on natural development is smaller than 0.25 mm, gravel packing may be more desirable as the screen's transmitting capacity may not be great enough to supply the desired yield.
2. Alternating formations of fine, medium, and coarse sediment. In these formations, it is often difficult to determine precisely the position and thickness of each individual layer and to design a multiple-slot screen corresponding to the stratification.
3. Poorly cemented sandstone aquifers. These formations may continuously lose fine material resulting in a sand-pumping well. Another reason for gravel packing a sandstone aquifer is that the formation material usually provides little or no lateral support for the screen.

The gravel pack is designed on the basis of sieve analyses of aquifer samples. If aquifer samples from different depths show considerable variation in gradation, the gravel-pack design should be based to be stable against the finer-grade samples. Numerous investigators and agencies have experimented to develop formulae or criteria that will result in a stable gravel-pack gradation. According to Anderson (1995), the following criteria have generally been found satisfactory in actual practice:

1. Aquifer material with a *uniformity coefficient* less than 2.5: Use uniform gravel-pack material with a uniformity coefficient less than 2.5 and with the D_{50} of the gravel pack 4 to 6 times the D_{50} of the aquifer (Figure 11.10). If uniform gravel pack is not available, use a gravel pack with uniformity coefficient between 2.5 and 5 and with the D_{50} of the gravel pack not more than 9 times the D_{50} of the aquifer.
2. Aquifer material with uniformity coefficient between 2.5 and 5: Use uniform gravel-pack material with uniformity coefficient less than 2.5 and with the D_{50} of the gravel pack not more than 9 times the D_{50} of the aquifer. If uniform gravel pack is not available, use a gravel pack with uniformity

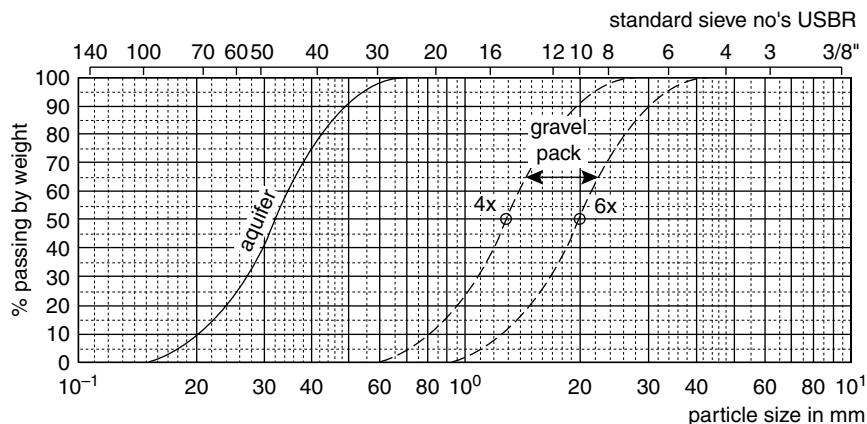


FIGURE 11.10 Aquifer and filter pack gradation.

coefficient between 2.5 and 5 and with the D_{50} of the gravel pack not more than 12 times the D_{50} of the aquifer.

3. Aquifer material with uniformity coefficient greater than 5: Multiply D_{30} of the aquifer by 6 and 9 and locate these points on the sieve analysis graph. Draw two parallel lines through these points having a uniformity coefficient of 2.5 or less, and specify gravel-pack material that will fall between these lines.

Although smaller thicknesses already fulfill the objective of a gravel pack, the thickness of a gravel pack should at least be 76 mm to ensure that a continuous layer of filter material will surround the entire screen. Under most conditions, the upper limit of gravel-pack thickness should be about 200 mm as the energy created by the development procedure must be able to penetrate the pack to repair the damage done by drilling, break down any residual drilling fluid on the borehole wall, and remove finer particles near the borehole.

11.5.6 Sand Trap

The sand trap is the section of blind pipe at the bottom of the screen section. Its function is to store sand and silt entering the well during pumping; this will occur even if the well has been properly developed. The length of the sand trap is usually of the order of a few meters (2–6 m). The diameter of the sand trap is usually the same as that of the screen section.

11.5.7 Pump

Centrifugal pumps are by far the most important class of pumps. Originally designed as a pump to be located at or near land surface for suction lift or booster service, it soon was adapted to installation under water in wells, first by long shafts extensions in large caissons (vertical turbine), and later in compact form as the familiar deep-well (submersible) turbine pump. A typical line-shaft pumping unit comprises a driver (power unit), discharge head, column (discharge) pipe and shaft, and one or more stages of *impellers* in a bowl assembly. The driver is commonly a vertical hollow-shaft electric motor, while the shaft is centered in the column pipe. A typical submersible pumping unit comprises a submersible electric motor directly coupled to the bowl assembly and a column (discharge) pipe.

The major factors determining the selection of a pump are the required discharge rate, the required head to be delivered by the pump, and the pump efficiency; the total dynamic head is made up of (1) the water-level depth inside the pumped well, as discussed earlier in this section; (2) the above ground lift; and (3) head losses due to friction and turbulence in the discharge pipelines. The required brake horsepower

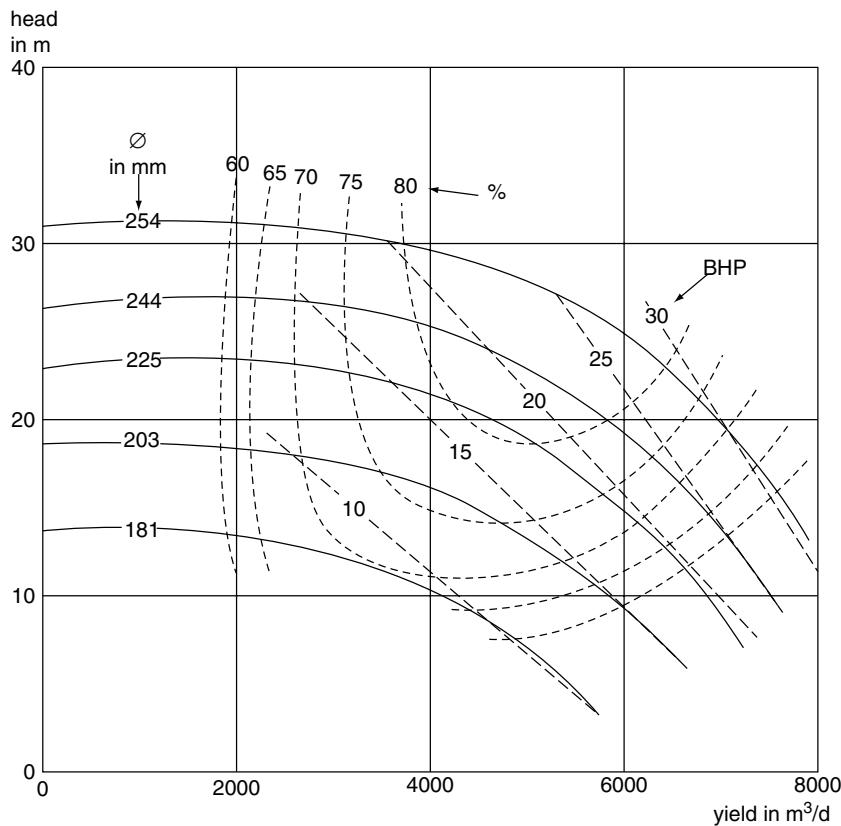


FIGURE 11.11 Representative performance curve for a vertical turbine pump operating at 1800 rpm.

at the pump shaft is defined as

$$P_s = \frac{\rho g Q H}{\eta} \quad (11.25)$$

where P_s is the power to be delivered to the shaft of the pump in Watts, Q is the pump discharge in m^3/sec , H is the total dynamic head delivered by the pump in meters, ρ is the density of water in kg/m^3 , g is the acceleration due to gravity in m/sec^2 , and η is the dimensionless pump efficiency.

Most pump manufacturers provide performance curves for each individual impeller design and for various rotation speeds. Generally, these curves are drawn for a single stage (pumping unit). These curves typically include (1) a head-capacity curve relating the total dynamic head to the discharge rate, (2) a curve relating efficiency to discharge rate, and (3) a curve relating brake horsepower requirement to discharge rate. From the performance curves, a particular impeller is chosen that can provide the desired yield for the total dynamic head at reasonable efficiency. The performance curve shown in Figure 11.11 indicates that a single-stage turbine pump with a 254 mm impeller operates at a peak efficiency of about 80% within a total head of 23 to 30 m and delivers from 6400 to 3700 m^3/d , respectively. If the required head fluctuates within this range, the pump operates efficiently and the power required varies by 7 brake horsepower.

In selecting a turbine pump, it is usually desirable to pick an impeller type that has a steep rather than flat head-capacity curve. This is particularly if the pumping lift will change significantly during the seasons of the year, or if the pumping lift is expected to increase over future years. With a steep curve, a moderate increase in pumping lift will not result in a large decrease in the discharge rate, whereas with a flat curve, even a slight increase in pumping lift can reduce the discharge by 50% or more. With steep head-capacity

curves, the peak efficiency might be lower, but the average pumping conditions would occur at a higher efficiency.

11.5.8 Design Optimization

The design of a well is rather straightforward when the water transmitting properties of the aquifer system are limited with respect to the required discharge rate. The well will be drilled to the bottom of the aquifer and all the productive layers of the aquifer will be screened using screens with the largest possible slots and open area. When the required pumping capacity is not restricted by the water transmitting properties of the aquifer system, that is, with highly productive, thick aquifers, and modest pumping requirements, the design becomes more complex, especially when well fields are considered as is usually the situation with wells for irrigation and drainage purposes. Some of the optimizing options are then:

1. If larger pumps are installed, fewer pumps will be required, which generally results in lower investment costs; on the other hand, larger capacity pumps may result in considerably higher drawdowns and thus higher energy costs.
2. Water can be brought to the land surface by a well with a short screen or a long screen. A short screen involves low investment costs and high energy costs because of greater drawdowns in the wells, while a long screen entails relatively higher investment costs but lower energy costs.

The optimization procedure involves examining the different well configurations that satisfy the design criteria, and, for each of these, calculating the investment costs and the annual costs of operation and maintenance. The present value of these costs is determined by applying an annual rate for discounting costs and an interest rate. The configuration that yields the lowest present value is then selected. Boehmer and Boonstra (1994) provides a checklist on how the investment costs can be calculated for a well with a pre-fixed pumping capacity and located in a given well-field layout.

With increasing screen lengths, investment costs rise but the drawdown and the pumping costs fall. The cost per cubic meter of water first decreases, owing to the decreasing head losses caused by the decreasing partial penetration of the aquifer, and leads to lower pumping costs. Having reached a minimum, the price of water rises, as the decreasing energy costs in the borehole no longer compensate for the higher investment costs. The calculations are repeated for different screen diameters, screen types, pump engine, and types of energy. Figure 11.12 shows the relation between the cost of water, investment and reinvestment costs, energy costs, and operation and maintenance costs.

Finally, the design with the lowest costs per cubic meter drainage water is selected. Obviously, the more types of screens, engines, pumps, and energy available, the better the results of the optimization procedure will be. The number of calculations required to arrive at a final result is large and complex and can best be handled by an optimum well-field-design computer program as advised by Boehmer and Boonstra (1994).

11.5.9 Groundwater Remediation Wells

As mentioned above, well designs can be optimized for the cost of a single well, but also for a well field. In a well field, it is important to realize that over-designed wells will negatively affect other wells in the surroundings. Optimization can then be obtained on economics. However, if the objective is not only economics, the optimization will not necessarily result in the cheapest option.

In the past decades, attention has been shifting from water resources development to water resources management. This includes an increased interest in maintaining or improving groundwater quality. At some locations, especially where groundwater pollution is decreasing groundwater resources for drinking water, remediation wells are installed. The objective of these wells is to either intercept pollution plumes before they reach the drinking water wells, or, when installed near the source of the pollution, even to reverse the hydraulic gradients, so that pollution is prevented from further spreading, and returns to the

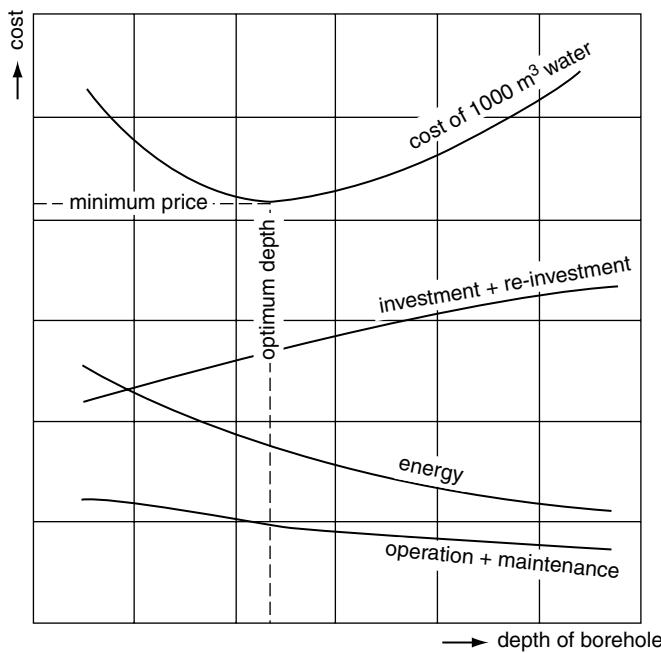


FIGURE 11.12 Costs of 1000 m^3 water as a function of borehole depth.

point of origin. A well field designed for this situation will preferably be designed to have sufficient overlap in the sphere of influence to protect human health.

Many design and construction parameters that are used for water development wells are also applicable to the remediation wells. Screening depth and length should be designed to intercept the maximum amount of pollution at the lowest possible costs. The well field should be designed to have maximum coverage over the pollution plume. Pump capacity should be selected in combination with the water treatment plant on the surface, to prevent higher flow rates than the capacity of the treatment plant. One special point of attention is that the gravel pack around the well is sealed at the bottom and the top of the screen. Especially in a multi-layer aquifer system, it is of major importance to prevent pollution to spread to other systems through the well.

Another special point of importance is the heterogeneity of the aquifer system. Although with current methods it is impossible to describe the heterogeneity accurately, it is possible to obtain a statistical likeliness of the geology (Carle and Fogg, 1996). Especially in groundwater remediation this is important, as small clusters of clay in a sandy aquifer could have a long retention time of pollutants.

11.5.10 Monitoring Wells

Another well type that is important in the management of groundwater is the monitoring well. Groundwater monitoring, preferably by the groundwater users themselves, may result in local groundwater management without the need of intervention from governing bodies (Steenbergen, 2003). Groundwater monitoring may be done using existing wells, which gives a good impression of the effects of groundwater pumping. However, if more detailed information is needed on either vertical or horizontal groundwater flows, additional monitoring wells could be installed.

Nested piezometers are often used to determine the vertical gradients in an aquifer system. Nested piezometers are a set of observation wells with varying depth, usually 3–4. As they are designed to measure the pressure at a defined depth in the aquifer system, it is important to design the well, and especially the screen depth. The screens should be installed at locations where a gradient could be expected, for example,

just above an aquitard, and just below. During the construction of the observation wells, the sealing is, like in the remediation wells, very important. As gradients in the aquifer system are expected (otherwise there would be no need to install nested piezometers), one should prevent the piezometers from being conduits for flow between different aquifers.

The spatial design of a well-monitoring field should also be based on the expected outcomes. When low regional lateral groundwater flow is expected, observation wells can be installed at large distances from each other. The area covered by monitoring wells should preferably be surrounded by obvious hydrological boundaries for the groundwater system, like rivers, large irrigation canals, or mountain ranges. Observation wells usually have a small diameter; it is advisable, however, to allow for a sufficient size so that portable submersible pumps can be lowered into the well. Observation wells can then also be used for groundwater sampling. Note that for groundwater sampling it is important to develop an observation well similar to a production well. Well development is described in the next paragraph.

11.6 Well Construction and Maintenance

Various well drilling methods have been developed as geologic conditions range from hard rock such as granite and dolomite to completely unconsolidated sediments such as alluvial sand and gravel. Particular drilling methods have become dominant in certain areas as they are the most effective in penetrating the local aquifers and thus offer cost advantages. In many cases, however, the drilling contractor may vary the usual drilling procedure depending on the depth and diameter of the well, type of formation to be penetrated, sanitation requirements, and principal use of the well. It is obvious, then, that no single drilling method is best for all geologic conditions and well installations. In most cases, the drilling contractor is best qualified to select the particular drilling procedure for a given set of construction parameters. Successful drilling is both an art developed from long experience and the application of good engineering practices.

Well construction usually comprises four of five distinct operations: drilling, installing the casing, placing a well screen and filter pack, if required, grouting to provide sanitary protection, and developing the well to insure sand-free operation at maximum yield. Well drilling and installation methods are so numerous that only some of the basic principles and their applications can be described in this section. More detailed information on well construction methods can be found in reference books such as those of Driscoll (1986) and Huisman (1975).

11.6.1 Construction Methods

In the various construction methods, a distinction can be made between shallow and deep wells. Shallow wells, generally less than 15 m in depth, are constructed by boring, driving, or jetting. These methods are now briefly described and their applications are given in Table 11.10. Bored wells are constructed with hand-operated or power-driven augers. The drive shaft is built up from rods; its upper part, the so-called kelly, has a square cross-section to receive the necessary torque from a rotary table. At the bottom, the auger is provided with a cutting face; for emptying the auger, the drive shaft should be dismantled and reassembled afterwards. A continuous-flight power auger has a spiral extending from the bottom of the hole to the surface. Cuttings are carried to the surface as on a screw conveyor, while sections of the auger may be added as depth increases. Driven wells are constructed by driving a pointed screen with attached pipe directly into the waterbearing formation. Water enters the well through a drive point at the lower end of the well. This consists of a screened cylindrical section protected during driving by a steel cone at the bottom. Driving can be done with a sledge, drop hammer, or air hammer. As driving proceeds and the well point sinks into the ground, succeeding sections of pipe are screwed on top of the screen. Jetted wells are constructed by the cutting action of a downward-directed stream of water. The high-velocity stream washes the earth away, while the casing, which is lowered into the deepening hole, conducts the water and

cuttings up and out of the well. The soil material is subsequently removed in a settling basin after which the water is picked up again by the jetting pump.

Most deep, high-capacity wells are constructed by cable-tool percussion drilling or by one of the several rotary methods. These methods are now briefly described and their applications are given in Table 11.10. Cable-tool drilling machines, also called percussion rigs, operate by repeatedly lifting and dropping a heavy string of drilling tools into the borehole. The drill bit breaks or crushes consolidated rock into small fragments, whereas the bit primarily loosens the material when drilling in unconsolidated formations. In both instances, the reciprocating action of the tools mixes the crushed or loosened particles with water to form a slurry at the bottom of the borehole. If little or no water is present in the penetrated formations, water is added to form a slurry. Slurry accumulation increases as drilling proceeds and eventually it reduces the impact of the tools. When the penetration rate becomes unacceptable, slurry is removed at intervals from the borehole by a sand pump or bailer. In water well drilling, the depth capability for cable rigs ranges from 100 to 1500 m.

The direct rotary drilling method was developed to increase drilling speeds and to reach greater depth in most formations. The borehole is drilled by rotating a bit, and cuttings are removed by continuous circulation of a drilling fluid as the bit penetrates the formation. The bit is attached to the lower end of a string of drill pipe, which transmit the rotating action from the rig to the bit (Figure 11.13). In the direct rotary system, drilling fluid is pumped down through the drill pipe and out through the ports or jets in the bit; the fluid then flows upward in the annular space between the hole and drill pipe, carrying the cuttings in suspension to the surface. The fluid is cleaned in a settling pit before it is recirculated; drilling fluids include air, clean water, and scientifically prepared mixtures of special-purpose materials.

In direct rotary drilling, the viscosity and uphole velocity of the drilling fluids are the controlling factors in removing cuttings effectively. Because of limitations in pump capacity and therefore effective cuttings removals, most direct rotary machines used to drill water wells are limited to boreholes with a maximum diameter of around 600 mm. This size may not be sufficient for high-capacity wells, especially those that are to be filter packed. Also, as hole-diameters increase past 600 mm, the rate of penetration by direct rotary machines becomes less satisfactory. To overcome the limitation on hole-diameter and drilling rate, reverse circulation machines were designed. The design of a reverse circulation rig is essentially the same as that of the direct rotary rig except most pieces of equipment are larger. For unconsolidated formations, the reverse circulation rig is probably the most rapid drilling equipment available. It requires a large volume of readily available water.

Rotary drillings can also be accomplished with compressed air instead of drilling mud. The technique is rapid and convenient for small-diameter holes in consolidated formations where a clay lining is unnecessary to support the wall against caving in. An important advantage of the air rotary method is its ability to drill through fissured rocks with little or no water.

TABLE 11.10 Well Construction Methods and Applications

Method	Materials	Depth (m)	Diameter (cm)	Yield (m^3/d)
Bored	unconsolidated	10–25	5–90	15–500
Driven	silt, sand	15	3–10	15–200
Jetted	silt, sand	15	4–8	15–150
Percussion	all	450	8–60	15–15,000
Direct rotary	all	450	8–45	15–15,000
Reverse rotary	unconsolidated	60	40–120	2500–20,000
Air rotary	consolidated	250	15–45	2500–15,000

Source: From U.S. Soil Conservation Service, 1969. *Engineering Field Manual for Conservation Practices*. Department of Agriculture, Washington, DC. 995 p. With permission.

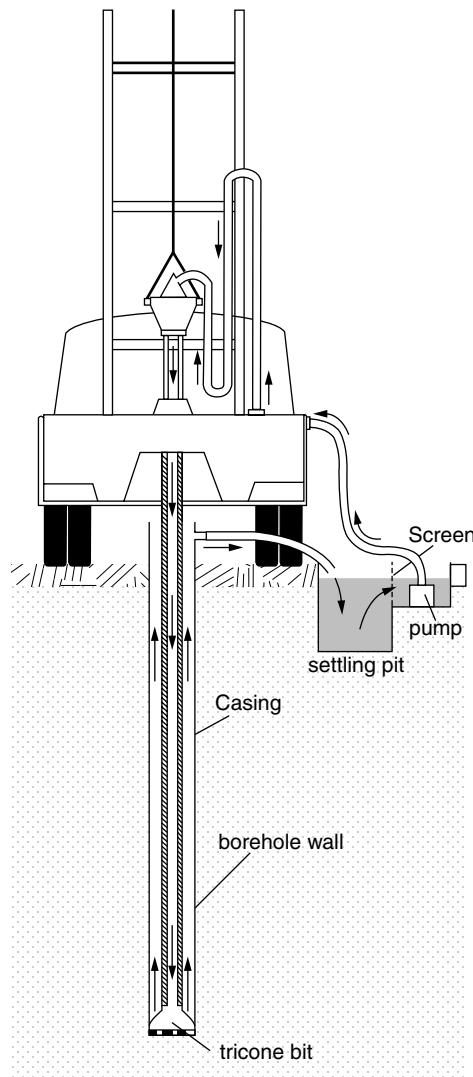


FIGURE 11.13 Schematic diagram of a direct rotary circulation system.

Finally, it should be noted that the depth, diameter, and yield of wells as presented in Table 11.10, are only indications of the order of magnitude of what can be expected. In reality, larger diameters and depth may be obtained, whereas the yield is primarily depending on the actual geology and availability of groundwater.

11.6.2 Well Development

The principal purpose of well development is to remove the fine materials adjacent to the well bore, to increase porosity and hydraulic conductivity of the aquifer and gravel pack, to remove any mud cake or compacted zone as a result from the actual drilling, and to minimize or eliminate sand pumping.

Upon completion of drilling, most wells require development to reach maximum efficiency. This is particularly true of wells producing from unconsolidated aquifer materials and those in which an artificial gravel pack has been placed around the well screens. In addition, many wells may require periodic redevelopment to restore production capacity that has been lost as a result of such factors as encrustation

of screens, clogging of screens by bacterial deposits, or migration of fine particles into a gravel pack. The following discussion summarizes some developments and procedures.

The simplest method of removing finer material from water-bearing formations is by overpumping, that is, pumping at a higher rate than the well will be pumped during exploitation. Overpumping, by itself, seldom produces an efficient well as most of the development action takes place in the most permeable zones closest to the top of the screen. The same applies to a certain extent to surging/backwashing. It consists of pumping a well at a high rate for a short period, shutting down the pump to allow water in the column to fall and backwash the screen, and then repeating the process until the discharge is clear. Although overpumping and backwashing techniques are widely used, and in certain situations may produce reasonable results, their overall effectiveness in high-capacity wells is rather limited when compared with other development methods.

Another method of development is to force water to flow into and out of the screen by operating a plunger up and down the casing, similar to a piston in a cylinder. The tool normally used is called a surge block. The block is usually made of discs of belting bolted between steel plates, or of wood. The outside diameter of the block should be only slightly less than the inside diameter of the screen. Before development, the well should be bailed to make sure that water will flow into it. Surging should start above the screen and move progressively downward to reduce the possibility of sand locking the surge block. The initial motion should be relatively gentle with long strokes and at a slow rate. After surging above the screen, the hole should be cleaned and surging starts at the lower end of the screen, gradually working upward until the entire screen has been developed. The process is then repeated with a faster stroke until little or no sand can be pulled into the well. Surging procedures produce good results for screen installations in zones having good porosity and hydraulic conductivity.

Many drillers use compressed air to develop wells in consolidated and unconsolidated formations. The practice of alternatively surging and pumping with air became feasible with the great number of rotary drilling rigs equipped with large air compressors. In air surging, air is injected into the well to lift the water to the surface. As it reaches the top of the casing, the air supply is shut off, allowing the aerated water column to fall. The effectiveness is often enhanced, particularly with long screens, by using a double-packer tool that confines the pumping to a short length of the screen. Equipment can also be designed to provide a stronger backwash effect in addition to the pumping. As with surge blocks the well should be periodically cleaned.

High-velocity jetting is a very effective way to develop sand and gravel aquifers and artificial gravel packs. In practice, jetting with water is almost always accompanied by simultaneous air-lift pumping so that clogging of the formation does not occur (Figure 11.14). A jetting tool is lowered on a drop pipe and water pumped down at high pressure to produce nozzle velocities of 30 m/sec or more, with a minimum desirable velocity of about 45 m/sec. Driscoll (1986) provides data for nozzles of several sizes at different operating pressures. The tool should have two to four nozzles, spaced equally around the circumference and directed horizontally toward the inside of the well screen. The nozzles should extend to within 25 mm from the inside of the screen. In operation, the jetting tool is lowered into the lower part of the screen, rotated slowly, and raised slowly throughout the length to be developed. The process is then repeated until development is complete.

Various acids and chemicals can be very effective in well development under certain conditions, both upon completion of drilling and for maintenance. Acid and chemicals can be used alone, or in combination with other methods such as surge blocks or water jetting. Acids can be very effective in limestone and dolomite aquifers by enlarging fractures adjacent to the well bore and thereby significantly reducing drawdown at a given pumping rate. Acids can also be effective in removing encrusting deposits on well screens (see section on maintenance). Acids frequently used in well development and rehabilitation are hydrochloric acid and sulfamic acids. The usual procedure is to introduce a volume of acid several times that of the section of screen to be treated, followed by frequent agitation, and then removed by pumping after 6 to 8 h. Chemicals frequently used in well development are polyphosphates and surfactants (wetting agents). The polyphosphates act to deflocculate and disperse clays and drilling muds. They are typically introduced as a solution of from 0.5 to 1.5 kg polyphosphate in 100 l of water in conjunction with surge

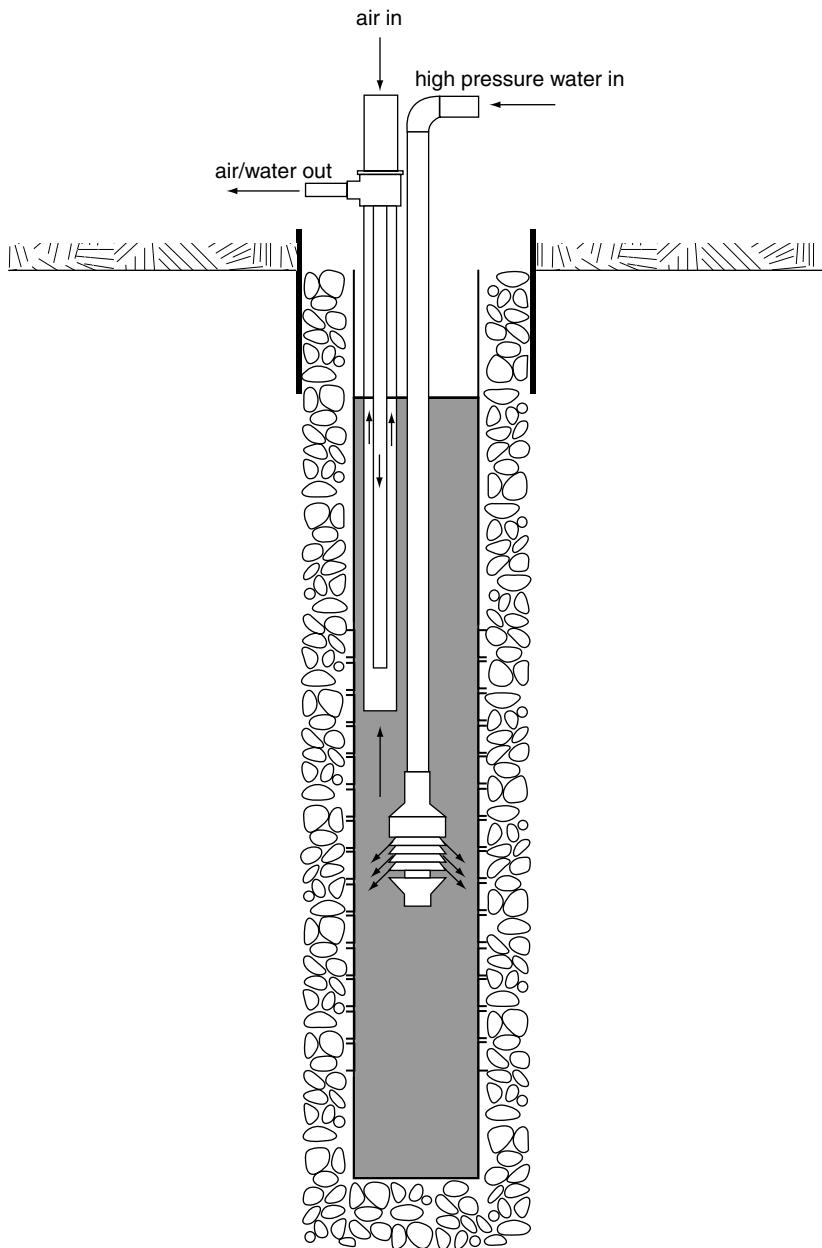


FIGURE 11.14 Schematic diagram of high-velocity jetting with air-lift pumping system.

block or water jetting development. The addition of wetting agents aids in allowing the polyphosphates to penetrate clays.

11.6.3 Maintenance

The performance of a well usually declines after some years of operation, resulting in higher drawdowns and higher pumping costs. The well is in need of rehabilitation when the specific capacity of the well becomes so small that the pumping costs increase or the discharge rate of the well can no longer be

maintained. Before that time, the well needs to be rehabilitated. An effective well-maintenance program begins with good records being kept of the well's construction, including good records of the geological conditions, the position and types of aquifers and aquiclude, water quality, and the specific capacity of the well, determined during well testing.

Every type of well requires its own maintenance program. To evaluate the performance of a well, Driscoll (1986) provides a checklist. A significant decrease in specific capacity or an increase in the pumping of sand indicates that the well needs rehabilitation or restoration to its original performance. In general, rehabilitation measures are most successful when the well performance has not deteriorated too badly, or the specific capacity has not decreased too much. If the specific capacity of the well has declined by 25% of its original value, it is time to carry out a rehabilitation program. To determine the right moment for well rehabilitation, periodic monitoring of well performance should be done in term of short standard tests. Complete well records can be kept at relatively minor expense, and these are indispensable in determining the causes of well failure and selecting the maintenance and rehabilitation program.

The major causes of a reduction in well performance are: (1) a reduced well yield due to chemical encrustation or clogging of the screen due to bacteriological activity; (2) plugging of the formation around the well screen by fine particles of clay and sand in the pores; (3) pumping of sand due to poor well design or corrosion of the well screen; and (4) collapse of the well screen due to chemical or electrolyte corrosion of metal well screens.

Chemical and biological encrustations are major causes of well failure. Water quality and flow velocity through the screen openings determine the occurrence of encrustation. Chemical encrustation is caused by the precipitation of carbonates, mainly calcium carbonate, or iron hydroxides, which block the screen openings. Carbonate precipitation is caused by the release of carbon dioxide from the water owing to a pressure decline in the water caused by the drawdown in the well. Iron dissolved in groundwater may precipitate from the water on the well screen as oxygen is introduced into the water when the well is pumped. Another reason for the precipitation of iron may be the presence of iron bacteria in the water. Chemical encrustation can best be removed by treating the well with a strong acid solution that chemically dissolves the encrusting materials so that they can be removed from the well by pumping. Hydrochloric acid and sulphuric acid can be used. Chlorine, a strong oxidizing agent, inhibits the growth of iron bacteria. The use of hypochlorite is a relatively safe and convenient alternative to chlorine gas. The occurrence of iron bacteria in wells can be prevented by disinfecting the well and the pump immediately after installation.

Physical plugging by clay and silt particles can best be prevented by proper well development after the well screen has been installed. The removal of fine particles from the formation immediately around the screen can best be achieved by washing and brushing the screen with dispersing compounds such as sodium tripolyphosphate and other types of polyphosphates.

Sand pumping causes the abrasion of pump bowls that leads to the failure of the pump. Sand pumping results from over-sized slots in screens, over-sized gravel pack, corrosion of the well screen, inadequate development of the well, or too high entrance velocities, causing the transport of sand from the formation toward the well. One of the above conditions, or a combination of them, results in sand from the formation entering the well. Remedy this problem may be uneconomic: it may be better to drill a new well. The best alternative, if possible, is to replace the screen or to place an inner screen inside the original well screen.

Corrosion of well screens can severely reduce the lifetime of a well. Chemical corrosion occurs especially when metal well screens are used in aggressive and saline water loaded with gasses like hydrogen sulphide, carbon dioxide, and oxygen. Corrosion can be prevented by applying nonmetal screens, or, when the water is not aggressive, only metal screens of stainless steel and low-carbon steel.

Finally, to pump water from a well in the most economic way, proper maintenance of pumps and engines is a prerequisite. Pump and engine manufacturers prescribe periodical maintenance of their products. Maintenance procedures depend on the pump type. They include the adjustment and replacement of impellers, bearings, stuffing boxes, and bowl assemblies. A complete analysis of pump and engine maintenance is beyond the scope of this chapter, so readers are referred to the maintenance procedures specified by manufacturers.

Glossary

- Drainable surplus:** the quantity of water that flows into a groundwater reservoir in excess of the quantity that flows out under natural conditions.
- Impeller:** devices that accelerate the water within the pump to build pressure.
- Specific capacity:** the yield of a well per unit drawdown, usually expressed as cubic meters per day per meter drawdown. It is a measure for the productivity of a well.
- Uniformity coefficient:** ratio of the 40% retained particle size of the sediment and the 90% retained size. The lower its value, the more uniform is the grading between these limits.

References

- Anderson, K.E. 1995. *Ground Water Handbook*. National Ground Water, Dublin, OH, 401 p.
- Boehmer, W.K. and Boonstra, J. 1994. Tubewell drainage systems. *Drainage Principles and Applications*. 2nd ed. H.P. Ritzema (ed.). ILRI-Publication 16, Wageningen, pp. 931–964.
- Boonstra, J. and Kselik, R.A.L. 2001. *SATEM 2002: Software for Aquifer Test Evaluation*. ILRI Publication 57, Wageningen, 148 p.
- Carle, S.F. and Fogg, G.E. 1996. Transition Probability-Based Indicator Geostatistics. *Mathematical Geology*, Vol. 28(4), pp. 453–476.
- Driscoll, F.G. 1986. *Groundwater and Wells*. St. Paul, Johnson Division, Minnesota, 1089 p.
- Hantush, M.S. 1964. Hydraulics of wells. *Advances in Hydroscience*, Vol. 1, pp. 281–432.
- Huisman, L. 1975. *Groundwater Recovery*. 2nd ed. McMillan, London, 336 p.
- Jacob, C.E. 1947. Drawdown test to determine effective radius of artesian well. *Transactions of the American Society of Civil Engineers*, Vol. 112, Paper 2321, pp. 1047–1064.
- Lennox, D.H. 1966. Analysis of step-drawdown test. *Journal of the Hydraulics Division, Proceedings of the American Society of Civil Engineers*, Vol. 92(HY6), pp. 25–48.
- Peterson, D.F., Israelson, O.W. and Hansen, V.E. 1952. *Hydraulics of Wells*. Bulletin 351. Utah Agricultural Experimental Station, 48 p.
- Rorabaugh, M.J. 1953. Graphical and theoretical analysis of step-drawdown test of artesian well. *Transactions of the American Society of Civil Engineers*, Vol. 79, separate no. 362, 23 pp.
- Sheahan, N.T. 1971. Type-curve solution of step-drawdown test. *Ground Water*, Vol. 9, pp. 25–29.
- Steenbergen, F. van 2003. Local groundwater regulation. Water Praxis Document Nr. 14. ARCADIS-Euroconsult, Arnhem, The Netherlands. 25 p.
- U.S. Environmental Protection Agency. 1975. *Manual of Water Well Construction Practices*. EPA-570/9-75-001. Office of Water Supply, Washington, DC. 156 p.
- U.S. Soil Conservation Service. 1969. *Engineering Field Manual for Conservation Practices*. Department of Agriculture, Washington, DC. 995 p.

Further Information

- Driscoll (1986) provides a wealth of practical information on well hydraulics, well drilling, well design, well pumps, well maintenance and rehabilitation, and groundwater monitoring that cover a wider range of conditions than here presented.
- Boonstra and Kselik (2001) presents a software package that facilitates the analysis of step-drawdown data, including both the Jacob's method as well as the Rorabaugh's method.

12

Sea Water Intrusion into Coastal Aquifers

12.1	Introduction.....	12-1
12.2	Mathematical Modeling of Sea Water Intrusion	12-6
12.3	Numerical Modeling and Computer Code	12-11
	The Numerical Model • The FEAS Code	
12.4	Examples of Sea Water Intrusion	12-15
	Example A: Sea Water Intrusion — A Typical Case • Example	
	B: Sea Water Intrusion with Upconing due to Pumping •	
	Example C: Sea Water Intrusion in a Double-Layered Aquifer	
12.5	Introduction to the Management of a Coastal Aquifer	12-22
	Glossary.....	12-26
	Acknowledgement	12-27
	References	12-27
	Further Information	12-29

Jacob Bear
*Techion-Israel Institute of
Technology*

Quanlin Zhou
*Lawrence Berkeley National
Laboratory*

The objective of this chapter is to present the problem of sea water intrusion into a coastal aquifer, together with the conceptual and complete mathematical models that describe it. Examples are presented of how some typical sea water intrusion cases that are described by the models are solved by the computer code FEAS. A brief discussion is presented on the management of a coastal aquifer.

12.1 Introduction

Coastal aquifers constitute important sources of fresh water in many parts of the world, especially in arid and semiarid zones. Often, coastal areas are also heavily urbanized, a fact that makes the need for fresh water in such areas even more acute. However, the proximity and contact with the sea require special attention when planning the management of such aquifers. In fact, in many coastal aquifers, the intrusion of sea water has become one of the major constraints on groundwater management. As sea water intrusion progresses, the part of the aquifer close to the sea becomes saline, and pumping wells close to the coast start pumping saline water and have to be abandoned. Also, the area above the intruding sea water wedge is, usually, lost as a source of fresh water (by natural replenishment).

Since the famous works of Badon-Ghyben (1888) and Herzberg (1901), and the less known work of Du Commun (1828) (Konikow and Reilly, 1999), extensive research has been carried out, leading to the understanding of the mechanisms that govern sea water intrusion. The dominant factors are (1) the flow regime in the aquifer above the intruding sea water wedge, (2) the variable density, and (3) hydrodynamic

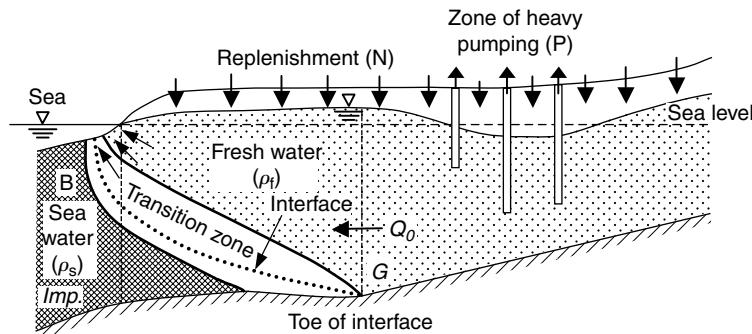


FIGURE 12.1 A typical cross section of sea water intrusion in a coastal aquifer. (After Bear, 1979, *Hydraulics of Groundwater*. McGraw-Hill, New York, p. 569. With permission.)

dispersion. Reviews of the phenomenon of sea water intrusion, and of the research that has been carried out on this subject, both theoretical work and field and laboratory investigations, may be found in many books and publications, and will not be repeated here (e.g., Bear, 1972, 1979; Reilly and Goodman, 1985; Bear and Verruijt, 1987; Bear et al., 1999; Cheng and Ouazar, 2004).

Briefly, under normal conditions in a coastal aquifer, excess fresh water (i.e., natural and artificial recharge minus pumping) is discharged into the sea. This means that close to the coast, a seaward hydraulic gradient exists in the aquifer. Due to the presence of sea water in the aquifer under the sea, a zone of contact is formed between the lighter fresh water flowing towards the sea, and the heavier sea water in the aquifer. A typical cross section, with natural replenishment, pumping, and a transition zone, is shown in Figure 12.1. The detailed shape of the transition zone, from fresh water to sea water, depends also on whether this zone is advancing inland or retreating. In all cases, the domain in the aquifer that is occupied by sea water has, usually, the form of an advancing or receding wedge. One should note that, like all figures that describe aquifers, these are also highly distorted figures, not drawn to scale.

Sea water and fresh water are often referred to as “miscible liquids,” although, actually, both constitute a *single* liquid phase — water, H_2O — with different concentrations of *total dissolved matter* (salt, TDM). For the sake of simplicity, we shall continue to refer to them as two liquids — fresh water and sea water. Hence, the passage from the portion of the aquifer that is occupied by the former to that occupied by the latter takes the form of a *transition zone*, rather than a *sharp interface*. Under certain circumstances, depending on the extent of sea water intrusion and on certain aquifer properties, this transition zone, which is primarily a result of *hydrodynamic dispersion* of the dissolved matter, may be rather wide. Under other conditions, it may be rather narrow, relative to the aquifer’s thickness, and the passage from the zone occupied by fresh water to that occupied by sea water may be *approximated* as a *sharp interface*. Often, the term “interface” is used for the iso-density surface that is midway between fresh water and sea water. In this article, the term “interface” will, sometimes, be used interchangeably with “transition zone.”

Under natural undisturbed conditions in a coastal aquifer, a state of equilibrium is maintained, with a steady-state sea water zone and a zone of flowing fresh water above it. The transition zone is also fed from below by sea water (Figure 12.1). When the discharge to the sea is reduced as a result of pumping fresh water from a coastal aquifer, water levels (or the piezometric heads in a confined aquifer) close to the sea are lowered and the transition zone rises. The sea water wedge and the transition zone advance landward, until a new equilibrium is reached. Wells that operate within the sea water wedge and the transition zone pump saline water and have to be abandoned. When pumping takes place in a well located above the transition zone, the latter *upcones* towards the well. Unless the well is at a sufficient distance above this zone and the rate of pumping is sufficiently small, the well will eventually pump saline water.

We should emphasize here that the geological structure and configuration of a coastal aquifer are usually much more complex than the schematic one shown in Figure 12.1. Often, especially close to the sea, the coastal aquifer is divided into a number of sub-aquifers by impervious or semi-pervious layers,

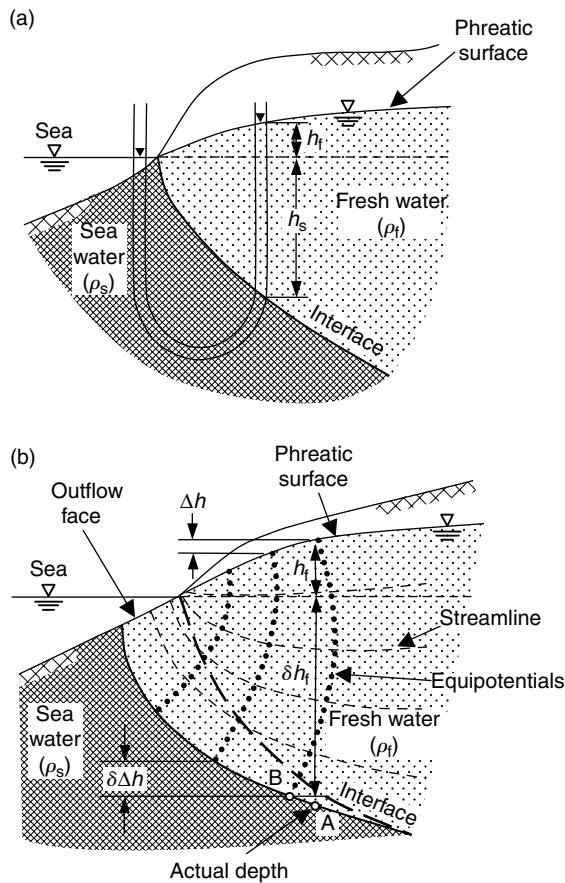


FIGURE 12.2 (a) The Ghyben–Herzberg approximation, and (b) the actual sharp interface.

with leakage taking place through the latter. The flow in such multilayered aquifers and the shape of the transition zone(s) are often much more complex than shown in Figure 12.1. An example of sea water intrusion in a multilayered coastal aquifer is demonstrated in Section 12.4.

As emphasized so far, a transition zone always exists between fresh water and sea water. Moreover, nowadays, we also have enough tools to handle three-dimensional (3D) transition zone models. Nevertheless, it is useful to use the “sharp interface approximation” of this zone to discuss certain basic features of sea water intrusion. Here, we shall present the concept of a sharp interface to introduce the relationship between seaward flow of fresh water and the extent of sea water intrusion. We shall do so by a simple steady-state analysis.

We start by introducing the famous Ghyben–Herzberg rule that describes the relationship between the fresh water flow to the sea and the extent of sea water intrusion, say, as expressed in terms of the length of the sea water wedge. Figure 12.2a shows a vertical cross section normal to the coast, with a sharp interface separating the two fluids as envisaged by the *Ghyben–Herzberg assumption*. Essentially, Ghyben and Herzberg assumed that under steady-state conditions, a static equilibrium exists, with stationary sea water and a hydrostatic pressure distribution in the seaward-flowing fresh water. This means that the flow is (essentially) horizontal and the equipotentials (i.e., surfaces of equal piezometric head) are vertical. This, in fact, is identical to the Dupuit assumption of essentially horizontal flow in aquifers. With the notation of Figure 12.2a, we have

$$h_s = \delta h_f \quad (12.1)$$

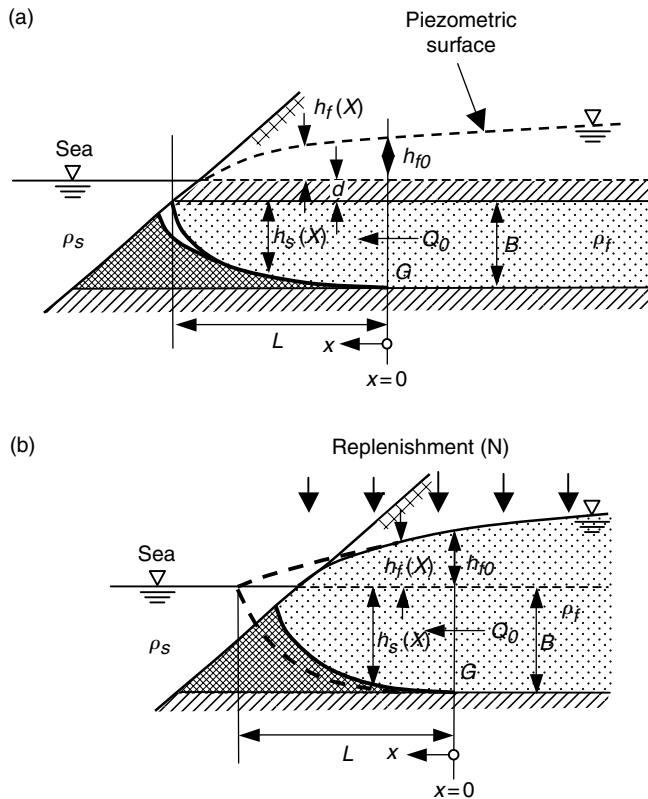


FIGURE 12.3 The interface (a) in a confined, and (b) in a phreatic coastal aquifer.

where

$$h_f = z + \frac{p}{\rho_f g}, \quad h_s = z + \frac{p}{\rho_s g}, \quad \delta = \frac{\rho_f}{\rho_s - \rho_f} \quad (12.2)$$

in which p is the pressure, h_f and h_s denote the fresh water and sea water piezometric heads, in the fresh water zone and in the sea water wedge, respectively; ρ_f and ρ_s denote the respective densities, assumed constant, z denotes the vertical coordinate (positive upward), and g denotes gravity acceleration. Note the discrepancy from the Ghysen–Herzberg interface shape as the interface approaches the coast. The discrepancy stems from the difference between the actual shape of the equipotentials and the vertical ones assumed by the Dupuit approximation. In a confined aquifer, h_s in Equation 12.1, is the depth of a point on the interface below sea level, while h_f is the fresh water piezometric head.

We now consider the simple case of essentially horizontal steady flow normal to the coast in a confined aquifer with the stationary interface shown in Figure 12.3a. For this case, the fresh water flow equation, a combination of the steady-state mass balance equation and the flux equation, is written:

$$\frac{\partial Q_f}{\partial x} \equiv \frac{\partial}{\partial x} \left(-K_f h_s \frac{\partial h_f}{\partial x} \right) = 0 \quad (12.3)$$

where Q_f is the discharge rate of fresh water to the sea, per unit length of coast, K_f is the hydraulic conductivity in fresh water ($=(\rho_f g k / \mu_f)$) k is the permeability and μ_f is the fresh water dynamic viscosity).

With the relationship between h_f and h_s :

$$h_s + d = \delta h_f \quad (12.4)$$

and the boundary conditions:

$$Q_f \equiv Q_0, \quad h_f = h_{f0}, \quad \text{and} \quad h_s = B, \quad \text{at } x = 0 \quad (12.5a)$$

$$h_s = 0, \quad \text{at } x = L \quad (12.5b)$$

we have the solution:

$$Q_0 x = \frac{K_f (B^2 - h_s^2)}{2\delta}, \quad Q_0 L = \frac{K_f B^2}{2\delta} \quad (12.6)$$

Equation 12.6, which describes a *parabola*, clearly shows the relationship between the length of sea water intrusion, L , and the discharge to the sea, Q_0 (as well as the piezometric head, h_{f0} , above the toe of the interface (G)).

For the phreatic aquifer shown in Figure 12.3b, the flow equation, under the same conditions, takes the form:

$$-\frac{\partial Q_f}{\partial x} + N = 0, \quad Q_f = -K_f(h_f + h_s) \frac{\partial h_f}{\partial x} \quad (12.7)$$

where N is the replenishment. With Equation 12.1 and the following boundary conditions:

$$Q_f = Q_0, \quad h_f = h_{f0}, \quad h_s = h_{s0} = \delta h_{f0}, \quad \text{at } x = 0 \quad (12.8a)$$

$$h_s = h_f = 0, \quad \text{at } x = L \quad (12.8b)$$

we have the solution:

$$h_{f0}^2 - h_f^2 = \frac{2Q_0 x + Nx^2}{K_f(1+\delta)} \quad (12.9a)$$

$$h_{f0}^2 = \frac{2Q_0 L + NL^2}{K_f(1+\delta)} \quad (12.9b)$$

$$Q_0 = \frac{K_f B^2}{2L} \frac{1+\delta}{\delta^2} - \frac{NL}{2} \quad (12.9c)$$

Again, we see here the relationship between the drainage of fresh water to the sea and the length of sea water intrusion. Reducing Q_0 means an increased L . By controlling h_{f0} at the toe G, say, by means of artificial recharge, the water table may be lowered landward, without any additional sea water intrusion.

Altogether, also in the general case of 3D flow with a transition zone, the extent of sea water intrusion, measured, say, by the length of the intruding wedge on the aquifer's bottom, is related to the rate of fresh water drained to the sea. As this rate is reduced (by pumping), the extent of sea water intrusion increases. This makes the problem of sea water intrusion into a coastal aquifer a *management problem*. In such a problem, we seek, for example, to determine the maximum annual net rate of pumping from the aquifer (i.e., pumping minus artificial recharge) such that a certain goal is achieved (e.g., maximize net benefits), without violating specified (legal, hydrologic, technical, etc.) constraints. Thus, in managing a coastal aquifer, the issue is not to stem sea water intrusion, but to determine the optimal extent of sea water intrusion. Usually, we wish to maximize pumping in order to achieve certain, say, economic, goals, without violating specified constraints. Models of flow, solute transport, and optimization are used in

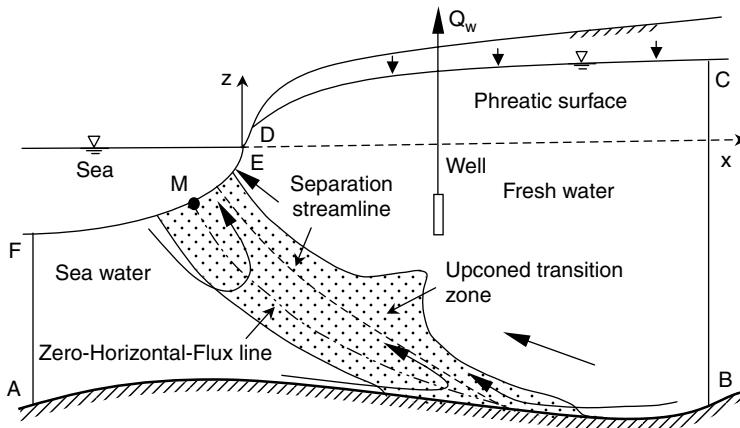


FIGURE 12.4 The transition zone with upconing in a simplified vertical cross section of a coastal aquifer, normal to the coast line.

planning aquifer management. An essential constraint to be satisfied, as part of the optimization model, is the mathematical flow and solute transport model, which takes into account the dependence of water density on salt concentration (= variable density flow).

In the past, the *sharp interface approximation* was introduced, primarily, to enable relatively easy solutions, both analytical and numerical, of certain simple sea water intrusion problems of practical interest. However, nowadays, with the availability of new improved numerical techniques, including methods for coping with the nonlinearities that are inherent in the transition zone model, and with fast and large memory computers (even PCs), numerical solutions of 3D models that take the transition zone into account should not pose special difficulties. There is also no reason to limit the models to (vertical) two-dimensional (2D) flow domains. Indeed, a number of models and computer codes that consider sea water intrusion as a solute transport problem have already been developed (e.g., Sorek and Pinder, 1999). In Section 12.4, we present a number of typical sea water intrusion cases solved by the computer code FEAS (Zhou, 1999; Bensabat et al., 2000; Bear et al., 2001; Zhou et al., 2001, 2005).

Plenty of information on sea water intrusion into coastal aquifers (e.g., the management problem, modeling with the sharp interface approximation, analytical solutions, modeling as a variable density flow and transport problem, numerical solutions, and discussions on specific numerical codes) can be found in Bear et al. (1999). A few comments on the management of a coastal aquifer will be presented in Section 12.5.

12.2 Mathematical Modeling of Sea Water Intrusion

As already mentioned in the introduction, *in reality*, the aquifer domain occupied by only sea water, and the aquifer domain occupied by only fresh water are separated by a *transition zone*. This is a consequence of the fact that the two “miscible” liquids are, actually, single liquid — water — with different concentrations of dissolved salts. The width of the transition zone is dictated by three phenomena (a) advection of fresh water towards the sea (or, under certain conditions, landward), (b) recirculation of sea water and mixed water, and (c) hydrodynamic dispersion (dispersion and molecular diffusion) in the transition zone. Across it, the salinity of the water varies from that of fresh water to that of sea water. The width of this zone grows as it is being displaced in response to changes in the flow regime and in the discharge of water to the sea. The transition zone is also fed by a flux of salt from the sea water zone.

Figure 12.4 shows a phreatic coastal aquifer with a transition zone between sea water and fresh water. The considered flow domain is ABCDEMFA. The detailed conceptual and mathematical models of flow and solute transport are presented in Chapters 4, 18, and 19 of this Handbook and will not be repeated

here. The new features here are that the flow and the solute transport models are *coupled* as the density of the liquid continuously varying in response to the changes in dissolved salt concentration, and that here we apply these models to the particular case of sea water intrusion in a coastal aquifer. We usually refer to such a model as “variable density flow and transport model.” In what follows, we shall present this model, assuming isothermal conditions.

The mathematical model describing sea water intrusion in a coastal aquifer consists of (a) mass balance equation for the water, (b) flux equation for the water of variable density (Darcy's law), (c) mass balance equation for the dissolved salts, and (d) flux equation for the dissolved salts. The first two equations are often combined as a single *flow equation* for the water. The last two equations may be combined to form a single mass balance equation for the dissolved salts, often referred to as the *advection-dispersion equation*, or *the transport equation*. In addition, the complete model includes (a) constitutive equations that relate the liquid's density and dynamic viscosity to the total dissolved salt concentration, and (b) initial and boundary conditions (Bear, 1999; Zhou, 1999; Zhou et al., 2005).

Although, in principle, the flow model can be written in terms of pressure, p , as a primary variable, a more (numerically) efficient model is obtained by expressing the flow model in terms of a *reference (fresh water) piezometric head*, $h_f(x, y, z, t)$, as defined in Equation 12.2, using ρ_f as a reference density (Frind, 1982). To generalize the discussion on variable density flow, we consider the Total Dissolved Solids (TDS) in the range where the volume of water may vary with concentration. To facilitate the discussion, we introduce a *normalized salt mass fraction*, $C(x, y, z, t)$, defined by

$$C = \frac{\omega - \omega_f}{\omega_s - \omega_f} \quad (12.10)$$

in which ω and ω_s denote the salt (or TDS) mass fraction (i.e., mass of the dissolved salt per unit mass of fluid) in water (=salt solution) and in sea water, respectively, and ω_f is the fresh water mass fraction, used as a reference mass fraction. Note that $c = \rho\omega$, in which c denotes the TDS concentration (=mass of TDS per unit volume of fluid).

The constitutive equation that expresses the relationship, $\rho = \rho(p, \omega)$, between the fluid density, the pressure, and the salt mass fraction, is

$$\rho = \rho_0 \exp[\beta'_p(p - p_0) + \beta'_\omega(\omega - \omega_0)] \quad (12.11)$$

where ρ_0 , p_0 , and ω_0 are reference values of density, pressure, and salt mass fraction, respectively, $\beta'_p = (1/\rho)\partial\rho/\partial p$ is the coefficient of water compressibility (at constant salt mass fraction), and $\beta'_\omega = (1/\rho)\partial\rho/\partial\omega$ is a coefficient that introduces the effect of change in salt mass fraction on the fluid's density (at constant pressure). Usually, we select $\omega_0 = \omega_f$ (i.e., equal to the mass fraction of fresh water). The reference pressure is such that for fresh water at $p = p_0$, the density is $\rho = \rho_f$, as the density of fresh water is selected as *a reference density*.

The linearized approximation of Equation 12.11 is

$$\rho = \rho_0[1 + \beta''_p(p - p_0) + \beta''_\omega(\omega - \omega_0)] \quad (12.12)$$

where $\beta''_p = (1/\rho_0)\partial\rho/\partial p$ and $\beta''_\omega = (1/\rho_0)\partial\rho/\partial\omega$. In what follows, we shall assume that for the range of pressures considered here, $\beta''_\omega|\Delta\omega| \gg \beta''_p|\Delta p|$, so that we may employ the approximation:

$$\rho = \rho_f(1 + \beta_c C), \quad \beta_c = \beta''_\omega(\omega_s - \omega_f) \quad (12.13)$$

where β_c may be referred to as a *density difference factor* (dimensionless). In spite of the above assumption, we do take into account the effect of pressure in the expression for the specific storativity appearing in the mass balance equation.

To introduce the effect of mass fraction on the fluid's dynamic viscosity, we may use the constitutive relationship for dynamic viscosity in the form (Lever and Jackson, 1985):

$$\mu = \mu_f \mu_r = \mu_f (1 + 1.85\omega - 4.1\omega^2 + 44.50\omega^3) \quad (12.14)$$

in which the viscosity, μ_f , corresponds to $\omega = 0$, and μ_r is defined by Equation 12.14.

The *specific discharge relative to the solid*, $q_r \equiv \phi(\mathbf{V} - \mathbf{V}_s) = \mathbf{q} - \phi\mathbf{V}_s$, is expressed by Darcy's law (Bear, 1972; Bear and Bachmat, 1990)

$$\mathbf{q}_r = \phi(\mathbf{V} - \mathbf{V}_s) = -\frac{\mathbf{k}}{\mu}(\nabla p + \rho g \nabla z) \quad (12.15)$$

in which \mathbf{V} and \mathbf{V}_s denote the velocity vectors of the water and of the solid matrix, respectively, ϕ denotes the porosity, the second rank tensor, \mathbf{k} , denotes the permeability, and \mathbf{q} denotes the specific discharge vector. We usually assume, also here, that $\mathbf{V}_s \approx 0$, $\mathbf{q}_r \approx \mathbf{q}$. In terms of h_f and C , Darcy's law can be rewritten in the form:

$$\mathbf{q} = -\frac{\mathbf{K}_f}{\mu_r}(\nabla h_f + \beta_c C \nabla z), \quad \mathbf{K}_f = \frac{\rho_f g \mathbf{k}}{\mu_f} \quad (12.16)$$

with \mathbf{K}_f denoting the reference hydraulic conductivity.

For a single fluid phase, for example, water, of variable density, Bear (1972, 1979) and Bear and Bachmat (1990), as many others, present the general mass balance equations in the form:

$$\frac{\partial \phi \rho}{\partial t} = -\nabla \cdot (\rho \mathbf{q} - \phi \mathbf{D} \nabla \rho) + (\rho_R Q_R - \rho Q_P) \quad (12.17)$$

in which Q_R and Q_P , denote, symbolically, the rates of injection of water into and withdrawal of water from the aquifer (dimension 1/T), ρ_R denotes the density of the injected water, t is the time, and \mathbf{D} , a second rank symmetric tensor, denotes the *coefficient of dispersion* (dimensions L²/T). When water is withdrawn and injected through (point) wells, the symbolic source and sink terms in Equation 12.17 may, symbolically, be written as:

$$\rho_R Q_R = \sum_n \rho_{Rn} Q_{Rn}(\mathbf{x}_n, t) \delta(\mathbf{x} - \mathbf{x}_n) \quad (12.18a)$$

$$\rho Q_P = \sum_m \rho_m Q_{Pm}(\mathbf{x}_m, t) \delta(\mathbf{x} - \mathbf{x}_m) \quad (12.18b)$$

in which Q_{Rn} and ρ_{Rn} are the injection rate, and the density of the water injected through a well at point \mathbf{x}_n , respectively, and Q_{Pm} and ρ_m are the pumping rate and density of the water pumped through a well at point \mathbf{x}_m , and δ denotes the *Kronecker delta*.

Note that Equation 12.17 includes a term, $\mathbf{J}^{*\rho}$, denoting the *dispersive flux of the total fluid mass*, which can be rewritten in the form:

$$\mathbf{J}^{*\rho} = -\mathbf{D} \cdot \nabla \rho = -\rho_f \beta_c \mathbf{D} \cdot \nabla C \quad (12.19)$$

In the case of a fluid of variable density, significant density gradients may develop. For example, it is possible that a rather narrow transition zone will develop, with a relatively large density gradient across it, therefore with flow that is more or less normal to such a gradient, significant lateral dispersion may take place. Bear and Bachmat (1990, p. 290) present and discuss a method, using appropriate Peclet numbers (that define the ratio between advective and dispersive fluxes), for examining the conditions under which the dispersive flux of the total mass may be neglected as being much smaller than the advective one. Note

that *there is no diffusive flux of the total dissolved mass*. In the case of sea water intrusion, we may encounter a rather narrow transition zone, with flow parallel to it. The appropriate Peclet number may be less than or not much larger than one, so that we cannot conclude, *a-priori*, that advection dominates over dispersion.

In view of the relationship $\rho = \rho(p, C)$, we make use of Equation 12.16 and Equation 12.19 to modify Equation 12.17, rewriting it in terms of the reference piezometric head, h_f , and the normalized mass fraction, C , in the form:

$$S_0 \frac{\partial h_f}{\partial t} + \phi \beta_c \frac{\rho}{\rho_f} \frac{\partial C}{\partial t} = \nabla \cdot \left[\frac{(1 + \beta_c C)}{\mu_r} \mathbf{K}_f \cdot (\nabla h_f + \beta_c C \nabla z) + \phi \beta_c \mathbf{D} \cdot \nabla C \right] + \frac{\rho_R}{\rho_f} Q_R - (1 + \beta_c C) Q_P \quad (12.20)$$

in which $S_0 = \rho g(\phi \beta_p + \alpha)$ denoting the *aquifer's elastic storativity*, with α denotes the aquifer's (vertical) compressibility. Other forms are also possible.

For a fluid of variable density, the general mass balance equation for the TDS in the water can be written in the form:

$$\frac{\partial \phi \rho \omega}{\partial t} = -\nabla \cdot (\rho \omega \mathbf{q} - \phi \rho \mathbf{D}_h \cdot \nabla \omega) + \rho_R \omega_R Q_R - \rho \omega Q_P \quad (12.21)$$

in which \mathbf{D}_h denotes the *coefficient of hydrodynamic dispersion* (a second rank symmetric tensor), $\mathbf{D}_h = \mathbf{D} + \mathbf{D}_{\text{diff}}^*$, and $\mathbf{D}_{\text{diff}}^*$ denotes the *coefficient of molecular diffusion in a porous medium*. In an *isotropic* porous medium, the *coefficient of dispersion*, D , is related to the fluid's velocity and to the *longitudinal and transversal dispersivities*, α_L and α_T (dimension L), respectively, by

$$D_{ij} = \alpha_T V \delta_{ij} + (\alpha_L - \alpha_T) \frac{V_i V_j}{V} \quad (12.22)$$

in which δ_{ij} denotes the *Kronecker delta*, and $V = |V|$.

The salt transport equation, Equation 12.21, can be written in terms of the normalized mass fraction, C , with $\omega_f = 0$ (Bear, 1999; Zhou et al., 2005):

$$\frac{\partial \phi \rho C}{\partial t} = -\nabla \cdot (\rho C \mathbf{q} - \phi \rho \mathbf{D}_h \cdot \nabla C) + \rho_R C_R Q_R - \rho C Q_P \quad (12.23)$$

Expanding the first-order derivatives in Equation 12.23 and making use of the mass balance equation for water, Equation 12.17, we can rewrite the transport equation in the form:

$$\phi \rho \frac{DC}{Dt} = \nabla \cdot (\phi \rho \mathbf{D}_h \cdot \nabla C) + \rho_R Q_R (C_R - C) \quad (12.24)$$

in which DC/Dt is the material derivative with respect to solute particles. Altogether, we have to solve the variable density flow equation, Equation 12.20, and the salt transport equation, Equation 12.24, simultaneously, for the two primary variables: $h_f(x, y, z, t)$ and $C(x, y, z, t)$, making use of Equations 12.13 and 12.14 to express ρ and μ , and Equation 12.16 to express \mathbf{V} or \mathbf{q} .

The initial conditions for the variable density flow and salt transport equations are

$$h_f(x, y, z, 0) = h_{f0}(x, y, z), \quad C(x, y, z, 0) = C_0(x, y, z) \quad (12.25)$$

where $h_{f0}(x, y, z)$ and $C_0(x, y, z)$ are known distributions.

The conditions on the boundary segments (shown in Figure 12.4) for the flow equation, Equation 12.20, neglecting the dispersive flux of the total fluid mass, are:

$$\begin{aligned}
 \text{On AB : } q_n &\equiv \mathbf{q} \cdot \mathbf{n} = 0, & & \text{(impervious bottom)} \\
 \text{On BC : } h_f &= h_{fp}, \text{ or } \mathbf{q} \cdot \mathbf{n} = q_{np}, & & \text{(land-side lateral boundary)} \\
 \text{On CD : } \rho q_n &= -\rho_N N n_z + \frac{(\phi\rho - \theta_{w0}\rho_N)}{|\nabla F|} \frac{\partial h_f}{\partial t}, & & \text{(phreatic surface)} \\
 \text{On DE : } h_f &= \varsigma \gamma_{DE}, & & \text{(seepage face)} \\
 \text{On EF : } h_f &= \beta_c H_{sea}, & & \text{(sea bottom)} \\
 \text{On FA : } h_f &= \beta_c H_{FA}, & & \text{(sea-side lateral boundary)} \tag{12.26}
 \end{aligned}$$

where h_{fp} is the prescribed reference head on a Dirichlet-type boundary, q_{np} is the prescribed fluid flux on a 2nd-type boundary, \mathbf{n} ($= \nabla F/|\nabla F|$, with $F(x, y, z, t) = 0$ representing the equation of a boundary surface) is the unit outward normal vector, with the components n_x, n_y, n_z in 3D, N is the rate of replenishment, ρ_N is the density of the replenishment, θ_{w0} is the irreducible water content assumed to prevail in the unsaturated zone, ς_{DE} is the elevation at a point on the seepage face above sea level, H_{sea} and H_{FA} are the depth of sea water at a point on the sea bottom and the sea-side lateral boundaries, respectively. If we wish to take into account the dispersive flux of the fluid's mass, we replace $\rho\mathbf{q}$ by $\rho\mathbf{q} - \phi\rho\beta_c\mathbf{D} \cdot \nabla C$.

The boundary conditions for the salt transport equation are:

$$\begin{aligned}
 q_n^d &\equiv -\mathbf{D}_h \cdot \nabla C = 0, & & \text{On AB and DE} \\
 C &= 0, & & \text{On BC} \\
 q_n^d &= -\left(\frac{\rho_N}{\rho} C - C_N\right)\left(N n_z + \frac{\phi_{w0}}{|\nabla F|} \frac{\partial h_f}{\partial t}\right), & & \text{On CD} \\
 C &= 1, & & \text{On FA} \tag{12.27a}
 \end{aligned}$$

where q_n^d is the hydrodynamic dispersive flux normal to a 2nd- or 3rd-type boundary, and C_N is the normalized mass fraction in the replenishment water.

On the sea bottom boundary, EF, we may have either only landward flow from the sea, in which case the condition there is the same as on FA, or we may have both an inflow portion, MF, and an outflow portion, EM, separated by a point with zero-normal fluid flux. For the inflow portion, we usually employ either a 3rd-type condition

$$q_n^d \equiv -\mathbf{D}_h \cdot \nabla C = (1 - C)q_n \tag{12.27b}$$

or a Dirichlet condition

$$C = 1 \tag{12.27c}$$

For the outflow portion, the assumption is often made that the fluid concentrations are identical (i.e., continuous) on both sides of this boundary. Since the fluid fluxes are also identical, the condition becomes a 2nd-type condition:

$$q_n^d \equiv -\mathbf{D}_h \cdot \nabla C = 0 \tag{12.27d}$$

The (moving) point M , between the inflow and outflow portions of the boundary, is unknown a priori. In a numerical solution, during each iteration, we check whether flow along the sea bottom is directed inward or outward, and then assign the appropriate boundary condition accordingly.

We note an *inconsistency* in the salt mass fraction specified on the sea bottom segment EM between the flow boundary condition (based on the assumption of hydrostatic pressure of sea water [$C = 1$] on the sea side of EM) and the transport boundary condition (based on the assumption of equality of mass fraction [$C < 1$]). As a consequence, a large error in the salt mass balance may occur in a numerical simulation, when such an inconsistency occurs in the specified flow and transport boundary conditions. To overcome this inconsistency, we may assume the presence of a *buffer zone* on the sea bottom, which contains out-flowing water. We then take the density of this fluid into account when determining the head condition on this boundary. Iterations may be required in a numerical solution. The thickness of this buffer zone is a calibration parameter.

Let us demonstrate the solution of this flow and transport model for three example cases of sea water intrusion in coastal aquifers.

12.3 Numerical Modeling and Computer Code

As discussed earlier, the density-dependent flow equation, Equation 12.20, and the salt transport equation, Equation 12.24, in the mathematical model presented in Section 12.2, are coupled and nonlinear. In most cases, the aquifer is also heterogeneous and anisotropic, with rather irregular boundaries. Obviously, no analytical solutions are possible for the two simultaneous equations. A numerical solution is needed, using a computer program. A number of computer codes have been developed and published in the literature (e.g., Bear et al., 1999), such as FEFLOW (Diersch, 1988), MOCDFENSE (Sanford and Konikow, 1985), SUTRA (Voss, 1984; Voss and Souza, 1987), SEAWAT (Langevin, 2003; Langevin et al., 2003), DSTRAM (Huyakorn et al., 1987), CODESA-3D (Galeati et al., 1992; Putt and Paniconi, 1995), and SWIFT (Ward et al., 1984; Ma et al., 1997). However, in this chapter, we'll solve the examples by first transforming the mathematical model into finite element numerical models, which are, in turn, solved by using the computer code called FEAS (a Finite Element Aquifer Simulator). The code has been successfully used for a number of practical modeling projects in the coastal aquifer in Israel (Zhou, 1999; Bensabat et al., 2000; Bear et al., 2001; Zhou et al., 2001, 2005).

12.3.1 The Numerical Model

In the numerical model, the Galerkin finite element method (FEM) is employed to discretize the density-dependent flow equation, Equation 12.20, and the salt transport equation, Equation 12.24, and their boundary conditions, and to compute the reference head and salt mass fraction fields (Huyakorn and Pinder, 1983). In addition, the Galerkin FEM is also used to discretize the flux equation, Equation 12.16, and to compute the specific discharge and velocity field (Zhou et al., 2001). The numerical model consists of the density-dependent flow model, the consistent velocity model, the salt transport model, and the *modified Picard iterative scheme*. The salt transport equation is solved within the Eulerian–Lagrangian framework in two successive steps. The advective (Lagrangian) mass fraction is first computed by using the *adaptive pathline-based particle tracking technique* (Zhou, 1999; Bensabat et al., 2000). The remaining dispersion equation is then solved by using the Galerkin FEM. This method can be applied to any transport problem with an arbitrary mesh Peclet number in the range from 0 to infinity. The Picard method is employed for linearizing the coupled nonlinear density-dependent flow and salt transport equations. The iterative sequential procedure consists of three steps: first the density-dependent flow equation is solved for the reference head field; then the consistent specific discharge scheme is used to solve for the velocity field; finally the salt transport equation is solved for the salt mass fraction field. In each step, the iterative process continues until a satisfactory convergence is reached.

To apply the Galerkin FEM, we represent the two primary variables (reference fresh water head and salt mass fraction) and the specific discharge vector, using the basis functions:

$$h_f = h_f(x, y, z, t) \approx N_j(x, y, z) h_{fj}(t) \quad (12.28a)$$

$$C = C(x, y, z, t) \approx N_j(x, y, z) C_j(t) \quad (12.28b)$$

$$q_i = q_i(x, y, z) \approx N_j(x, y, z) q_{ij} \quad (12.28c)$$

where the $N_J(x, y, z)$ are the global basis functions, J is the nodal numbering, $h_{fJ}(t)$, and $C_J(t)$ are the nodal values of reference head and mass fraction at the t time step, respectively, and q_{IJ} is the i th component of the J th nodal specific discharge vector. Note that repeated indices indicate nodal summation.

12.3.1.1 Flow Model

Application of the Galerkin criterion and Green's theorem to the flow equation, Equation 12.20, and using the flux boundary condition on the 2nd- and 3rd-type boundaries lead to

$$\begin{aligned} \int_{\Omega} \frac{\partial N_I}{\partial x_i} K_{ij} \frac{\partial N_j}{\partial x_j} h_{fj} d\Omega + \int_{\Omega} N_I N_j S_0 \frac{dh_{fj}}{dt} d\Omega &= - \int_B N_I \frac{\rho}{\rho_f} q_n dB \\ + \int_{\Omega} N_I \frac{\rho_R Q_R - \rho Q_P}{\rho_f} d\Omega - \int_{\Omega} \frac{\partial N_I}{\partial x_i} K_{ij} \beta_c C \frac{\partial z}{\partial x_j} d\Omega - \int_{\Omega} N_I \frac{\phi \beta_c \rho}{\rho_f} \frac{\partial C}{\partial t} d\Omega \end{aligned} \quad (12.29)$$

where B is the boundary of the considered domain Ω , q_n is the outward normal flux on the boundary, and $K_{ij} = \rho g k_{ij}/\mu$, which depends on the fluid's density. We assume that the dispersive flux of the total fluid mass is negligible in the flow equation, Equation 12.20, in comparison with the advective flux. Equation 12.29 can be written in the matrix–vector form:

$$[A_{IJ}] \{h_{fJ}\} + [B_{IJ}] \left\{ \frac{dh_{fJ}}{dt} \right\} = \{F_I\} \quad (12.30a)$$

where $[A_{IJ}]$, $[B_{IJ}]$, and F_I are the stiffness (or flux) matrix, the fluid's mass storage matrix, $\{h_{fJ}\}$ is the reference head solution vector, and the right-hand side load vector, respectively, given by

$$[A_{IJ}] = \sum_e A_{IJ}^e = \sum_e \int_{\Omega^e} K_{ij}^e \frac{\partial N_I^e}{\partial x_i} \frac{\partial N_j^e}{\partial x_j} d\Omega \quad (12.30b)$$

$$[B_{IJ}] = \sum_e B_{IJ}^e = \sum_e \int_{\Omega^e} S_0^e N_I^e N_J^e d\Omega \quad (12.30c)$$

$$\{F_I\} = \{F_I^B\} + \{F_I^S\} + \{F_I^G\} + \{F_I^C\} \quad (12.30d)$$

$$\{F_I^B\} = \sum_e F_I^{Be} = - \sum_e \left(\int_{B^e} N_I^e \frac{\rho}{\rho_f} q_n dB \right) \quad (12.30e)$$

$$\{F_I^S\} = \sum_e F_I^{Se} = \sum_e \int_{\Omega^e} N_I^e \left(\frac{\rho_R}{\rho_f} Q_R - \frac{\rho}{\rho_f} Q_P \right) d\Omega \quad (12.30f)$$

$$\{F_I^G\} = \sum_e F_I^{Ge} = - \sum_e \int_{\Omega^e} \frac{\partial N_I^e}{\partial x_i} K_{ij}^e \beta_c N_J^e \frac{\partial z}{\partial x_j} d\Omega C_J \quad (12.30g)$$

$$\{F_I^C\} = \sum_e F_I^{Ce} = - \sum_e \int_{\Omega^e} N_I^e \phi \beta_c \frac{\rho}{\rho_f} N_J^e d\Omega \frac{dC_J}{dt} \quad (12.30h)$$

where Ω^e is the element domain with boundary B^e , the superscript e represents the elemental values, and \sum_e represents the global assembly of the elemental matrices and vectors over the total number of elements, $\{F_I^B\}$ is the global right-hand vector contributed from the conditions on all boundaries except Dirichlet ones, $\{F_I^S\}$ is the global vector for the sources/sinks injected/pumped by wells, $\{F_I^G\}$ is due to the body force produced by the difference of salt mass fraction in the vertical direction, and $\{F_I^C\}$ is the

term representing the fluid's mass change caused by changes in mass fraction. In the numerical model, the incorporation of $\{F_I^B\}$ and $\{F_I^S\}$ into the global system is separated from the global assembly of element systems.

For each element, we employ the backward difference time integration scheme. Thus, the resulting system of element matrix equations takes the form:

$$[AB_{IJ}^e]\{h_{fJ}^{t+1}\} = \{F_I^{e*}\} \quad (12.31a)$$

with

$$[AB_{IJ}^e] = [A_{IJ}^e] + \frac{1}{\Delta t}[B_{IJ}^e] \quad (12.31b)$$

$$\{F_I^{e*}\} = \{F_I^{Ge}\}^{t+1} + \{F_I^{Ce}\}^{t+1} + \frac{1}{\Delta t}[B_{IJ}^e]\{h_{fJ}^t\} \quad (12.31c)$$

After global assembly of element matrix systems and treatment of boundary conditions and well's sources/sinks, we have a set of nonlinear algebraic equations for the density-dependent flow

$$[AB_{IJ}]\{h_{fJ}^{t+1}\} = \{F_I^*\} \quad (12.32a)$$

with

$$[AB_{IJ}] = \sum_e [AB_{IJ}^e] \quad (12.32b)$$

$$\{F_I^*\} = \sum_e \{F_I^{e*}\} + \{F_I^S\} + \{F_I^B\} \quad (12.32c)$$

This system is nonlinear due to (1) the effects of varying fluid's density (and dynamic viscosity) on the hydraulic conductivity, K_{ij}^e , and on the specific storativity, S_0^e , and (2) the presence of two additional coupling terms: $\{F_I^G\}$ and $\{F_I^C\}$. In a phreatic aquifer, the nonlinearity also comes from the moving phreatic surface. The resulting system of algebraic equations is symmetric positive definite, and an incomplete Cholesky conjugate gradient algorithm (ICCG) (Ajiz and Jennings, 1984; Barrett et al., 1994) is used to solve the linear algebraic equations.

12.3.1.2 Transport Model

The Eulerian–Lagrangian method is used for the simulation of salt transport. This method is based on the decomposition of the mass fraction field into two components: advection and dispersion. The advection (Lagrangian) mass fraction is computed using the *adaptive pathline-based particle tracking technique* (Zhou, 1999; Bensabat et al., 2000). The remaining dispersion problem is solved by the Galerkin FEM.

By performing the Galerkin FEM procedure on the transport equation, Equation 12.24, in the Eulerian–Lagrangian framework, we obtain a system of ordinary differential equations for salt transport:

$$[T_{IJ}]\{c_J\} + [E_{IJ}] \left\{ \frac{DC_J}{Dt} \right\} = \{G_I\} \quad (12.33a)$$

where DC_J/Dt is the material derivative with respect to the velocity of solute particles, $[T_{IJ}]$, $[E_{IJ}]$, and $\{G_I\}$ are the dispersive transport matrix, the solute mass storage matrix, and the transport right-hand side

vector, respectively, given by:

$$[T_{IJ}] = \sum_e T_{IJ}^e = \sum_e \int_{\Omega^e} \frac{\partial N_I^e}{\partial x_i} \phi \rho D_{hij} \frac{\partial N_J^e}{\partial x_j} d\Omega \quad (12.33b)$$

$$[E_{IJ}] = \sum_e E_{IJ}^e = \sum_e \int_{\Omega^e} \phi \rho N_I^e N_J^e d\Omega \quad (12.33c)$$

$$\{G_I\} = \sum_e G_I^e = \sum_e \left(\int_{\Omega^e} N_I^e \rho_R Q_R (C_R - N_J^e C_J) d\Omega + \int_{B^e} N_I^e \phi \rho D_{hij} \frac{\partial C}{\partial x_j} n_i dB \right) \quad (12.33d)$$

The global right hand vector, $\{G_I\}$, results from well's sources/sinks and boundary conditions, assuming that there are no volumetric or areal sources/sinks within the solution domain.

For each element, we have the material derivative in Equation (12.33a) rewritten as

$$\frac{DC_J}{Dt} = \frac{C_J^{t+1} - C_J^L}{\Delta t} \quad (12.34)$$

where C_J^L is the Lagrangian concentration, obtained by the adaptive backward particle tracking technique and mass fraction interpolation (Bensabat et al., 2000). Using the backward difference time integration scheme, we obtain a system of element matrix:

$$[TE_{IJ}^e] \{C_J^{t+1}\} = \{G_I^*\} \quad (12.35a)$$

with

$$[TE_{IJ}^e] = [T_{IJ}^e] + \frac{1}{\Delta t} [E_{IJ}^e] \quad (12.35b)$$

$$\{G_I^*\} = \frac{1}{\Delta t} [E_{IJ}^e] \{C_J^L\} \quad (12.35c)$$

The contribution from the boundary conditions and well's sources/sinks is incorporated into the global matrix and vector in the same way as for the flow problem. After global assembly of elemental systems, the treatment of boundary conditions and well's sources/sinks, we have a set of nonlinear algebraic equations

$$[TE_{IJ}] \{C_J^{t+1}\} = \{G_I^*\} \quad (12.36a)$$

with

$$[TE_{IJ}] = \sum_e [TE_{IJ}^e] \quad (12.36b)$$

$$\{G_I^*\} = \sum_e \{G_I^*\} + \{G_I\} \quad (12.36c)$$

The coefficient matrix $[TE_{IJ}]$ is a function of salt mass fraction (or fluid's density) and flow velocity, and the right-hand vector $\{G_I^*\}$ depends on flow velocity.

12.3.1.3 Calculation of Specific Discharge

In the calculation of specific discharge, or velocity, for a density-dependent flow problem, the gradient of the reference head, $\partial h_f / \partial x_j$, and the gravity term, $\beta_c C (\partial z / \partial x_j)$, in the vertical direction in Darcy's law, Equation 12.16, must have the same order of approximation within an element. This means that a

consistent approximation is obtained by using one-order higher basis functions for the reference head than that for the fluid density (salt mass fraction) in the vertical direction. When both the reference head and the salt mass fraction are approximated by linear interpolation functions in the vertical direction, spurious vertical velocities are produced (Voss and Souza, 1987; Herbert et al., 1988). To avoid the spurious vertical velocity, we use the element-averaged salt mass fraction in the flux equation, Equation 12.16.

By applying the standard Galerkin FEM to the water flux equation, Equation 12.16, we obtain the linear algebraic equations:

$$[P_{IJ}] \cdot \{q_{ij}\} = \{R_{il}^h\} + \{R_{il}^c\} \quad (12.37a)$$

with

$$[P_{IJ}] = \sum_e P_{IJ}^e = \sum_e \int_{\Omega^e} N_I^e N_J^e d\Omega \quad (12.37b)$$

$$\{R_{il}^h\} = \sum_e R_{il}^{he} = - \sum_e \int_{\Omega^e} N_I^e \frac{K_{fij}^e}{\mu_r} \frac{\partial N_J^e}{\partial x_j} d\Omega \{h_{fj}\} \quad (12.37c)$$

$$\{R_{il}^c\} = \sum_e R_{il}^{ce} = - \sum_e \int_{\Omega^e} N_I^e \frac{K_{fij}^e}{\mu_r} \beta_c C_e \frac{\partial z}{\partial x_j} d\Omega \quad (12.37d)$$

where $[P_{IJ}]$ is the global coefficient matrix, $\{q_{ij}\}$ is the nodal unknown vector of the i th component of \mathbf{q} , $\{R_{il}^h\}$ is the right-hand side load (flux) vector, contributed by the reference head gradient, $\{R_{il}^c\}$ is the load vector, contributed by the body-force term, C_e is the element-averaged salt concentration, and $K_{fij} (\equiv \rho_f g k_{ij} / \mu_f)$ is the reference hydraulic conductivity at reference density, ρ_f , and reference viscosity, μ_f .

12.3.2 The FEAS Code

The numerical model described above can be solved by the FEAS code, which was developed for simulating sea water intrusion in 2D and 3D coastal aquifers. FEAS, written in C++, is a code family for aquifer simulations based on the finite element techniques. The FEAS code has been verified using a number of benchmark problems, such as the Henry problem, the Elder problem, and the Hydrocoin problem, and has been applied to sea water intrusion in coastal aquifers and upconing problems (Zhou, 1999; Bear et al., 2001; Zhou et al., 2005).

We shall demonstrate the solution of this flow and transport model for three cases of sea water intrusion in 3D coastal aquifers.

12.4 Examples of Sea Water Intrusion

12.4.1 Example A: Sea Water Intrusion — A Typical Case

This example involves the simulation of a sea water wedge under natural equilibrium conditions and sea water intrusion under pumping conditions in a phreatic aquifer. The aquifer strip is 11,000 m long (perpendicular to the coast) and 1,000 m wide, with a sloping sea bottom of 5%. The depth of the aquifer bottom (with a slope of 0.8%) varies from 20 m below sea level at the land-side boundary, to 100 m at the coast, and further to 108 m depth at the vertical sea-side boundary. The aquifer's top is the ground surface above the phreatic aquifer. Under natural equilibrium conditions, a fresh water flux of 0.015 m/day enters the aquifer, and the fresh water natural replenishment is at a rate of 0.1 m/yr. The aquifer is homogeneous, with a hydraulic conductivity of 30 m/day and a porosity of 0.2. The longitudinal dispersivity is 10 m and the transversal dispersivity is 1 m. Three wells, located at 5,000, 7,000, and 9,000 m from the coast, respectively, pump, when active, a total of one-third of the total natural replenishment. Note that in all figures representing a cross section normal to the coast, the figures show only part of the total investigated aquifer.

In addition to the above boundary conditions, the aquifer bottom is assumed to be impervious to both flow and transport. The vertical, sea-side boundary and the sea bottom are assumed to be at hydrostatic pressure imposed by the sea water. For transport, the mass fraction of fresh water is specified at the land-side boundary, and the mass fraction of sea water is specified at the vertical, sea-side boundary. At the sea bottom, a mixed-type boundary condition is specified. This type of condition can be automatically switched from the 1st-type condition (Equation 12.27c) to the 2nd-type condition of zero dispersive flux (Equation 12.27d), depending on the direction of the normal groundwater flux. Initially, fresh water on the inland side and sea water on the sea side of the coast (with a vertical sharp interface) is used for the sea water wedge case under natural equilibrium conditions, whose steady-state solution is considered as initial conditions for the sea water intrusion case under pumping. The FEAS code is used to simulate the transient coupled density-dependent flow and salt transport until a steady-state condition is obtained for each case.

Figure 12.5a shows the groundwater flow field, with reference head contours and streamlines originating at both the sea water and fresh water zones. Also shown is the salt mass fraction field (in a vertical cross section) under natural equilibrium conditions. With the total fresh water discharged to the sea, the sea water wedge is formed in a region within approximately 610 m from the coast. The fresh water zone receives fresh water recharged from the land-side boundary and from natural replenishment, producing a seaward fresh water flow over the sea water wedge. In the sea water wedge, a recirculating sea water flow develops. Sea water fills the wedge primarily from the sea bottom, and is swept away back to the sea within the transition zone. As demonstrated by the streamlines, the fresh water and sea water flow originating from their respective zones are separated sharply by a “*separation streamline*,” which is the envelope of all streamlines originating at the sea. The point of the separation streamline at the aquifer bottom is referred to as its “toe,” a term usually used in a sharp interface model. At the toe of the separation streamline, the contrast of horizontal hydraulic gradient between the seaward fresh water flow and the landward sea water flow is balanced, and the entire system is at an equilibrium condition. Unlike the sharp separation of sea water and fresh water flow, a wide transition zone exists between the sea water zone and the fresh water zone, because of hydrodynamic dispersion. Within the transition zone salt mass fraction varies between sea water and fresh water. The shape and width of the transition zone depend on both the flow field and the velocity-dependent dispersion. Note that the 50% salt mass fraction is not close to the separation streamline, which may be considered as the sharp interface in the case of no dispersion.

As shown in Figure 12.5b, by pumping fresh water far away from the near-coast region affected by sea water, the elevations of the phreatic surface are lowered. The lowered fresh water hydraulic head in the sea water wedge disturbs the natural equilibrium conditions, inducing further sea water movement into the coastal aquifer. The sea water intrusion starts from the toe of the separation streamline at the aquifer bottom, where the reference head on the sea side of the sea water wedge is higher than the reference head on the inland side lowered by pumping. The landward gradient of reference head leads to advective intrusion of sea water. Once the sea water advances landward, the sea water intruded in the expanded wedge moves upward along the advancing separation streamline. The upward sea water movement results in higher salt mass fraction in the overlying zone, above the already intruded zone. Meanwhile, dispersion also induces an increase in salt mass fraction in this zone. Both of advection and dispersion facilitate the widening of the intruding sea water wedge in the vertical direction. The higher salt mass fraction and wider transition zone gradually increases the reference head in the expanded intruding wedge. Meanwhile, the intruding wedge also changes the reference head of the fresh water zone in contact with it by reducing the cross-sectional area for fresh water flow to the sea, and by reducing the reference head immediately inland of the sea water wedge. As a result, a further landward movement of sea water occurs. The speed of sea water intrusion is fast immediately after the initiation of the disturbance caused by pumping. With the intruding sea water wedge becoming wider and extending landward, the landward gradient of the reference head on the sea side of the wedge becomes smaller, and the sea water intrusion is slower. A new equilibrium is finally reached by the contrast in the reduced landward fresh water flow and the recirculating sea water flow. As shown in Figure 12.5, the sea water wedge changes from 610 to 860 m from

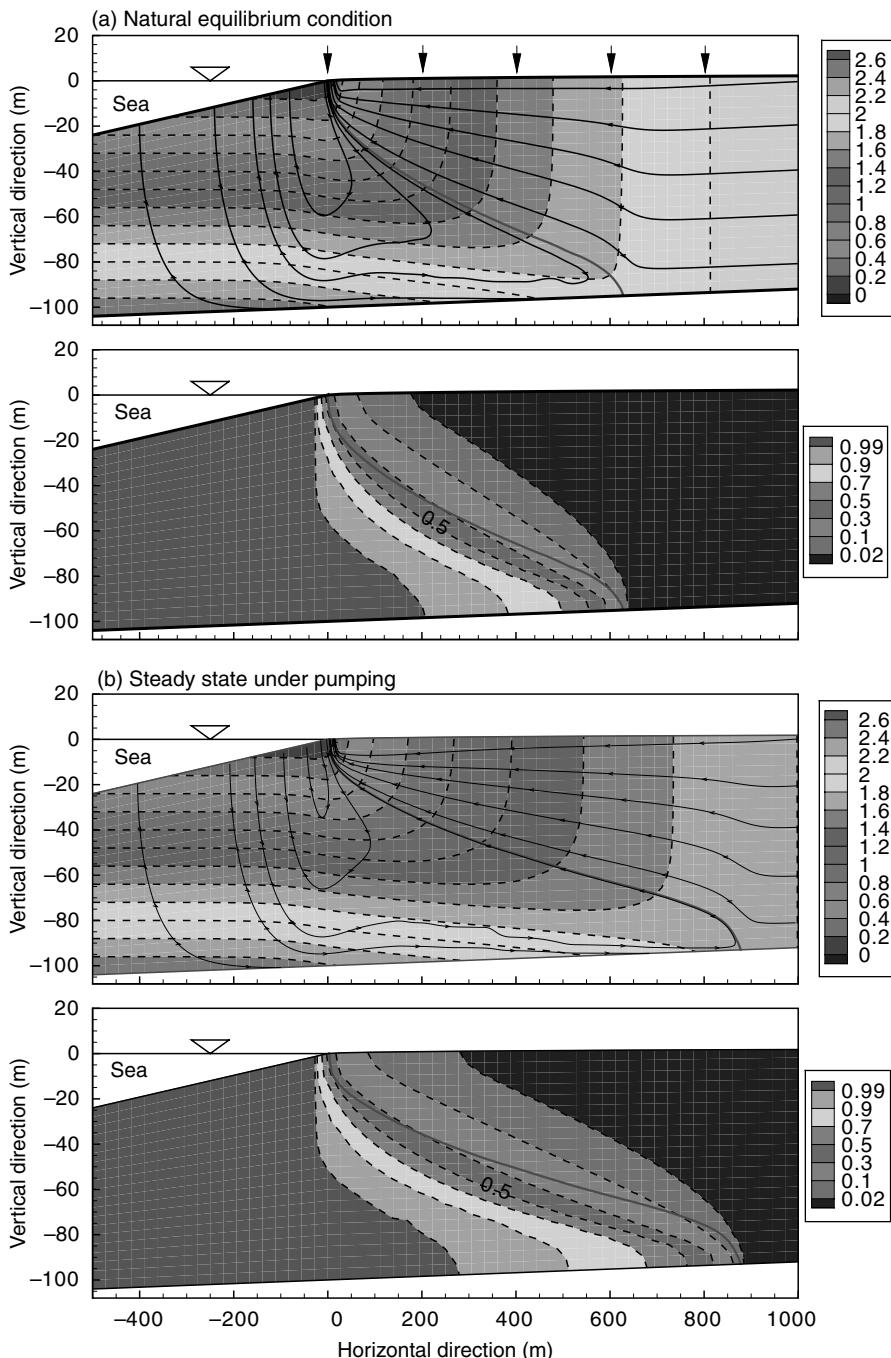


FIGURE 12.5 (See color insert following page 23-52) Numerical model results for (a) a sea water wedge under natural equilibrium conditions (no pumping) and (b) steady-state sea water intrusion under pumping conditions (three pumping wells). In each case, the figure shows the reference head distribution and streamlines in the top figure, and salt mass fraction in the bottom figure, as well as the separation streamline (in red solid line).

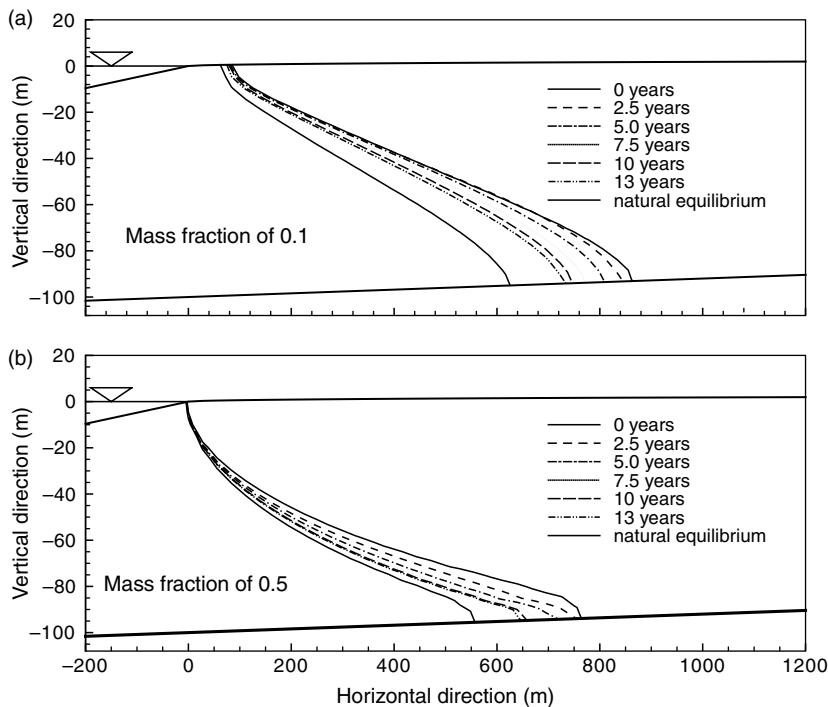


FIGURE 12.6 Transient solution of the receding sea water wedge in a phreatic aquifer after the shutoff of the three pumping wells (Example A).

the coast by pumping a third of the total fresh water discharge. The water table at the land-side boundary is lowered from 10.0 m above the sea level to 6.4 m.

Figure 12.6 shows the transient receding process of the intruding sea water wedge after the shutoff of the three pumping wells. After the wells' shutoff, the hydraulic head in the fresh water zone increases, disturbing the equilibrium maintained under the pumping conditions. This higher hydraulic head results in a higher seaward flow of fresh water, sweeping the intruding transition zone. The recovery occurs for the entire transition zone. However, the recovery process is relatively slow. After 13 years, the transition zone is approximately halfway between the initial equilibrium state under pumping conditions and that under natural equilibrium conditions. Moreover, the recovery is slowing down as it proceeds. The entire recovery process may take very long time, as demonstrated by Bear et al. (2001).

12.4.2 Example B: Sea Water Intrusion with Upconing due to Pumping

This example involves the simulation of regional sea water intrusion, combined with local transition zone upconing in a confined aquifer. The aquifer strip is 4,000 m long perpendicular to the coast, 1,000 m wide parallel to the coast, and 200 m thick. The aquifer extends 1,000 m seaward of the coast (only 500 m is shown in Figure 12.7 and Figure 12.8). The aquifer top and bottom are assumed to be horizontal surfaces. The sea bottom has a slope of 5%, varying from 0 m at the coast to 50 m below sea level at the vertical, sea-side boundary. The top and bottom boundaries of the aquifer are impervious to both groundwater flow and salt transport. An inflow flux of 0.05 m/day takes place through the vertical land-side boundary, whereas the vertical sea-side boundary and the sea bottom are assumed to be at the hydrostatic pressure of sea water. For transport, the salt mass fraction is 0 and 1 at the land-side boundary and the sea-side boundary, respectively. As in Example A, a mixed-type boundary condition is specified at the sea bottom. Initially, the aquifer is full of fresh water and sea water on the inland side and on the sea side of the coast, respectively. A uniform distribution of (isotropic) hydraulic conductivity of 30 m/day and a porosity of

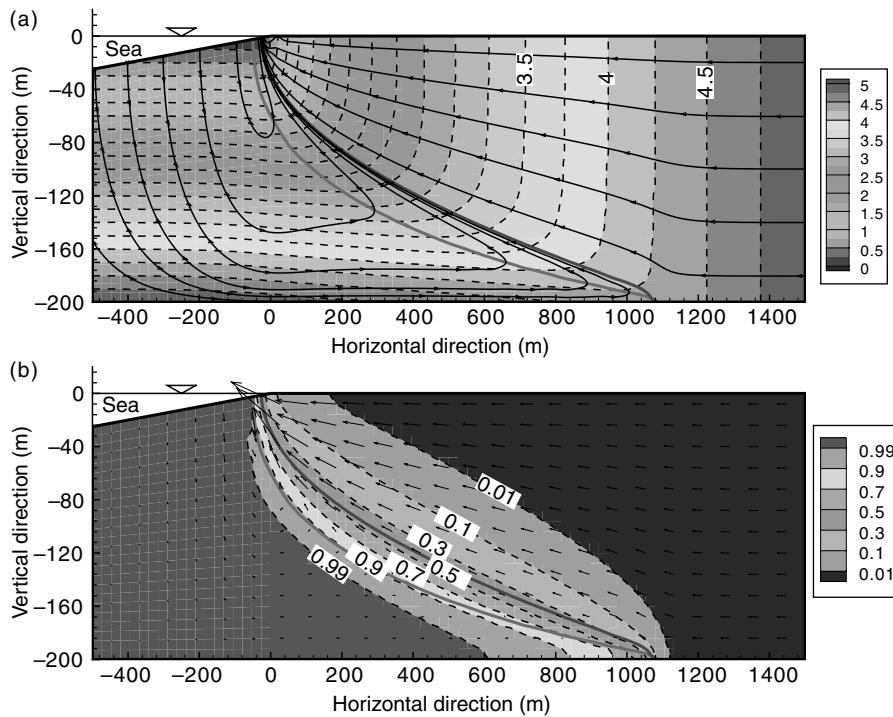


FIGURE 12.7 (See color insert following page 23-52) Steady-state solution of a sea water wedge in a confined aquifer for Example B: (a) the reference hydraulic head distribution (in dashed lines), with the separation streamline (in red line) separating streamlines (in solid lines) originating at the fresh water zone and the sea water zone, and the zero-horizontal-flux line (in pink line), and (b) the mass fraction distribution within the transition zone, and the groundwater flow velocity vector.

0.2 are used for this confined aquifer. The longitudinal and transversal dispersivities used are 10 m and 1 m, respectively. No molecular diffusion is considered, as it is negligible in comparison with dispersion. The steady-state solution of the sea water intrusion problem is obtained by simulating the transient density-dependent flow and salt transport using the FEAS code.

Figure 12.7 shows the distributions of reference hydraulic head, velocity vector, and salt mass fraction in a vertical cross section, and streamlines originating at the sea bottom, the sea-side boundary, and the land-side boundary. Fresh water flows to the sea uniformly before it reaches the sea water wedge. Above the sea water wedge, the total discharge of fresh water flows through a smaller cross section, with an increasing velocity as the sea is approached. At the coast, the cross section available for fresh water reaches its minimum, and the fresh water velocity reaches its maximum value. Sea water enters the aquifer primarily from the sea bottom; a small fraction of total sea water flux comes from the vertical sea-side boundary. As demonstrated by the velocity field and streamlines, a recirculating sea water flow occurs. For a streamline close to the coast, sea water moves downward, then turns into the transition zone, and migrates further within the transition zone to the outflow face at the sea bottom. A streamline starting at a horizontal axis of -500 m at the sea bottom shows that sea water flows downward, and then moves horizontally to the transition zone. After sea water passes through the lower portion of the transition zone, it changes direction sharply, moving to the sea. As demonstrated by the velocity magnitude, most of the sea water entering into the aquifer system comes from the small portion of the sea bottom close to the coast. Farther away from the coast, only a small quantity of sea water feeds into the aquifer system; this can also be seen from the reference head distribution 500 m away from the coast. The total sea water flow entering the transition zone moves back to the sea through a small strip of the transition zone. This strip is bounded by the lower zero-horizontal-flux line on the sea water side, and by the upper separation streamline

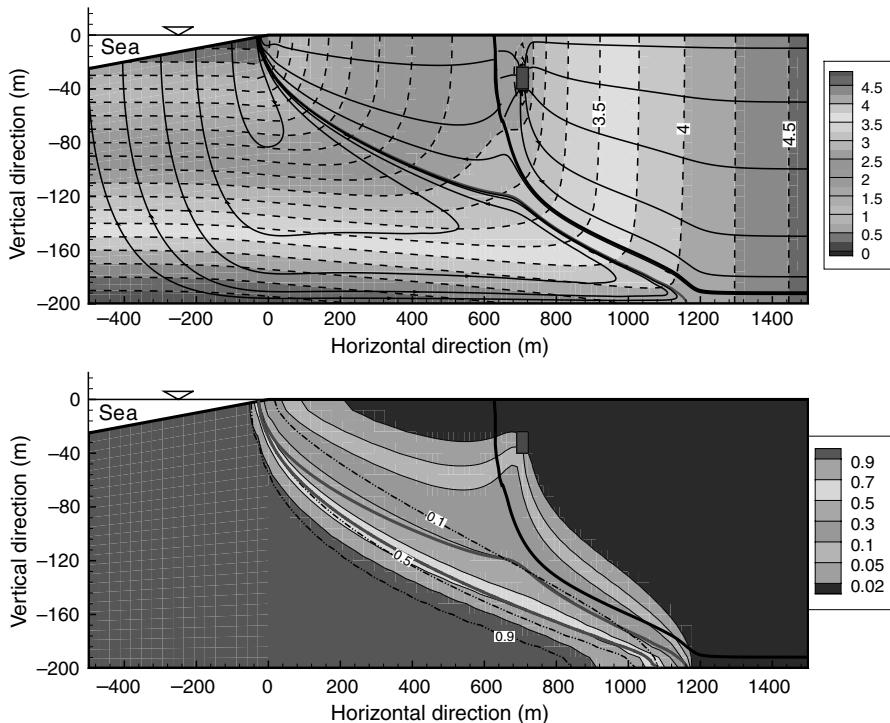


FIGURE 12.8 (See color insert following page 23-52) Transient sea water upconing in a confined aquifer at the end of 8.9 yr in Example B: (a) contours of reference hydraulic head (dashed lines), the separation streamline (red line), and streamlines (solid black lines) originating from the fresh water zone and from the sea water zone, and the capture zone (thicker solid black line), and (b) mass fraction contours (solid lines) with sea water upconing towards the pumping well, in comparison with the initial steady-state transition zone (dashed lines), and the separation streamlines for the upconing and the initial steady-state condition.

on the fresh water side (see Figure 12.7a). The “separation streamline” separates the recirculating sea water flow from the regional seaward fresh water flow, whereas the “zero-horizontal-flux” line, by definition, separates landward sea water flow on its sea side from sea-ward flow on its land side. Both lines are located within the transition zone. At the aquifer bottom, these two lines merge because no sea water flow enters this strip. Toward the sea, more sea water enters this strip, widening it. The strip reaches its maximum value at a depth of 120 m below sea level. Closer to the sea, the strip becomes narrower, but with a higher flow velocity representing more sea water flow accumulated. At the sea bottom, this strip is 22 m wide, ranging from $x = -28$ m to $x = -50$ m.

The length of the sea water wedge at the bottom is about 1080 m, with a wide transition zone. The transition zone is practically defined here as the horizontal separation between 0.99 and 0.02 salt mass fraction contour lines. It is about 500 m wide at the aquifer bottom, 700 m wide at 100 m below sea level, and 70 m at the sea level. The 0.5 contour line is close to the separation streamline, although the horizontal separation between them increases with depth. The 0.9 contour line is close to the zero-horizontal-flux line up to the depth of 140 m. This means that sea water is swept away primarily through the small strip between the zero-horizontal-flux line and separation streamline. In addition, a fraction of sea water mass dispersed into the area on the inland side of the separation streamline is swept away with fresh water flow originating in the land-side fresh water zone. In the case of no dispersion, all sea water flow is swept away through the strip between the zero-horizontal-flux line and the separation streamline, whereas fresh water flows to the sea on the inland-side of the separation streamline. This streamline may be considered as the sharp interface between sea water and fresh water. In the case in which fresh water and sea water are under hydrodynamic conditions, the zero-horizontal-flux line is different from the separation streamline, forming the pathway for recirculating sea water. Only under the Ghyben–Herzberg

assumption of hydrostatic pressure of sea water in a sharp interface model, is the zero-horizontal-flux line identical to the separation streamline, because no sea water flow is involved.

In addition to the steady-state solution of sea water intrusion in a confined aquifer, local sea water upconing induced by pumping from a well located at some distance above the interface zone, and the decay of the upconed transition zone after the well's shut-off is also simulated. The pumping well is located above the transition zone at $x = 706$ m, and is screened from -24 m to -40 m below sea level. A time-dependent pumping rate is used: a constant rate of $3000 \text{ m}^3/\text{day}$ is specified before the salt mass fraction in the pumped water reaches 3% of sea water; a constant rate of zero is specified after the well's shutoff. The transient density-dependent flow and salt transport is simulated for the upconing process under pumping and for the decay process after the well's shutoff, for a total of 13 yr. Note that this single well in the investigated aquifer strip really represents a well in an array of wells parallel to the sea spaced 1000 m apart.

Figure 12.8 shows the groundwater flow in the vertical cross section, with reference hydraulic head, streamlines, and salt mass fraction, as produced by the pumping well at the end of 8.9 yr; at that time, the mass fraction in the pumped water reaches 3% of sea water. The pumping reduces the hydraulic head at the well's screen, inducing fresh water flow toward it. In this vertical cross section, we note that most of the fresh water that enters through the land-side boundary is intercepted by the pumping well (see the capture zone shown in Figure 12.8), although only 30% of the total fresh water flow (in 3D) is intercepted and pumped out by the well. There is still a thin zone of fresh water flowing from the land-side boundary to the sea between the streamline originating immediately below the capture zone and the separation streamline. In the vicinity of the well of this thin zone, the pumping leads to a higher vertical velocity beneath the well, above the separation streamline. The pumping also reduces the hydraulic head in the area inland of the well. As a result, the separation streamline moves upward by approximately 30 m at $x = 706$ m, from the steady-state condition to its position at the end of 8.9 yr from the beginning of pumping. As the boundary condition at the sea bottom and the vertical sea-side boundary remains unchanged, a higher total sea water flow rate occurs under the pumping condition, resulting in more salt mass stored in the aquifer in the form of the salt upconed zone and within the intruding sea water wedge.

Above the separation streamline, a zone of sea water upconing toward the pumping well occurs. For example, the 5% iso-salinity contour reaches the lower portion of the well's screen, whereas the 10% contour is close to the pumping well, when salt mass fraction in the pumped water reaches 3% at the end of 8.9 years. The apparent shape of upconing sea water can be seen for lower iso-salinity contours, whereas higher iso-salinity contours (say, 30%) do not show the shape of upconing. This is consistent with the findings of Zhou et al. (2005). The development of the upconed sea water zone toward the pumping well depends on both advection and dispersion. The iso-salinity contours of lower salt mass fraction are similar to the streamlines; this is particularly true in the vicinity of the pumping well. The higher vertical velocity brings mixed water of higher salt mass fraction to the areas close to the well. However, advection alone is not sufficient to salinify the pumping well, because no streamline originating at the sea reaches the well and the well is far above the separation streamline. Dispersion also plays a key role in salinizing the well. A transition zone, with mixed water far above the separation streamline provides a possibility for the well to be more easily salinized than under the sharp interface condition. Both advection and dispersion contribute to the increase in the salinity of the pumped water, as described by the transition zone model.

The pumping also induces some additional regional sea water intrusion. For example, the 50% iso-salinity contour moves further landward by 90 m at the aquifer bottom. Other higher-mass fraction iso-salinity (say, 0.9) contours also slightly move landward, although the width of the transition zone in terms of high mass fraction (say, from 99 to 30%) remains unchanged. This is unlike the transition zone for lower mass fractions (say $\leq 30\%$), which changes significantly as a result of pumping. The separation streamline also moves toward the pumping well, with a small upconed zone beneath the pumping well. However, this upconing is far less significant than that of the lower mass fraction contours. The latter is critical for the salinization of pumping wells in coastal aquifers. This may also indicate that a sharp interface model may significantly overestimate the critical pumping rate.

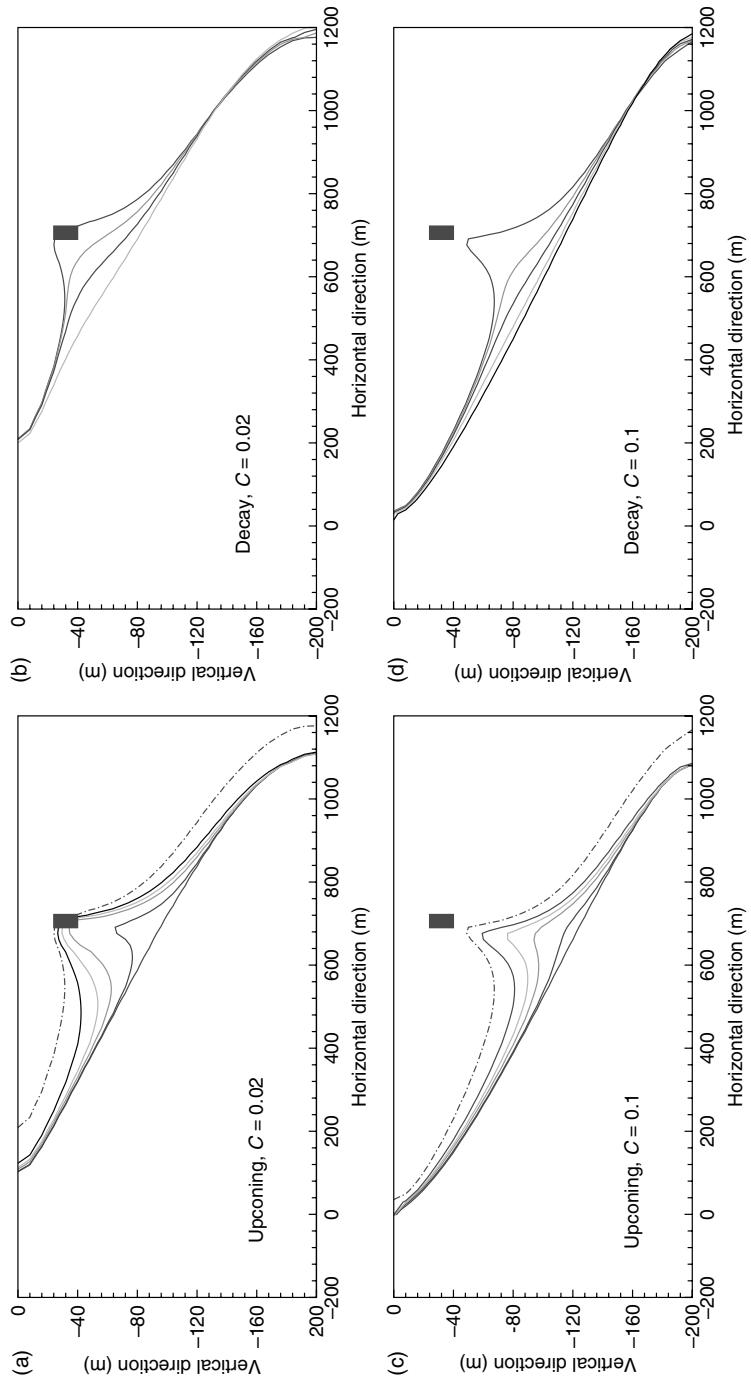


FIGURE 12.9 (See color insert following page 23-52) Transient upconing and decay processes in a confined aquifer (Example B). The times for the upconing process are 53, 239, 439, 839, and 3239 days (well's shutoff), and 3239, 3439, 3639, 4039, and 4639 days for the decay process, all from the beginning of the pumping.

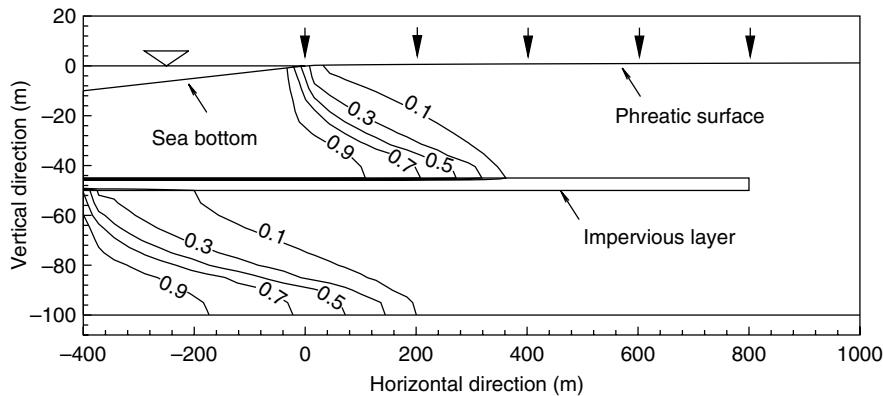


FIGURE 12.10 Steady-state solution of the multiple sea water wedges in a double-layered phreatic aquifer (Example C).

As shown in Figure 12.9, at the early stage of upconing, the mixed water of low mass fractions upcones toward the pumping well more significantly than the rise of the higher mass fraction contours. During this stage, only the mixed water in the vicinity of the well, is affected by pumping whereas the mixed water in the area close to the aquifer bottom and that close to the coast does not change significantly. This is because pumping above the transition zone lowers the hydraulic head on its land-side, and this does not change the balance between fresh water and sea water at the toe of the separation streamline. The enhanced sea water flow from the sea brings more sea water to the upconed sea water zone. This gradually increases the hydraulic head below the transition zone. In the second stage, the increase in hydraulic head by the upconed sea water violates the gradient balance at the toe, thus leading to additional regional sea water intrusion. Following the well's shutdown, the sea water upconed zone is swept seaward quickly because of the strong fresh water flow. The upconing occurs beneath the pumping well, although it slightly extends to its sea side, whereas the decay process occurs primarily by the seaward sweep. Particularly interesting is the observation that four years after the well's shutdown, the intruding sea water wedge in the lower portion of the aquifer has not changed significantly. This may indicate that the recovery of the aquifer from the upconing is slow, although it will eventually be completely recovered (Bear et al., 2001).

12.4.3 Example C: Sea Water Intrusion in a Double-Layered Aquifer

This example simulates sea water intrusion in a double-layered aquifer. An impervious layer extends from the sea-side boundary to 800 m inland from the coast. The layer, about 5 m thick, is located in the middle of the phreatic aquifer. This aquifer is 4000 m long, and 100 m thick. The bottom is a horizontal surface, while the top is the ground surface with the phreatic surface below it. The replenishment of a rate of 50 mm/yr is specified on the phreatic surface, and a flow rate of 0.01 m/day is specified on the land-side boundary. The other flow boundary condition is the same as in Example A. The transport condition on the vertical sea-side boundary and the sea bottom is specified as the mixed-type condition (Equation 12.27). The aquifer properties remain unchanged from Example A.

Figure 12.10 shows the steady-state solution of sea water intrusion in the double-layered phreatic aquifer. We observe two sea water wedges, separated by the impervious layer. As the flow conditions at the seaside boundary are different for the upper phreatic aquifer and the lower confined one, their reference head fields are different. A higher reference head occurs in the lower confined aquifer.

12.5 Introduction to the Management of a Coastal Aquifer

Management of a considered system means making decisions concerning the planning and operation of the system to achieve specified goals, without violating technical and non-technical constraints imposed on the

managed system. The objectives and constraints are dictated by the *policy* that governs the management of the system. An aquifer should be regarded as a system that has to be managed. Depending on the prevailing (say, regional or national) water policy, a considered aquifer may be managed as a stand-alone system, or as an element in a large water resources system that may include a number of aquifers and surface water sources that are managed as a single system. Treated sewage and desalinated sea water, or brackish water, may also constitute sources of water in the system. In such a system, water from one area may be transported and supplied in another area, and water from one source, say a surface water source, may be conveyed to be artificially injected into an aquifer (e.g., for storage purposes), in the same or in another region, to improve the overall efficiency of the considered water resources system.

Strictly, management of an aquifer means making decisions concerning the pumping and artificial recharge of the aquifer — total volume of annual withdrawal, rates, and location of wells, quality of pumped water, and the like, subject to imposed technical and non-technical constraints. These decisions are made so as to maximize or minimize the *objective function* (or functions in a multi-objective case) dictated by the imposed policy.

Examples of possible objective functions are:

- Maximize the sustainable yield of the aquifer.
- Minimize the cost of supplying the specified demand for water of a given population.

Here, *sustainable yield* of an aquifer is the annual quantity of water that can be withdrawn from the aquifer, *year after year*, as a constant volume, or as one that varies according to some rule, without violating specified constraints, especially *that the source will be preserved (quality and quantity) forever*.

Examples of possible constraints are:

- Dictated range of allowed water levels (may vary with location).
- Minimum spring discharge, or discharge to wetlands, to a stream or to a lake.
- Maximum permissible salt concentration (may vary with location) in pumped water, or in the aquifer.
- Maximum cost of supplied unit volume of water.

Examples of non-technical constraints are:

- Water rights and legal aspects.
- Allocation of water to consumers according to sector/region/type of users.
- Maximum costs, limited budget, and other economic considerations.

The management of a considered aquifer requires a calibrated hydrogeological model that describes flow and solute/contaminant transport phenomena that take place in it. This model will provide the response of the aquifer to any proposed management scheme, thus enabling the planner or decision maker to eliminate those schemes that violate the imposed constraints, and to evaluate each proposed management scheme in terms of the objective function. In fact, satisfying the flow and transport model constitutes a constraint in the management model. Unfortunately, a practical management model that incorporates this constraint is still missing and trial and error approaches are often employed. Having a calibrated hydrogeological model for a given aquifer means (1) having a good conceptual-mathematical-numerical model and a verified computer code to solve it, and (2) sufficient data to perform model calibration (=parameter estimation).

An important outcome of aquifer calibration is an estimate of the natural replenishment, whether from precipitation, as inflow through aquifer boundaries, say, from adjacent aquifers, or as inflow from streams and lakes. The replenishment may be estimated as some average value, as statistics of values, or as a relationship to precipitation (for which measured information is, usually, more readily available). Under certain conditions, the inflow from adjacent aquifers can be related to water levels in these aquifers. Knowledge of the natural replenishment is essential for determining the sustainable yield of a considered aquifer. In principle, unless imported water is added by artificial recharge techniques, in the long run, the

sustainable yield cannot exceed the natural replenishment. In fact, it must be lower, as we must allow for flushing of accumulated salts/contaminants.

With these introductory remarks in mind, let us now consider the management of a coastal aquifer. As already emphasized above, essentially, the basic approach to the management of a coastal aquifer is the same as for any other aquifer, except that the presence of the sea constitutes a continuous threat to the quality of groundwater in the vicinity of the sea through the mechanism of *sea water intrusion* (Emch and Yeh, 1998; Das and Datta, 1999; Rao et al., 2004). We shall focus on the aquifer as a stand-alone system.

In a coastal aquifer, the sustainable yield is determined so as to allow the interface (= interface zone) to encroach up to some predetermined optimal distance. The management approach should be seeking to optimize sea water intrusion according to some objective function and specified constraints, rather than to stem sea water intrusion.

For the sake of simplicity, let us assume steady flow, a sharp interface, and the Ghyben–Herzberg relationship (equivalent to the assumption of essentially horizontal flow) between the elevations above sea level of the water table in a phreatic aquifer and the depth of the interface below sea level. In this simplified Ghyben–Herzberg model, presented in Section 12.1, the volume of fresh water displaced by the advancing interface between two steady states is equal to 40 times (when density of fresh water and sea water is 1.000 and 1.025 g/cm³, respectively), the volume of water between the two steady states of the water table. Thus, for the interface to advance, large quantities of ground water (between the initial and final interfaces) have to be removed from the aquifer. In reality, however, under unsteady flow conditions, if we start from some steady state, the changes in water table elevations, say in response to excessive pumping — annual or a long-term trend — are much faster, with a corresponding rise of the interface.

As emphasized above, the extent of sea water intrusion is dictated by the level of withdrawal from the aquifer. As such, it is a decision variable in the management problem. With this in mind, a coastal aquifer may be in one of three situations:

- *Virgin conditions.* (Practically) no pumping and no/little sea water intrusion takes place. We wish to develop the aquifer by pumping up to its full or optimal (sustainable) yield, with/without artificial recharge (assuming that the natural replenishment is known).
- *Developing conditions.* Pumping and sea water intrusion take place, but no wells have been salinized (yet). We wish to continue to operate the aquifer and develop it up to its sustainable yield. We may need more wells, or move wells to new locations.
- *Excessive sea water intrusion conditions.* Withdrawal exceeds sustainable yield, and significant sea water intrusion takes place. Some wells have already been salinized. We wish to restore the aquifer to sustainable yield conditions.

In the first and second cases, withdrawal can be increased up to the estimated sustained yield. In fact, as explained earlier, for a number of years, the withdrawal may exceed the ultimate sustainable yield (obviously, if consumers are available). We often refer to this excess pumping as “pumping the one-time-reserve.” In the third case, withdrawal has to be reduced. To speed up restoration, withdrawal may at first, and for some years, be reduced to below sustainable yield, or artificial recharge increased. We have to take into account two important factors characterizing the restoration process:

- Changes in the position of the interface are slow, so that the restoration to sustainable yield with an optimal position of the interface may take many years.
- Furthermore, as in reality no sharp interface exists, as the transition zone moves seaward, it tends to widen, and it may take a very long period to really flush the saline water and reduce salinity near the coast to desirable levels (Bear et al., 2001).

An important feature that should not be overlooked stems from the fact that as the transition zone retreats, and fresh water replaces saline water in the void space, the clay fraction of the soil disperses, and the permeability is reduced.

As emphasized above, the extent of sea water intrusion depends on the flow rate of fresh water allowed to discharge to the sea. Obviously, this discharge means maintaining an appropriate seaward gradient

of water levels (or, piezometric heads in a confined aquifer). By maintaining an appropriate water table elevation at a certain distance from the sea, we can arrest the intruding sea water wedge at that (optimal) distance from the sea, thus preventing the wedge from advancing farther inland. Beyond the “fresh water mound” that can be maintained at a desired distance from the sea, we can manage the aquifer with water levels fluctuating above and even below sea level. Obviously, the issue is, then, how to maintain this mound. One obvious way is to control pumping so as to allow a certain fresh water discharge to the sea. Another option is to maintain the mound by artificial recharge through infiltration basins or an array of wells along a line parallel to the coast at the desired distance from the coast. Recharge can be implemented with imported water from other regions, and even by treated sewage (obviously, taking into account that part of the water will also flow landward).

Other “engineering” techniques have been proposed from time to time for controlling, and even preventing sea water intrusion. Among them we can mention a “trough” parallel to the coast, produced by pumping (Huisman and Olsthoorn, 1983), and an impervious or semi-pervious barrier constructed by various geomechanical techniques parallel to the coast (Todd, 1980). The latter will be effective only if it extends over a significant length.

A well located above the sea water wedge, usually, pumps saline water, even when its screen is located at some distance above the interface (or transition zone). This is due to the phenomenon known as interface “upconing.” However, if the screen is sufficiently high and the rate of pumping sufficiently small, the well will pump fresh water, thus reducing the actual discharge to the sea. In Israel, this technique — called “coastal collector” — takes the form of an array of low discharge shallow wells along a line located at a rather small distance from the sea.

As emphasized above, many coastal aquifers are contaminated from point and distributed sources at ground surface. Furthermore, groundwater in coastal aquifers, especially close to the coast, is, to some extent, saline, because of airborne salts originating in the sea. Aquifer water draining to the sea will prevent the accumulation of salts and other contaminants in the aquifer. The management of the aquifer should take this aspect into account. However, there is no need to elaborate on this topic, as, with respect to contamination, the coastal aquifer is not different from any other aquifer (except for the differences already mentioned earlier). Because of the threat to water quality in the underlying aquifer, there is a strong interaction between aquifer management and issues related to land use, with the objective of minimizing the threat to groundwater quality.

To summarize, the management of a coastal aquifer requires a calibrated model of the hydrogeology of the aquifer, preferably a 3D variable density flow and transport model. Such models are readily available nowadays. The selected management scheme should be based on an objective function and constraints dictated by the policy established within the framework of some system higher in the hierarchy, for example, the regional or national one. **The basic feature of management of a coastal aquifer is the determination of a sustainable yield that will arrest the sea water wedge at a desired distance from the sea by allowing a certain discharge of fresh water to the sea. A fresh water mound and a coastal collector are useful tools for increasing the efficiency of aquifer management.**

Glossary

Coastal aquifer management Determining the spatial distribution of water withdrawal and artificial recharge in a costal aquifer, as well as the quality of the pumped water, taking into account the relationship between the extent of sea water intrusion and the rate of fresh water discharge to the sea.

Coastal collector An array of pumping wells parallel to the coast, at a rather small distance from the latter, and with screens located at some distance above the assumed sharp interface (or the transition zone), such that when they pump at a sufficiently small rate, they do not become salinized by the upconing interface.

Density difference factor The ratio of density difference between (pure) sea water and fresh water to the density of fresh water.

Fresh water Water in a coastal aquifer, at a sufficient distance from the coast such that its salinity is not affected by the presence of sea water.

Ghyben–Herzberg assumption The assumption that in the region of sea water intrusion, with a sharp interface separating between the fresh water and the sea water bodies, fresh water equipotentials are vertical (i.e., flow is horizontal) and sea water is stationary. This leads to a fixed ratio between the elevation of the water table (or piezometric head in a confined aquifer) and the depth of the interface at that point, which depends on the density of sea water.

Reference piezometric head The piezometric head, with the density taken as some reference density, usually that of fresh water.

Sharp interface approximation The assumption that, although fresh water and sea water are miscible fluids, the body of fresh water and that of sea water in a coastal aquifer are separated by a sharp surface.

Sea water intrusion The intrusion (or additional intrusion) of sea water into a coastal aquifer as a consequence of a reduction in the fresh water flow discharged to the sea.

Sea water wedge The shape of the stationary (or steady state) body of sea water intruding into a coastal aquifer, under the sharp interface approximation

Separation streamline The 3D surface that sharply separates between streamlines originating from the inland fresh water zone and recirculating streamlines originating from the sea water zone. This separation surface is located within the transition zone. In the sharp interface approximation, the separation streamline coincides with the sharp interface.

Transition zone The zone in a coastal aquifer across which the salinity of the water varies from that of fresh water to that of sea water.

Upconing The rise of an assumed sharp interface, or of a transition zone, towards a well that pumps water from a coastal aquifer, with its screen located above that interface or transition zone.

Variable density flow Flow in a porous medium domain, in which the determination of the flux takes into account that the fluid's density depends, among other factors, on the concentration of total dissolved salts.

Zero-horizontal-flux surface The surface of zero horizontal flux, separating between the seaward and landward horizontal components of flow. The upward vertical flow through this surface eventually moves toward the sea through the volume between this surface and the overlying separation streamline surface.

Acknowledgment

The authors wish to thank Dr. Christian Langevin of the USGS for his useful review comments and suggestion, which helped us to improve this chapter. Thanks are also due to Dr. Jacob Bensabat, EWRE, who allowed us to use the FEAS code, developed jointly by Drs. Quanlin Zhou and Jacob Bensabat, for the simulation of the typical sea water intrusion cases presented in this chapter.

References

- Ajiz, M. A. and Jennings, A. 1984. A robust incomplete Choleski conjugate gradient algorithm. *Int. J. Numer. Methods Eng.* 20:949–966.
- Barrett, R., Berry, M., Chan, T. F., Demmel, J., Donato, J. M., Dongarra, J., Eijkhout, V., Pozo, R., Romine, C., and Van der Vorst, H. 1994. *Templates for the Solution of Linear Systems: Building Blocks for Iterative Methods*. SIAM, Philadelphia.
- Badon-Ghyben, W. 1888. *Nota in Verband met de Voorgenomen Putboering Nabij Amsterdam* (Notes on the probable results of well drilling near Amsterdam). *Tijdschr. Kon. Inst. Ing.*, The Hague.
- Bear, J. 1972. *Dynamics of Fluids in Porous Media*. American Elsevier, New York (also Dover Publications, 1988), p. 764.
- Bear, J. 1979. *Hydraulics of Groundwater*. McGraw-Hill, New York, p. 569.

- Bear, J. 1999. Conceptual and mathematical modeling, in *Seawater Intrusion in Coastal Aquifers — Concepts, Methods and Practices*, J. Bear, A. H.-D. Cheng, S. Sorek, D. Ouazar, and I. Herrera (Eds.). Kluwer Academic Publishers, Dordrecht, p. 625.
- Bear, J. and Bachmat, Y. 1990. *Introduction to Modeling of Transport Phenomena in Porous Media*. Kluwer Academic Publishers, Dordrecht, p. 553.
- Bear, J. and Verruijt, A. 1987. *Modeling Groundwater Flow and Pollution*. D. Reidel Publishing Co., Dordrecht, p. 414.
- Bear, J., Cheng, A. H.-D., Sorek, S., Ouazar, D., and Herrera, I. (Eds.) 1999. *Seawater Intrusion in Coastal Aquifers — Concepts, Methods and Practices*. Kluwer Academic Publishers, Dordrecht, p. 625.
- Bear, J., Zhou, Q., and Bensabat, J. 2001. Three dimensional simulation of seawater intrusion in heterogeneous aquifers, with application to the coastal aquifer of Israel, *Proc. First Inter. Conf. on Saltwater Intrusion and Coastal Aquifers — Monitoring, Modeling and Management*, Essaouira, Morocco, April 23–25.
- Bensabat, J., Zhou, Q., and Bear, J. 2000. An adaptive pathline-based particle tracking algorithm for the Eulerian–Lagrangian method. *Adv. Water Resour.* 23:383–397.
- Cheng, A.H.-D. and Ouazar, D. 2004. *Coastal Aquifer Management: Monitoring, Modeling, and Case Studies*. Lewis Publishers, CRC Press, Boca Raton, FL, p. 280.
- Diersch, H.-J. 1988. Finite element modeling of recirculating density driven saltwater intrusion processes in groundwater. *Adv. Water Resour.* 2:25–43.
- Das, A. and Datta, B. 1999. Development of multiobjective management models for coastal aquifers. *J. Water Resour. Plan. Manage.* 125:76–87.
- Du Commun, J. 1828. On the cause of fresh water springs, fountains, etc. *Amer. J. Sci. Arts* 14:174–175.
- Emch, P. G. and Yeh, W. W.-G. 1998. Management model for conjunctive use of coastal surface water and ground water. *J. Water Resour. Plan. Manage.* 124:129–139.
- Frind, E. O. 1982. Simulation of long-term transient density-dependent transport in groundwater. *Adv. Water Resour.* 5:73–98.
- Galeati, G., Gambolati, G., and Neuman, S. P. 1992. Coupled and partially coupled Eulerian–Lagrangian model of freshwater-seawater mixing. *Water Resour. Res.* 28:149–165.
- Herbert, A. W., Jackson, P., and Lever, D. A. 1988. Coupled groundwater flow and solute transport with fluid density strongly dependent upon concentration. *Water Resour. Res.* 24:1781–1795.
- Herzberg, A. 1901. Die Wasserversorgung einiger Nordseebaden (The water supply on parts of the North Sea coast in Germany). *Z. Gasbeleucht. Wasserversorg.* 44:815–819, 824–844.
- Huisman, L. and Olsthoorn, T. N. 1983. *Artificial Groundwater Recharge*. Pitman Adv. Publishing Program, p. 320.
- Huyakorn, P. S. and Pinder, G. F. 1983. *Computational Methods in Subsurface Flow*. Academic Press, Orlando, p. 473.
- Huyakorn, P. S., Andersen, P. F., Mercer, J. W., and White, Jr., H. 1987. Saltwater intrusion in aquifers: development and testing of a three-dimensional finite element model. *Water Resour. Res.* 23: 293–312.
- Konikow, L.E. and Reilly, T.E. 1999. Seawater intrusion in the United States, in Bear *et al.*, 1999.
- Langevin, C. D. 2003. Simulation of submarine ground water discharge to a marine estuary: Biscayne Bay, Florida. *Ground Water* 41:758–771.
- Langevin, C. D., Shoemaker, W. B., and Guo, W. 2003. MODFLOW-2000, the U.S. Geological Survey modular ground-water model — Documentation of the SEAWAT-2000 version with the variable-density flow process (VDF) and the integrated MT3DMS transport process (IMT). U.S. Geological Survey Open-File Report 03-426, p. 43.
- Lever, D. A. and Jackson, C. P. 1985. On the equations for the flow of a concentrated salt solution through porous medium. Harwell Rep. AERE-R, 11765. Her majesty's Stationary Off., London.
- Ma, T.-S., Sophocleous, M., Yu, Y.-S., and Buddemeier, R. W. 1997. Modeling saltwater upconing in a freshwater aquifer in south-central Kansas. *J. Hydrol.* 201:120–137.

- Putt, M. and Paniconi, C. 1995. Picard and Newton linearization for the coupled model of saltwater intrusion in aquifers. *Adv. Water Resour.* 18:159–170.
- Rao, S. V. N., Sreenivasulu, V., Bhallamudi, S. M., Thandaveswara, B. S., and Sudheer, K. P. 2004. Planning groundwater development in coastal aquifers. *Hydrol. Sci. J.* 49:155–170.
- Reilly, T. E. and Goodman, A. S. 1985. Quantitative analysis of saltwater–freshwater relationships in ground-water systems — a historical perspective. *J. Hydrol.* 80:125–160.
- Sanford, W. E. and Konikow, L. F. 1985. A two-constituent solute transport model for ground water having variable density. U.S. Geological Survey Water-Resources Investigations Report 85-4279, p. 89.
- Sorek, S. and Pinder, G. F. 1999. Survey of computer codes and case histories, in *Seawater Intrusion in Coastal Aquifers — Concepts, Methods, and Practices*, J. Bear, A. H.-D. Cheng, S. Sorek, D. Ouazar, and I. Herrera (Eds.). Kluwer Academic Publishers, Dordrecht, The Netherlands.
- Todd, D. K. 1980. *Groundwater Hydrology*, 2nd ed. John Wiley & Sons, New York, USA, p. 535.
- Voss, C. I. 1984. A finite-element simulation model for saturated–unsaturated, fluid-density-dependent ground-water flow with energy transport or chemically-reactive single-species solute transport. U.S. Geological Survey Water-Resources Investigations Report 84-4369, p. 409.
- Voss, C. I. and Souza, W. R. 1987. Variable density flow and solute transport simulation of regional aquifers containing a narrow freshwater–saltwater transport zone. *Water Resour. Res.* 23:1851–1866.
- Ward, D. S., Reeves, M., and Duda, L. E. 1984. Verification and field comparison of the Sandia Waste-Isolation Flow and Transport Model (SWIFT). NUREG/CR-3316, SAND83-154, Sandia National Laboratories, Albuquerque, New Mexico.
- Zhou, Q. 1999. Modeling seawater intrusion in coastal aquifers. Ph.D. thesis, Technion-Israel Institute of Technology.
- Zhou, Q., Bensabat, J., and Bear, J. 2001. Accurate calculation of specific discharge in heterogeneous porous media. *Water Resour. Res.* 37:3057–3069.
- Zhou, Q., Bear, J., and Bensabat, J. 2005. Saltwater upconing and decay beneath a well pumping above an interface zone. *Transp. Porous Media* DOI: 10.1007/s11242-005-0261-4, 61:337–363.
- Zhou, Q., Gelhar, L. W., and Jacobs, B. 2002. Comparison of the field-scale effective properties of two-phase flow in heterogeneous porous media obtained by stochastic analysis and numerical experiments, in *Proc. Int. Groundwater Symp. on Bridging the Gap between Measurements and Modeling in Heterogeneous Media*, Berkeley, California, March 25–29.

Further Information

The conceptual and mathematical modeling based on the sharp interface assumption can be found in many text books. Examples are Bear (1972, 1979) and Bear and Verruijt (1987). Recent applications of mathematical and numerical modeling to sea water intrusion in coastal aquifers can be found in the books of Bear et al. (1999) and Cheng and Ouazar (2004).

The FEAS code used for the typical examples in this chapter is a finite element aquifer simulator for sea water intrusion and brine transport in 2D and 3D aquifers. It can also be used to simulate coupled groundwater flow and heat transport, and multiphase flow for dense nonaqueous phase liquids (DNAPL) migration (as well as decoupled groundwater flow and solute transport in confined and unconfined aquifers). The application of FEAS to practical problems includes both sea water intrusion (Bear et al., 2001) and upconing (Zhou et al., 2005), and DNAPL migration in heterogeneous porous media (Zhou et al., 2002). For further information of the code, please contact Quanlin Zhou at Lawrence Berkeley National Laboratory or Jacob Bensabat at Environmental and Water Resources Engineering Inc., PO Box 6770, Haifa, 31067, Israel.

13

Groundwater and Heat Flow

13.1	Introduction.....	13-1
13.2	Basic Thermodynamics..... Heat as Energy • Temperature: Celsius and Fahrenheit • Heat Units: Calories and British Thermal Units • Thermal Conductivity • Thermal Resistivity • Specific Heat of Water • Three-Dimensional Hexagonal Structure of Water • Thermal Inertia • Viscosity and Heat • Relationship Between Thermal Energy and Mechanical Energy (Joule's Experiments)	13-2
13.3	Flow of Heat..... Thermal Gradients • Heat Flux • Conduction of Heat • Convection of Heat • Advection versus Convection • Diffusion of Heat	13-5
13.4	Sources of Heat Flow to and from Earth's Surface.... Solar Radiation • Geothermal Gradient and Heat Flow • Hydrothermal Activity • Geothermal Energy for Electric Power • Annual Air-Temperature Changes • Permafrost • Heat Exchange between Surface Water and Groundwater and Effects upon Aquatic Environments • Effects of Groundwater-Transported Heat upon Hydrogeochemistry	13-6
13.5	Applications	13-11
	Using Surface Heat Distributions to Locate Groundwater Resources • Using Temperature/Depth Profiles in Water Wells to Determine Vertical Velocities of Groundwater Movement • Distinguishing Different Aquifers by Temperature Profiles in Boreholes • Expansion of Formation Water in Deep Boreholes • Storage of Both Hot and Cold Water in Geological Formations	
	Glossary.....	13-13
	References	13-14
	Further Information	13-16

Darrell I. Leap
Purdue University

13.1 Introduction

Groundwater flow and heat flow are similar in that they both require gradients (hydraulic and thermal, respectively), hydraulic conductivity, and thermal conductivity, and both are assumed to take place

in a continuum in most conditions. Due to the similarities, both processes can be explained by very similar equations. Early investigators visualized these similarities and applied the already-established thermodynamic theories of Fourier and Newton to groundwater flow.

Thiem (1906) formulated equations of steady-state groundwater flow based on Fourier's steady-state equations of heat flow. Theis (1935) derived his famous similarity solution to the non-equilibrium diffusion equation for groundwater flow (Domenico and Schwartz, 1990) that had the same structure as the non-equilibrium heat flow equation as given by Fourier's similarity solution for non-equilibrium heat flow (Carslaw and Jeager, 1959; Ozisik, 1968). Aside from the historically and academically interesting parallels between the equations for both heat flow in solids (treated as a continuum) and groundwater flow in porous media (also treated as a continuum), there are applications of great practical importance and interest of groundwater flow and heat flow. Notable among these are: (1) groundwater flow and heat flow to streams and lakes, (2) heat flow in groundwater by advection and convection, (3) heat in groundwater and geochemical reactions including mineral deposition, (4) locating groundwater resources by heat profiles over an area, (5) estimating groundwater and contaminant velocities by temperature profiles in boreholes, (6) locating different aquifers by temperature profiling in boreholes, and (7) locating and analyzing geothermal heat resources for energy production.

13.2 Basic Thermodynamics

13.2.1 Heat as Energy

Energy is defined as the ability to produce change (e.g., do work), and work is expended energy (Leap, 2003). Potential energy is defined as energy available to do work but in storage, waiting to be used. The heat content (amount of heat energy) in a solid or fluid is most often measured by *temperature*—defined as the average kinetic energy of the particles in the solid or fluid being measured.

13.2.2 Temperature: Celsius and Fahrenheit

Temperature is measured by thermometers graded in the Celsius (formerly called “centigrade”) or Fahrenheit scale, or both. German-born scientist Gabriel Daniel Fahrenheit developed his scale in 1706 and Swedish astronomer Anders Celsius developed his in 1742 (Wilson and Arnold, 1958). The Fahrenheit scale is in common use in English-speaking countries, but not in England itself which now uses the Celsius scale as its primary temperature scale.

In the Fahrenheit scale, the freezing point of water is -32°F and the boiling point of water is 212°F . In the Celsius scale, the freezing point of water is 0°C and its boiling point is 100°C . Conversion from one scale to another can be accomplished by the formulas (Halliday and Resnick, 1970):

$$\text{Celsius to Fahrenheit} \quad T_{\text{F}} = 32^{\circ}\text{F} + (9/5)T_{\text{C}} \quad (13.1\text{a})$$

$$\text{Fahrenheit to Celsius} \quad T_{\text{C}} = (T_{\text{F}} - 32^{\circ}\text{F})/(9/5) \quad (13.1\text{b})$$

where C is Celsius, F is Fahrenheit, T_{F} is Fahrenheit temperature, and T_{C} is Celsius temperature.

A third temperature scale, the Kelvin scale, was invented in the 19th century by the English physicist, William Thompson who later became Lord Kelvin. This scale is used mostly in scientific work where it can be applied to entities having extremely wide ranges in temperature. This scale is also called the Absolute Thermodynamic Temperature Scale and has for its zero a temperature so cold that it is assumed that no heat energy whatsoever exists in the material being measured. Temperatures lower than this do not exist (Halliday and Resnick, 1970). On the Kelvin scale, water boils at 373.16°K and freezes at 273.16°K .

The Kelvin and Celsius scales are related by the formulas (Leap, 2003):

$$T_C = T_K - 273.16 \quad (13.2a)$$

and

$$T_K = T_C + 273.16 \quad (13.2b)$$

where T_K is temperature in degrees Kelvin.

13.2.3 Heat Units: Calories and British Thermal Units

The actual quantity of heat energy in a solid, liquid, or gas is given in both scientific and engineering units. Scientists more often use the term, calorie (cal) and engineers the term, British Thermal Unit (BTU). One calorie is the quantity of added heat that will raise the temperature of one gram of water from 14.5 to 15.5°C. Likewise, a kilocalorie (Kcal) is 1000 calories and that amount of heat will raise the temperature of one kilogram of water by one degree between the same temperature ranges. It is of interest to note that a calorie as a measure of food energy is actually a kilocalorie.

One BTU is the amount of added heat that will raise the temperature of one pound of water from 63 to 64°F. The relation between BTU's and calories is given by

$$1.000 \text{ Kcal} = 1000 \text{ cal} = 3.968 \text{ BTU} \quad (13.3)$$

13.2.4 Thermal Conductivity

Like hydraulic conductivity that expresses the ease with which a porous medium will transmit water, thermal conductivity (k) expresses the ease with which a solid, liquid, or gas transmits heat. Both porous media and heat-conducting media are considered to be continua. Hydraulic conductivity has the dimension of L/T where L is length and T is time. Thermal conductivity has the dimension of Kcal/sec meter°C.

13.2.5 Thermal Resistivity

Just as hydraulic resistivity is the inverse of hydraulic conductivity (K),

$$R_H = 1/K \quad (13.4)$$

the inverse of thermal conductivity is thermal resistivity, that is,

$$R_T = 1/k \quad (13.5)$$

13.2.6 Specific Heat of Water

The heat capacity of a substance is the quantity of heat (Q) added to the substance per unit temperature rise (ΔT), that is,

$$c = \text{heat capacity/mass} = \Delta Q/m\Delta T \quad (13.6)$$

The specific heat of water is 1 cal/g which is very high when compared with other liquids (Table 13.1).

TABLE 13.1 Comparison of Specific Heats of Various Substances.

Substance	Specific Heat (cal/g/°C)
Water	1.00
Alcohol	0.58
Sugar	0.27
Aluminum	0.21
Marble	0.21

Adapted from Stowe, 1996, *Exploring Ocean Science*, Wiley, Inc., New York.

13.2.7 Three-Dimensional Hexagonal Structure of Water

Water, H₂O, exists in the fluid state as a three-dimensional (3-D) lattice of molecules (Stowe, 1996). As water cools, the structure becomes more ordered and the molecules are closer together so that at the temperature of 4°C, water not yet frozen is at its most dense state of 1 g/cm³. As water becomes even colder and then freezes at 0°C, the 3D lattices expand slightly to accommodate a more orderly 3D lattice for the solid state. As a result, the ice is less dense than very cold not-yet-frozen water.

In the near-frozen state a little above 4°C, all the hexagons are not ordered, but in the frozen state, they are all locked together. As ice freezes and the 3D hexagon structures expand, there is created an intramolecular “porosity” that is greater than that of the unfrozen state. This increase of porosity makes ice less dense than liquid water and this is the reason that ice floats in water; the density of water is 1.00 g/cm³ and that of ice is 0.92 g/cm³.

13.2.8 Thermal Inertia

The high thermal inertia of water means that it will not warm fast due to the high amount of heat energy required to warm it. Likewise, it will not cool as fast as other liquids.

This property causes water in the seas, large rivers, and large lakes to have a moderating effect on the climate of lands that border them. Similarly, cool water pumped into aquifers in winter months gains little heat so that it is still cool for use in air-conditioning systems in the summer. This is due to the facts that (1) water has high thermal inertia, and (2) earth materials are poor conductors of heat that help to keep stored groundwater cool.

13.2.9 Viscosity and Heat

Dynamic viscosity, also called absolute viscosity, is defined as the resistance to shearing force by flowing liquids. This concept is demonstrated by moving one of two parallel plates separated by a fluid. Dynamic viscosity, μ , is given by

$$\mu = \tau / (dv/dy) \quad (13.7)$$

where τ is the shearing stress, v is the velocity of the upper and movable plate, and y is the vertical distance between the two parallel plates. Kinematic viscosity, ν , is defined as the dynamic viscosity of a fluid divided by the fluid's density, ρ :

$$\nu = \mu / \rho \quad (13.8)$$

This term is used primarily in the flow of gases in which density may change with flow whereas dynamic viscosity is used primarily in the flow of liquids.

Dimensions of dynamic viscosity are lb sec/ft² and those of kinematic viscosity are ft²/sec. Viscosity decreases with increasing temperature.

13.2.10 Relationship Between Thermal Energy and Mechanical Energy (Joule's Experiments)

James Joule (1818–1889) was an English physicist who proved the relationship between heat and mechanical energy by a series of ingenious experiments. He devised a series of paddle wheels in a container of water that were rotated by falling weights of known mass connected to pulleys that caused the paddle wheels to rotate when the weights were dropped. By measuring the temperature of the water before and after the weights fell and turned the paddle wheels, he was able to show the mechanical energy of the falling weights in terms of foot-pounds to the increase of temperature of the water. His experimental error was only 5% — remarkable for his day.

Today, scientists agree that 1 Kcal = 1000 cal = 4186 J, and 4186 J of mechanical energy will raise the temperature of 1 Kg of water from 14.5 to 15.5°C, the same amount raised by 1000 calories of heat (Halliday and Resnick, 1970). Therefore, water flowing in a surface stream or through porous media will be heated by the mechanical action of flow.

13.3 Flow of Heat

13.3.1 Thermal Gradients

The thermal gradient, dT/dx , is the rate of change of temperature with distance through a substance and k is the thermal conductivity or the ease with which the substance transmits heat; this term is discussed below.

From Equation 13.6, one can readily understand the similarities between the mathematical description of heat flow through a continuum and that of water through a continuum by the following 1D unidirectional, steady-state flow of water through a porous medium

$$dQ/dt = KA dH/dx \quad (13.9)$$

where dQ/dt is the volume of water flowing per unit time perpendicular to a unit area, A , H is the hydraulic head, K is the hydraulic conductivity, and dH/dx is the hydraulic gradient. Similarities also exist for transient flow and non-uniform flow of heat and water.

13.3.2 Heat Flux

Heat flux is simply the amount of calories of heat energy flowing perpendicular to a unit area that is perpendicular to the direction of heat flow.

13.3.3 Conduction of Heat

Heat is conducted by different substances at different rates depending upon the thermal conductivity of the substance and the thermal gradient through it. Rocks and soil are poor conductors of heat whereas, water conducts heat much faster. This is the reason the interior of Earth is so hot — rocks do not conduct heat fast enough to make it cooler.

13.3.4 Convection of Heat

As opposed to conduction of heat, convection of heat is heat transport by the actual movement of heated fluids such as gases and liquids. In this process, unlike conduction, heat can be transported

from a region of lower temperature to a region of higher temperature if the heated fluid flow is in that direction. Conduction of heat only takes place from a region of higher temperature to a region of lower temperature.

Modern geothermal heating systems employ convection to pump naturally warmer air or water from the atmosphere or groundwater reservoirs respectively into heating systems within buildings. Such systems are known as “heat-pump systems.”

13.3.5 Advection versus Convection

The terms, *advection* and *convection* are often confused and interchanged. “Convection” should be applied to movement of heat as just described and “advection” should be applied to movement of groundwater through porous media and transport of dissolved chemicals and contaminants within it.

13.3.6 Diffusion of Heat

Historically, *diffusion* referred to conduction of heat through a continuum such as a solid or liquid. The work and terminology of the French scientist, John Joseph Fourier (born 1768) carried through to the time of Theis (1935) when he developed the similarity solution (Theis solution) to the non-equilibrium “diffusion equation,” earlier called the Fourier Equation for the diffusion of heat.

Therefore, the difference between the “diffusion” of heat and the “conduction” of heat must be addressed here. In his 1807 work, Fourier showed that heat flux, q , should be proportional to the temperature gradient, dT/dx . $Q = -kdT/dx$ and k is “thermal conductivity” (Cussler, 1984). This relationship is closely parallel to Fick’s First Law of chemical diffusion (mass transport)

$$F = -D dC/dx \quad (13.10)$$

where F is the mass flux of a solute per unit area per unit time, D is the diffusion coefficient for the particular solute (area/time), C is the solute concentration, and dC/dx is the concentration gradient (mass/volume/distancce).

If concentration is changing with time, Fick’s second law is used:

$$\partial C/\partial t = D(\partial^2 C/\partial x^2) \quad (13.11)$$

and thus, it is apparent why early groundwater flow equations were called “diffusion equations.”

Today in the science of hydrogeology, including groundwater flow, the term “diffusion” is generally reserved for the actual molecular diffusion of solute along a concentration gradient. Conduction or conductance is generally applied to heat transport under a thermal gradient and “mass transport” is understood to mean the movement of solutes by both advection along a hydraulic gradient and molecular diffusion along a concentration gradient.

13.4 Sources of Heat Flow to and from Earth’s Surface

13.4.1 Solar Radiation

Heat energy from the Sun heats Earth by solar radiation. This radiation includes all wavelengths in the electromagnetic spectrum from long radio waves of 3 m to gamma rays of 10^{-12} m. The human eye can detect only radiation in the visible spectrum (i.e., white light).

Not all solar radiation reaching Earth’s atmosphere actually warms the surface. Only 50% reaches the surface and is absorbed by it. The remaining 50% is lost by backscattering into space, reflection by

TABLE 13.2 Concentric Layers of the Earth

Layer	Depth (Km)	Density Range (g/cm ³)	Average Density
Crust	0–30	2.5–3.3	2.8
Mantle	30–290	3.3–5.7	4.5
Outer Core	2900–5100	9.4–14.2	11.8
Inner Core	5100–6370	16.8–17.2	7.0

Source: From Leap, 2003, *Natural Sciences, A Study Guide*. The College Network, Indianapolis, IN.

clouds into space, reflection from Earth's surface, and absorption by the atmosphere and clouds (Leap, 2003).

Heat from solar radiation has a minor effect on groundwater. In recharge areas of permeable materials, rain during warm weather can add its heat to water being recharged from the surface to groundwater systems below. In the winter months, recharge is colder. It is for these reasons that near-surface groundwater is warmer during summer than in winter.

Cold recharge water carries greater concentrations of dissolved gases such as oxygen and carbon dioxide. The greater amount of oxygen causes greater oxidation of organic material in the soil and also of metals in mineral matter such as iron and manganese. The carbon dioxide causes increases solution of carbonate minerals.

13.4.2 Geothermal Gradient and Heat Flow

The Earth is composed of four layers as shown in Table 13.2.

The inner core is believed to be solid and the outer core, liquid. The mantle is plastic and deforms from the action of rising heat with consequent convection currents in the mantle. The crust is mainly solid except for a few isolated volcanic regions and perhaps some isolated pockets of magma that extend into the lower crust from the mantle.

The high densities of both the inner core and outer core and the increasing density from the surface to the core suggest that heavier elements such as mostly iron and nickel migrated toward the center of Earth while it was in a molten state. Along with these heavy elements, even heavier uranium also migrated downward.

The heat generated within Earth comes from (a) residual heat left over from Earth's formation, (b) heat from increasing pressure with depth, and most importantly, (c) heat from fission of radioactive elements such as uranium. Rocks have relatively low thermal conductivity and Earth releases its heat slowly. Nevertheless, it is this internal heat that causes rocks to melt in the mantle and causes convection currents in the mantle that then bring heat nearer to the surface to be released in localized areas from terrestrial and marine volcanoes and mid-ocean rifts. It is also Earth's internal heat that is the greatest source of heat in groundwater.

From the interior of Earth to its surface there is a geothermal gradient as heat is conducted to the surface. In the United States, the geothermal gradient varies from about 12 to 47°C (53 to 116°F) per kilometer (Keller, 2002). This gradient causes deeper groundwater in the United States to have an ambient year-round average temperature of about 13°C east of the Rocky Mountains. The steeper the geothermal gradient, the closer the hot rock is to the surface and the greater the flux of heat to the surface.

East of the Rocky Mountains the flux of heat to the surface is approximately 50 mW/m². However, in the Rocky Mountains and the coastal ranges of California, Oregon, and Washington heat flux ranges from 40 to 60 mW/m². Throughout most of the West, the flux varies from 60 to 100 mW/m² and in certain places, namely northern Nevada, the Cascade Mountains, the Geysers north of San Francisco, and several other isolated spots, the flux is greater than 100 mW/m² (Duffield et al., 1994).

13.4.3 Hydrothermal Activity

Where large fluxes of heat occur, groundwater is definitely warmer than in low-flux areas and often very hot. Heat is transmitted to aquifers by (1) direct conduction through rocks, especially water-saturated rocks from magma bodies at depth, and (2) by recharging surface water that comes into contact with hot rocks and is heated. The latter process is most common in geysers of hot water and steam such as those in Yellowstone National Park in Wyoming and elsewhere.

Hot springs occur in various parts of the United States, even where active volcanic emissions are long past. Some of the heat driving these springs is from past volcanic activity in areas where there is magma closer to the surface than in most areas.

The greatest number of hot springs and geysers are to be found in the western United States where the crust is geologically much younger than in the east and where volcanic activity occurs on a greater scale than in the east.

13.4.4 Geothermal Energy for Electric Power

This source of energy is utilized in areas where groundwater is heated to boiling or near-boiling by coming into contact with magma or with rocks heated by magma due to its close proximity.

The ancient Romans heated their baths and homes with water from hot springs and such uses are found today in Japan, Turkey, and Iceland. Geothermal energy for producing electrical power was first used in Italy in 1904. Electric power plants utilizing geothermal energy are still used today in the United States, New Zealand, Iceland, Mexico, and Japan (Encyclopedia Britanica, 2002).

Keller (2002) and Domenico and Schwartz (1990) discuss two ways in which geothermal energy is utilized to generate electric power with (1) vapor-dominated systems and (2) hot-water dominated systems. In vapor-dominated systems wells are drilled into aquifers sitting above hot magma in which the water is so hot that steam covers the water and is trapped between the water and an overlying layer of low-conductivity material. Extraction wells do not penetrate the hot water, but rather, terminate in the steam zone. The steam is extracted to power the turbines that in turn, spin the generators.

Spent steam is then condensed and the resulting water is then reinjected into the saturated part of the aquifer to be reheated so that new steam may be added to that already in the groundwater system. The hot-water aquifer is recharged naturally by the groundwater having a high hydraulic head in the recharge area.

In a hot-water dominated system, wells tap a hot-water aquifer above a magma body and separated from the magma by sedimentary or crystalline rocks. The hot water is brought to the surface where the pressure lower than in the aquifer, causes steam to escape from the water and the steam then drives turbines which drive generators.

The used water and condensed-steam water is reinjected into the hot groundwater system where it is recycled and reheated to be used again. The number of hot-water dominated systems is about 20 times the number of vapor-dominated systems. The latter kind are not common and have been identified only in the Geysers of California; Mt. Lassen National Park, California; and Yellowstone National Park of Wyoming.

In the United States, total power production from geothermal energy is about 3000 Mega Watts (MW) and more than 90% of this amount is generated in California. Approximately 2000 MW are generated at the Geysers Power Plant 145 Km north of San Francisco that has been in operation for many years (Keller, 2002). A typical geothermal well will produce 5 to 8 MW of power. Geothermal energy is only available as a site-specific source. At the present time, it supplies less than 1% of the electrical power produced in the United States.

Presently, fossil-fuel generated electrical power is less expensive to produce than that from geothermal sources. It is expected that geothermally-generated power is unlikely to exceed 10% of electrical power output in the near future in the United States. However, there are specific places in the United States where geothermal power can be of great utility.

Although the environmental impact of the development of geothermal energy is less troublesome than that of other forms of energy production, there are still environmental problems associated with

geothermal energy (Keller, 2002). Some impacts include gas emissions, on-site noise and land scars, thermal pollution from hot waste waters, and chemical pollution from saline and mineral-laden water.

Reinjecting spent-water back into hot-water aquifers might cause earthquakes. Finally, if too much hot water is withdrawn, it might in some cases cause subsidence of the surface. In spite of these possible, but occasional impacts, it is important to note that the geothermal plants produce less than 1% of nitrogen oxides and only 5% of carbon dioxide produced by coal-burning power plants for the same amount of energy produced (Duffield, et al., 1994). However, geothermal energy does not produce either the atmospheric particulate pollutants associated with use of fossil fuels or radioactive waste (Keller, 2002).

13.4.5 Annual Air-Temperature Changes

Domenico and Schwartz (1990) state that the thermal conductivity of air is only 0.006 cal/msec °C. This value is extremely low when compared with water that has a thermal conductivity of 0.11 cal/msec °C. Therefore, annual air-temperature changes have little or no effect upon groundwater temperatures in more temperate zones, but in high mountains and arctic regions, summer temperatures can cause melting of the upper layers of permafrost that can then re-freeze in colder months. This topic is discussed next.

13.4.6 Permafrost

Permafrost is perennially frozen earth often with included ice lenses that is found in polar latitudes and high mountains. In these areas, the mean annual temperature is low enough to make the ground temperature below 0°C (Fetter, 2001). Warm summer temperatures may cause the upper 1 to 2 m of the soil or rock to thaw and this is then known as the active layer. Below this layer the ground down to 400 m may be frozen and has likely been so since the Pleistocene Epoch (the “Ice Age”).

If such low temperatures and frozen conditions persist for two or more years, the frozen ground is then called permafrost (Cederstrom et al., 1953; Brandon, 1965; Williams, 1970).

At some depth, the maximum annual temperature is 0°C and this depth is called the permafrost table. The local depth of permafrost is dependent upon both the mean air temperature and the geothermal gradient. As a result, here may be some areas where zones of unfrozen ground exist sandwiched between the frozen layers.

Sloan and van Everdingen (1998) discussed in the following paragraphs, state that permafrost under the right temperature may contain a mixture of unfrozen water and ice, or it may contain only ice. The unfrozen water exists as thin films adsorbed on surfaces of mineral particles and the thicknesses of the films are functions of temperature (50 Å or more at 0°C; 9 Å at -5°C; and 3 Å at -193°C). It is nearly independent of the total water content (liquid plus solid), except in very dry soils.

Hydraulic conductivity in permafrost layers is a function of the temperature of the permafrost system, the thermal gradient and the available cross-sectional area of interconnected films of unfrozen water. Movement of groundwater is retarded in movement through permafrost, but permafrost acts more like a semi-confining bed of very low hydraulic conductivity and not like a completely confining bed of zero hydraulic conductivity. If the permafrost layers are fractured significantly, the higher overall hydraulic conductivity can be expected to be much higher than in unfractured permafrost.

Fetter (2001) describes groundwater below permafrost as confined. The permafrost acts as the confining layer and the confined groundwater can be artesian in nature. Saline water may occur below permafrost and is probably due to the fact that in areas where permafrost is continuous such as northern areas of Alaska, Canada, and Russia, the mean annual precipitation can range from 500 mm/yr to less than 100 mm/yr, thus restricting the amount of water available for recharge. Additionally, potential evapotranspiration exceeds precipitation over a large area of the permafrost region.

Fetter also reports that alluvial river valleys are good sources of groundwater in permafrost regions. These areas generally consist of sand and gravel and are generally unfrozen. Alluvium from large rivers will have more unfrozen ground that contains more available water than frozen ground along smaller tributaries and headwater streams.

Alluvial fans at the margins of mountains in Alaska composed of coarse sand and gravel are good sources of groundwater from below the permafrost close to rivers and lakes that may help to keep permafrost absent. Coarse-grained materials do not possess the large area of interconnected ice films on particles as found in small-grained deposits.

South-facing slopes that receive more sunlight than north-facing slopes in permafrost regions tend to be more productive for groundwater development.

13.4.7 Heat Exchange between Surface Water and Groundwater and Effects upon Aquatic Environments

Warm or hot groundwater discharging into a surface stream or lake by base flow can have both detrimental and beneficial effects on aquatic environments. This is especially obvious in geothermal areas like Yellowstone National Park where effluent from hot springs and geysers eventually empty into surface-water bodies. Certain species cannot live in excessively warm temperatures and others do not reproduce and grow well in waters of elevated temperatures; this includes both plant and animal species. However, some thermophilic bacteria such as *Thermus aquaticus* actually thrive in waters with temperatures near the boiling point such as those found in the hot springs of Yellowstone. Indeed the vibrant blue to red color bands observed within the parks, springs, and pools result from the presence of bacteria with varying temperature tolerances.

Rivers and other streams are probably less affected by warm groundwater or surface water discharge into them across the length of their reach as streams with their rather rapid flow rates can often either carry away pulses of warm water discharge or dilute them down to a lower temperature. At the point of discharge, however, warmer waters may be detrimental or beneficial depending on which types of biota are present within the stream.

Impacts are greatest in ponds and wetlands that are mostly nourished by groundwater. If this nourishing groundwater is too warm it can adversely affect or even kill wetland plants and aquatic animal species. It may also help some such as bacteria and algae (that "bloom" under warmer conditions).

For the past 35 yr, there has existed a pond in an abandoned gravel quarry on the campus of Purdue University. An impermeable bottom of clay has caused the pond to form. For many years, warm water from various water-cooled systems such as air conditioners and laboratory equipment has discharged into this pond so that the pond water is often at a temperature of 25°C. This warm water recharges the aquifer through the sides of the pond and the water then is captured by the campus water-supply wells with temperatures of 16 to 17°C, fully 5 to 6 degrees greater than the ambient groundwater temperature. This in turn diminishes the efficiency of water-cooled systems across the campus. The Indiana Department of Environmental Management will not allow the University to ditch the warm pond water to the nearby Wabash River for fear of its possible adverse impact on aquatic life in the river. Warm water from industrial plants, even though injected into the ground may find its way into surface-water bodies.

13.4.8 Effects of Groundwater-Transported Heat upon Hydrogeochemistry

In general, warmer groundwater will increase chemical activity and will cause greater dissolution of minerals in the rocks and soil through which it passes. This is especially true of minerals containing silica or silicates as well as others.

However, warm water will decrease solubility of other minerals containing carbonate as colder water will contain more dissolved carbon dioxide (CO_2) and will produce more carbonic acid (H_2CO_3) for dissolution of these minerals.

Warmer waters may actually result in degassing of CO_2 which in turn can cause mineral precipitation. If warmer groundwater moves toward the Earth's surface from great depth, as in volcanic or geothermal areas, it can cause precipitation of valuable mineral and ore deposits. Heated groundwater, dissolving minerals such as silicates, can increase both porosity and hydraulic conductivity in rocks.

Especially near convergent boundaries of tectonic plates, the high pressures and temperatures that cause melting of rocks also cause metals to be released from rocks that are partially molten. Hot water carries dissolved metal salts to host rocks where ores are deposited.

13.5 Applications

13.5.1 Using Surface Heat Distributions to Locate Groundwater Resources

Earth scientists have known for many years that groundwater modifies the heat flux coming to Earth's surface along the normal geothermal gradient. As a result, different researchers have developed governing equations and groundwater exploration methods to actually explain the effects of groundwater flow and geothermal heat flow and to explore for groundwater resources at depth.

Van Orstrand (1934) wrote that moving groundwater could transfer heat and thereby cause variations of temperature gradients. The natural heat flux to Earth's surface in the Midwest and most of the eastern United States is generally small, averaging less than 1 mW/m^2 at the surface. Bredehoeft and Papadopoulos (1965) stated that as the natural heat-flux density from Earth is usually small, the upward thermal gradient within Earth may be affected by the groundwater movement.

Basic equations for simultaneous transfer of heat and groundwater were derived by Stallman (1960). He also suggested that groundwater temperature measurements might be used to measure rates of groundwater movement.

Cartwright (1974a, 1974b) used soil temperatures to locate groundwater supplies beneath the areas tested for temperature. Generally, cooler soil temperatures indicated the presence of groundwater. This relationship between temperature and groundwater in areas of low geothermal gradient and low surface heat flux can be attributed to a combination of high specific heat and thermal inertia of water and low heat conductivity of rocks and soil. Therefore, surface water recharged to groundwater in cooler weather will not heat fast and will stand out by its low temperature. By the same principles, modern infra-red remote sensing from aircraft can locate cooler areas that lie over groundwater supplies.

13.5.2 Using Temperature/Depth Profiles in Water Wells to Determine Vertical Velocities of Groundwater Movement

Knowing the vertical velocity of groundwater from the surface downward is extremely important in estimating recharge rates from the surface and for estimating the velocity of downward migration of contaminants from the surface. Commonly, vertical velocity is determined with environmental tracers such as atmospheric tritium (Daniels et al., 1991). In these procedures, cores of soil are taken at various depths by core drilling and the water is extracted and analyzed for tritium that was recharged to the soil from the atmosphere during rain or snowmelt. An alternative method is to drill several holes in a cluster, each tapping a certain depth and then extracting water from each for tritium analysis.

In either case, the tritium collected at each sampling depth is then used to determine how far the water has traveled vertically by comparing its concentration with atmospheric concentration at recharge using the half-life of the isotope. In either method, considerable time and expense is necessary for drilling, coring, and analysis.

A much faster, more convenient, and less expensive method is determination of the vertical temperature profiles in the water columns of existing wells. The change of temperature with depth of the water column can then be used to estimate the vertical recharge rates alongside the wells. Bredehoeft and Papadopoulos (1965) derived type curves for estimating the flux of groundwater in one dimension by fitting to the type curves temperature data found in the wells with depth.

Later work by Stallman (1967) and Sorey (1971) also indicated that curvature in Earth's thermal profile can be caused by vertically moving groundwater transporting heat by advection. Other investigators have used groundwater temperatures to determine directions of groundwater flow; these include Norris and Spieker (1962a, 1962b), Winslow (1962), Schneider (1962a, 1962b, 1964, 1972), and Supkow (1970).

Taniguchi (1993) states that the groundwater fluxes estimated with type curves agreed well with estimates from hydraulic data and basin water-balance computations.

Arriaga (2002) in his research project for a Master of Science degree in the Department of Earth and Atmospheric Sciences at Purdue University used vertical temperature variations in existing wells around the Purdue University campus in West Lafayette, Indiana to determine the vertical velocity of groundwater in various locations. This research was motivated by the continuing expansion of building activities on the campus and the need to understand better the vertical recharge rates to aquifers underlying the campus, especially in the complex glacial stratigraphy with alternating aquifers and aquitards.

Arriaga (2002) measured vertical temperature profiles in 12 wells between September 2001 and April 2002. Temperatures were measured at 5-ft intervals and in some cases, 2-ft, and 1-ft intervals. The 12 wells provided continuous temperature information for the entire depth of the well and in some cases, for the entire saturated thickness of aquifers and aquitards.

The instrument used was a REELOGGERTM, a temperature probe on the end of a cable that registered well-water temperatures with a sensitivity of $\pm 0.02^\circ\text{C}$. As the probe was lowered, the water temperatures it sensed were recorded digitally on the surface for later processing by computer. Temperatures were taken in thermally-stable water columns of wells at depths opposite both aquifers and aquitards.

The flow of heat and groundwater in one dimension (vertical), assuming steady state is

$$(\partial^2 T / \partial z^2) - (c_w p_w v_z / k) (\partial T / \partial Z) = 0 \quad (13.12)$$

This is the vertical direction component of Stallman's (1960) equation for 3-D simultaneous transient flow of both heat and fluid in a fully-saturated, isotropic, and homogeneous saturated porous medium.

The analytical solution by Bredehoeft and Papadopoulos (1965) assumes an isothermal flow through aquitards with boundary conditions of

$$T_z = T_0 \quad \text{at } z = 0 \quad \text{and} \quad T_z = T_L \quad \text{at } z = L$$

where T_z is the temperature at any depth, T_0 is at $z = 0$ (uppermost temperature measurement), T_L is the temperature at $z = L$ (lowermost temperature measurement), L is the vertical distance over which temperatures are being observed, and v_z is vertical groundwater velocity.

Their solution is

$$(T_z - T_0) / (T_L - T_0) = f(\beta, z/L) = (e^{\beta(z/L)} - 1) / (e^\beta - 1) \quad (13.13)$$

where

$$\beta = c_w p_w v_z L / k \quad (13.14)$$

and is a dimensionless parameter that is positive or negative depending on whether v_z is downward or upward. Bredehoeft and Papadopoulos provide values of $f(\beta, z/L)$ for a range of parameters. They provide a set of type curves for different values of β .

The ratios $(T_z - T_0) / (T_L - T_0)$ are calculated from the measured temperature data and are plotted against the depth factor, z/L at the same scale as the type curve.

The value of β is dimensionless and is related to the curvature of a temperature-depth profile at several depths ranging from $z = 0$ to $z = L$. The thickness, L is less than or equal to the thickness of the semi-confining layer (water-saturated clay). It is through this layer that groundwater is leaking vertically with a velocity of v_z . Positive and negative values of β represent downward and upward leakage, respectively.

Once the value of β is known, the vertical groundwater velocity can be obtained by

$$v_z = k\beta / c_w p_w L \quad (13.15)$$

Typically, the value for thermal conductivity of water-saturated clay is roughly $k = 2 \times 10^{-3}$ cal/cm sec °C (Birch et al., 1942). Thermal conductivity of water-saturated sand and gravel outwash is $k = 4 \times 10^{-3}$ cal/cm sec °C. In water-saturated silts it averages $k = 2 \times 10^{-3}$ cal/cm sec °C (Wade, 1994).

Results of borehole water temperature measurements in the 12 wells showed the vertical velocities ranging from 1.71 to 1.86 cm/yr through aquitards and from 0.92 to 4.53 cm/yr in aquifers. In general, it was found that the aquifers with high horizontal conductivity had low vertical velocity components. In many aquitards the vertical velocities were greater than in aquifers as, according to the “Law of Tangents” (Fetter, 2001) and the continuity principle, water must take a shorter route through less permeable materials than through more permeable materials, and this necessitates faster vertical movement through an aquitard

$$K_1/K_2 = \tan \Theta_1 / \tan \Theta_2 \quad (13.16)$$

where K_1 is the hydraulic conductivity in layer 1, K_2 is the hydraulic conductivity in layer 2, Θ_1 is the angle made between impinging flow line and surface of layer 1, and Θ_2 is the angle made between the impinging flow line and surface of layer 2.

13.5.3 Distinguishing Different Aquifers by Temperature Profiles in Boreholes

In multiaquifer systems, some aquifers may be recharged from different areas, including melting surface snow or ice fields, or some may be deeper and warmer due to the geothermal gradient and greater heat flux into the aquifer. Such aquifers in many cases may be delineated from each other by their water temperatures if the temperatures are significantly different from one aquifer to another.

13.5.4 Expansion of Formation Water in Deep Boreholes

Often when a deep borehole (1000 to 2000 ft) is drilled into a deep aquifer and then left for some time and eventually pumped, the water level in the well will recover to an elevation slightly higher than the static water level before pumping. This behavior is due to the fact that before pumping, the water temperature equilibrates to a temperature lower than the temperature of the water in the aquifer and thus, the elevation of the water surface in the well is higher due to thermal expansion of the water during recovery after pumping.

13.5.5 Storage of Both Hot and Cold Water in Geological Formations

Cold water is stored in the sands and gravels of an abandoned channel of the Ohio River near Louisville, Kentucky. Recharged in winter months, this water is stored until warm weather periods and then used for air-conditioning purposes. Likewise, warm or hot water can be stored in permeable and porous formation for later extraction.

Storage of either hot or cold water in this manner can present some problems with temperature of the water or with chemical reactions between solutes in the water and minerals in the sediments into which it is injected. For example, hot water can speed up dissolution or precipitation of certain minerals. Cold water may cause dissolution of carbonate minerals. Precipitation of minerals can cause clogging of pores. Geological storage of either hot or cold water is possible as rocks have low thermal conductivities.

Glossary

British Thermal Unit (BTU) The quantity of heat that will raise the temperature of one pound of water by one degree Fahrenheit.

Calorie The quantity of heat that will raise the temperature of one cubic centimeter of water by one degree Celsius.

Celsius The temperature scale developed by Celsius and used mostly outside of the United States.

Continuum A continuous unbroken medium through which water or heat passes.

Diffusion Molecular diffusion is the spreading of the molecules of a solute in a liquid from a region of higher concentration to a lower concentration; heat diffusion is the spreading of heat through a continuum from a region of higher temperature to one of lower temperature.

Dynamic or Absolute Viscosity In liquids it is the resistance to shear as the liquid moves.

Energy The ability to produce change (e.g., work).

Fahrenheit The temperature scale developed by Fahrenheit and used mostly in English-speaking countries except England.

Geyser A particular tubular vent in areas of high geothermal heat that periodically gushes forth water and steam.

Heat Conduction The movement of heat through a continuum from molecule to molecule of the continuum simply under a thermal gradient.

Heat Content The actual number of calories in a sample of a material at any moment.

Heat Convection Movement of a mass of heat in a fluid (water or air) into an area that may be warmer or cooler than the fluid by forced fluid flow (e.g., by a fan).

Heat Flux The amount of heat moving perpendicular to a unit area per unit time.

Heat Pump A mechanical device that moves warm or cold air or water into a space for heating or cooling purposes respectively.

Hot-Water Dominated System A geothermal reservoir such as an aquifer, containing hot water completely filling the aquifer from top to bottom.

Hot Springs Springs from which issues hot water instead of water of ambient temperature.

Kelvin Scale A temperature scale developed by Lord Kelvin on which the lowest temperature possible, zero degree, is absolute zero at which there is no heat whatsoever in the system in question.

Kinematic Viscosity Used mostly in the study of flow of gases, it is the dynamic or absolute viscosity divided by the fluid density.

Non-Equilibrium Heat Flow Flow of heat down a thermal gradient such that it does not reach steady-state flow, but the flux changes with time and depletion of heat from storage.

Permafrost Ground containing water that is always frozen in very cold latitudes.

Potential Energy Energy that is available to produce change but is not being used.

Specific Heat The unique specific number of calories that will raise the temperature of a particular substance by one degree Celsius.

Temperature The measure of heat energy in a particular substance.

Thermal Conductivity The ease with which a substance will transmit heat.

Thermal Inertia The property that causes time delay in the warming or cooling of a fluid; water has a high thermal inertia because of its high specific heat.

Thermal Resistivity The inverse of the thermal conductivity.

Vapor-Dominated System A geothermal reservoir such as an aquifer in which steam sits between the top of the hot water and the bottom of the upper confining layer.

References

- Arriaga, M.A. 2002. Finding vertical groundwater velocities using temperature-depth profiles in West Lafayette, Indiana. *M.S. Thesis, Dept. of Earth and Atmospheric Sciences, Purdue University, West Lafayette, Indiana.*
- Birch, A.F., Shairer, J.F., and Spicer (eds). 1942. Handbook of physical constants. Geological Society of America, Special Paper.
- Brandon, L.V. 1965. Evidences of groundwater flow in permafrost regions. *Proceedings of the 1963 International Conference on Permafrost (Lafayette, Indiana)*, pp. 176–177, Pundue University, West Lafayette, IN.
- Bredehoeft, J.D. and Papadopoulos, I. 1965. Rates of vertical groundwater movement estimated from Earth's thermal profile. *Water Resources Research*, 1(2), 325–328.

- Carslaw, H.S. and Jaeger, J.C. 1959. *Conduction of Heat in Solids*, 2nd ed. Oxford Science Publications, Clarendon Press, Oxford, England.
- Cartwright, K. 1974a. Groundwater discharge in the Illinois Basin as suggested by temperature anomalies. *Water Resources Research*, 6(3), 912–918.
- Cartwright, K. 1974b. Tracing shallow groundwater systems by soil temperature. *Water Resources Research*, 10(4), 847–855.
- Cederstrom, D.J., Johnson, P.M., and Subitzky, S. 1953. *Occurrence and Development of Groundwater in Permafrost Regions*. U.S. Geological Survey Circular 275.
- Cussler, E.L. 1984. *Diffusion, Mass Transfer in Fluid Systems*, Cambridge University Press, Cambridge, England.
- Daniels, D.P., Fritz, S.J., and Leap, D.I. 1991. Estimating recharge rates through unsaturated glacial till by tritium tracing. *Groundwater*, 29(1), 26–34.
- Domenico, P.A. and Schwartz, F.W. 1990. *Physical and Chemical Hydrogeology*. John Wiley & Sons, New York.
- Duffield, W.A., Sass, J.H., and Sorey, M.L. 1994. *Tapping Earth's Natural Heat*. U.S. Geological Survey Circular 1125.
- Encyclopedia Britannica, 2002.
- Fetter, C.W. 2001. *Applied Hydrogeology*, 4th ed. Prentice-Hall, Inc., Upper Saddle River, NJ.
- Halliday, D. and Resnick, R. 1970. *Fundamentals of Physics*. John Wiley & Sons, Inc., New York.
- Keller, E.A. 2002. *Environmental Geology*, 2nd ed. Prentice-Hall Inc., Upper Saddle River, NJ.
- Leap, D.I. 2003. *Natural Sciences, A Study Guide*. The College Network, Indianapolis, IN.
- Norris, S.E. and Spieker, A.M. 1962a. Seasonal temperature changes in wells as indicators of semi-confining beds in valley-train aquifers. *U.S. Geological Survey Professional Paper*, 450-B, B101–B102.
- Norris, S.E. and Spieker, A.M. 1962b. Seasonal temperature changes in wells as indicators of semi-confining beds in valley-train aquifers. *U.S. Geological Survey Professional Paper*, 450-B, B101–B102.
- Ozisik, M.N. 1968. *Boundary Value Problems of Heat Conduction*. International Textbook Company, Scranton, PA.
- Schneider, R. 1962a. An application to the study of groundwater. *U.S. Geological Survey Water Supply Paper*, 1544-B, 1–16.
- Schneider, R. 1962b. Use of thermometry in hydrogeologic studies of glacial deposits at Worthington, Minnesota. *U.S. Geological Survey Water Supply Paper*, 73(10), 1305–1308.
- Schneider, R. 1964. Relation of temperature distribution to groundwater movement in carbonate rocks of central Israel. *Geological Society of America Bulletin*, 75(3), 209–216.
- Schneider, R. 1972. Distortion of the geothermal field in aquifers by pumping. *U.S. Geological Survey Professional Paper*, 800-C, C267–C270.
- Sloan, C.E. and van Everdingen, R.O. 1998. Chap. 31, Region 28, *Permafrost Region in The Geology of North America*, vol. O-2, Hydrogeology, The Geological Society of America.
- Sorey, M.L. 1971. Measurement of vertical groundwater velocity from temperature profiles in wells. *Water Resources Research*, 7(4), 963–970.
- Stallman, R.W. 1960. Notes on the use of temperature data for computing groundwater velocity. *Groundwater Notes* 39, U.S. Geological Survey. Washington, DC.
- Stallman, R.W. 1967. Flow in the zone of aeration. *Advances in Hydroscience*, Chow, V.T., ed. Academic Press, New York.
- Stowe, K. 1996. *Exploring Ocean Science*, 2nd ed. Wiley, Inc., New York.
- Supkow, D.J. 1970. Subsurface heat flow as a means for determining aquifer characteristics in the Tucson Basin, Pima County, Arizona. *Ph.D. Thesis, University of Arizona, Tucson*.
- Taniguchi, M. 1993. Evaluation of vertical groundwater fluxes and thermal properties of aquifers based on transient temperature-depth profiles. *Water Resources Research*, 29(7), 2021–2026.
- Theis, C.V. 1935. The lowering of the piezometric surface and the rate and discharge of a well using ground-water storage. *Transactions, American Geophysical Union*, 16, 519–524.
- Thiem, G. 1906. Hydrologische Methoden. *Gebhardt*, Leipzig.

- Van Orstrand, C.E. 1934. Temperature gradients in Problems of Petroleum Geology. *American Association of Petroleum Geologists*, 989–1021.
- Wade, S. 1994. A hydrothermal study to estimate vertical groundwater flow in the Canutillo Well Field between Las Cruces and El Paso, New Mexico. *M.S. Thesis, New Mexico Water Resources Research Institute*.
- Williams, J.R. 1970. Groundwater in the permafrost regions of Alaska. *U.S. Geological Survey Professional Paper*, 696.
- Wilson, R.E. and Arnold, R.D. 1958. Thermometry and pyrometry. *Handbook of Physics*, Chap. 3. Condon, E.U. and Odishaw, H., eds.
- Winslow, J.D. 1962. Effect of stream infiltration on groundwater temperatures near Schnectady, New York. *U.S. Geological Survey Professional Paper*, 450-C, C125–C128.

Further Information

In addition to the sources listed in the References, major journals in hydrogeology can be consulted such as:

Water Resources Research, American Geophysical Union, Washington, DC.

Groundwater, National Association of Groundwater Scientist and Engineers, Dublin, Ohio.

Hydrogeology Journal, co-sponsored by the Geological Society of America in Boulder, Colorado and the International Association of Hydrogeologists in Hanover, Germany.

Hydrological Science and Technology, American Institute of Hydrology, Smyrna, Georgia.

Bulletin of the Geological Society of America and Geology, both sponsored by the GSA, often contain papers on hydrogeology.

Special publication by the Geological Society of America and the American Geophysical Union are very good sources for specialized topics. Special publications by the national Geological Surveys of the United States, United Kingdom, Canada and many other countries are also good sources of information. State and Provincial Geological Surveys are ready sources of information.

14

Hydrogeological Characterization Using Geophysical Methods

14.1	Introduction.....	14-1
14.2	Common Geophysical Methods for Hydrogeological Characterization	14-3
	DC Resistivity Methods • Electromagnetic Induction • Ground Penetrating Radar • Seismic Methods • Seismic Petrophysics • Potential Field Methods • Borehole Geophysical Methods	
14.3	Hydrogeophysical Concepts.....	14-20
	Hydrogeological Mapping • Hydrogeological Parameter Estimation • Hydrological Process Monitoring	
14.4	Stochastic Forward and Inverse Modeling Using Hydrogeophysical Datasets.....	14-29
	Forward Modeling with Parameter Error • Inverse Modeling • Bayesian Estimators • Point Estimation with Known Parameters • Model Selection • Summary of Stochastic Techniques for Hydrogeophysical Investigations	
14.5	Summary	14-45
	Glossary.....	14-45
	Acknowledgments.....	14-46
	References	14-46
	Further Information	14-52

Susan S. Hubbard
*Lawrence Berkeley National
Laboratory*

Yoram Rubin
University of California

14.1 Introduction¹

The shallow subsurface of the earth is an extremely important geological zone, one that yields much of our water resources, supports our agriculture and ecosystems, and influences our climate. This zone also serves as the repository for most of our municipal, industrial, and governmental

¹Portions of this chapter are based on Hubbard and Rubin (2005) and Rubin and Hubbard (2005).

wastes and contaminants, intentional or otherwise. Safe and effective management of our natural resources is a major societal challenge. Contaminants associated with industrial, agricultural, and defense activities in developed countries, the increased use of chemical pollutants resulting from the technological development of countries with evolving market economies, the increased need to develop sustainable water resources for growing populations, and the threat of climate and land use change on ecosystems all contribute to the urgency associated with improving our understanding of the shallow subsurface.

It is widely recognized that natural heterogeneity and the large spatial variability of hydraulic parameters control infiltration within the vadose zone, groundwater storage, and the spread of contaminants. It is also well recognized that natural heterogeneity is typically great and can have multiple spatial scales of variability. Conventional sampling techniques for characterizing or monitoring the shallow subsurface typically involve collecting soil samples or drilling a borehole and acquiring hydrological measurements or borehole geophysical logs. When the size of the study site is large relative to the scale of the hydrological heterogeneity, or when the hydrogeology is complex, data obtained at point locations or within a wellbore may not capture key information about field-scale heterogeneity; especially in the horizontal direction. Integration of spatially extensive geophysical data with direct borehole measurements holds promise for improved and minimally invasive characterization and monitoring of the subsurface over a variety of resolutions and spatial scales.

The field of hydrogeophysics has developed in recent years to explore the potential that these methods have for mapping subsurface features, estimating properties, and monitoring processes that are important to hydrological studies, such as those associated with water resources, contaminant transport, and ecological and climate investigations. Hydrogeophysical investigations strive to move beyond the mapping of geophysical “anomalies” to provide information that can be used, for example, as input to flow and transport models. Hydrogeophysics builds on previous experience and developments associated with the mining and petroleum industries, which in the past 50 years have relied heavily on geophysical data for exploration of ore and hydrocarbons, respectively. As geophysics has been used as a tool in these industries for so long, there is a relatively good understanding of optimal data acquisition and reduction approaches as well as about petrophysical relationships associated with the consolidated, high pressure, and high temperature subsurface environments common to those industries. However, such subsurface conditions are quite different from the shallow, low temperature, low pressure, and less consolidated environments that typify most hydrogeological investigation sites. Indeed, the contrasts in shallow subsurface properties are often quite subtle compared to those of the deeper subsurface. Owing to the different subsurface conditions, the relationships that link geophysical attributes to subsurface parameters, as well as the geophysical attribute ranges, can vary dramatically between the different types of environments.

The goal of this chapter is to provide a description of practical tools for quantitative analysis of complex, heterogeneous aquifers using geophysical data and stochastic integration techniques. In Section 14.2, we review several geophysical techniques that are commonly used for hydrogeological investigation, and we provide examples of the geophysical information obtained using the methods. Section 14.3 describes hydrogeophysical concepts and approaches performed using geophysical datasets. This section gives examples of the use of geophysical data for mapping subsurface features, for estimating hydrogeological parameters, and for monitoring hydrological processes. When a detailed response analysis or confidence intervals for parameter estimation are needed, advanced stochastic techniques for integrating geophysical and hydrological data are often required. Section 14.4 develops a general framework for forward modeling using geophysical data under conditions of uncertainty by casting the hydrogeological prediction problem in a stochastic framework. This approach opens the door for estimating variables, such as permeability or concentration, in terms of their expected values and estimation variances (at a minimum) or, for a more complete characterization, through their probability distribution functions using integrated geophysical and hydrogeological datasets.

Our guiding principle throughout the chapter is the need for rational characterization approaches using geophysical data that recognizes the usual situation of data scarcity and subsurface heterogeneity.

14.2 Common Geophysical Methods for Hydrogeological Characterization

The purpose of this section is to review geophysical techniques that are most commonly used for hydrogeological characterization, including: electrical resistivity, electromagnetic, ground penetrating radar (GPR), seismic, potential field methods, and wellbore logging techniques. For each method, we provide a brief description of the underlying physical principles and instrumentation, common acquisition geometries, and general data reduction and interpretation methods. Also presented are several petrophysical relationships that may permit the transfer of geophysical measurements into estimates of hydrological parameters such as porosity and water content. These relationships are a topic of current research and thus are presented primarily to illustrate the potential that geophysical methods have for aiding subsurface characterization investigations, rather than to provide a recipe for geophysical-to-hydrogeological data transfer. We do not review other geophysical methods that offer great potential for hydrogeological characterization but that are not yet as extensively developed for such investigations, including induced polarization (IP), spontaneous potential (SP), nuclear magnetic resonance (NMR), and electroseismic methods.

We restrict our discussions to practical use and limitations of common geophysical methods; geophysical theory is beyond the scope of this book. However, references are given for each method that will allow readers to easily access more detailed information. For detailed information about acquisition, reduction, and interpretation of many geophysical methods, the reader is referred to Robinson and Coruh (1988), Burger (1992), and Telford et al. (1990).

14.2.1 DC Resistivity Methods

For groundwater studies, direct current (DC) resistivity, or electrical resistivity methods have been more frequently used than any other electrical method. Resistivity is a measure of the ability of electrical current to flow through materials; it is the inverse of electrical conductivity and is an intrinsic property of the material. Electrical resistivity is typically expressed in units of Ohm-meters ($\Omega\text{-m}$); resistivity values expressed Ohm-meters can be divided into 1000 to convert to conductivity values in milli-Siemans per meter (mS/m). Electrical resistivity methods involve the introduction of a time-varying DC or very low frequency ($<1 \text{ Hz}$) current into the ground between two current electrodes. In two-electrode systems, current flows from the positive to the negative current electrode. The currents can be envisioned to set up equipotential surfaces, with current flow lines running perpendicular to these surfaces. The fraction of current that penetrates to a particular depth is a function of the electrode spacing and the electrical resistivity distribution of subsurface materials (VanNostrand and Cook, 1966).

Most resistivity surveys utilize a four-electrode measurement approach. To obtain a value for subsurface resistivity, two potential electrodes are placed between the current electrodes, and the difference in potential or voltage (ΔV) is measured. This measurement, together with the known current (I) and a geometric factor (K), which is a function of the particular electrode configuration, can be used to calculate resistivity (ρ) for uniform subsurface conditions following Ohm's law:

$$\rho = \frac{\Delta V}{I} K \quad (14.1)$$

Equation 14.1 can also be applied when the resistivity of the ground is heterogeneous, but the more general term "apparent resistivity" (ρ_a) is used. The simplest four-electrode configuration is the Wenner array shown in Figure 14.1a, where the outer two current electrodes, labeled C1 and C2, apply a constant current, and the inner two potential electrodes, labeled P1 and P2, measure the voltage difference created by this current. The electrode spacing is a fixed value a , and the apparent resistivity of the heterogeneous

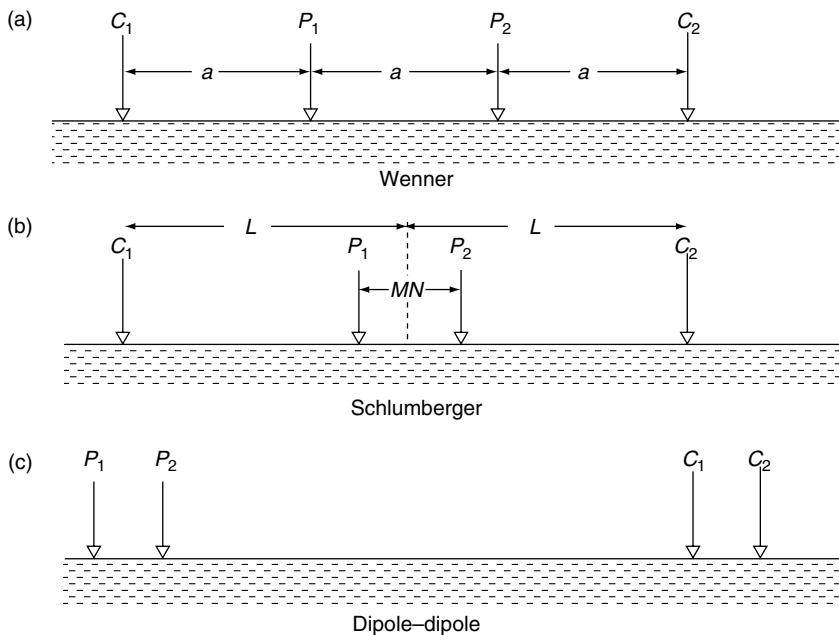


FIGURE 14.1 Common electrode configurations used to measure apparent resistivity of the subsurface. C1 and C2 are the current electrodes and P1 and P2 are the potential electrodes. (a) Wenner array. The outer two current electrodes apply a constant current, and the inner two potential electrodes measure the voltage difference created by the current. The electrodes maintain a constant spacing of “a.” (b) Schlumberger array. The spacing (MN) between the potential electrodes is much smaller than the spacing ($2L$) between the current electrodes. (c) Dipole–dipole array. The potential electrode pair functions independently from the current electrode pair, which enables a sampling in both the lateral and vertical direction of subsurface apparent resistivity. (From Burger, H. R., 1992, *Exploration Geophysics of the Shallow Subsurface*. Prentice Hall, Englewood Cliffs, NJ. With permission.)

subsurface sampled by this array can be computed using:

$$\rho_a = \frac{\Delta V}{I} 2\pi a \quad (14.2)$$

Another commonly used electrode configuration is the Schlumberger array (Figure 14.1b), where the spacing (MN) between the potential electrodes (P1, P2) is much smaller than the spacing ($2L$) between the current electrodes (C1, C2). In the dipole–dipole array, illustrated in Figure 14.1c, the current electrode pair and the potential electrode pair are closely spaced, and there is a significant distance between the sets. Unlike the Wenner and Schlumberger arrays, which collect data in either the *profiling* or the *sounding* mode depending on electrode array geometry, the potential and the current electrodes of the dipole–dipole array function independently of each other, permitting simultaneous profiling and sounding.

Profiling is undertaken by moving the entire array laterally along the surface by a fixed distance after each reading to obtain apparent resistivity measurement over a relatively constant depth as a function of distance. For each reading, the value of apparent resistivity is assigned to the geometric center of the electrode array. Profiling is useful for relatively rapid spatial coverage. Equipment to permit continuous electrical profiling has been used in Denmark to collect electrical measurements over large areas by pulling a string of electrodes with a vehicle (Sorensen, 1996); traverses as long as 10 km have been acquired in a single day using this towed-array approach. As profiles give lateral variations in electrical conductivity but not information about vertical distribution, the interpretation of profile data is generally qualitative, and the primary value of the data is to delineate gross geologic structure and stratigraphy.

Vertical electrical sounding (VES) curves give information about the vertical variations in electrical conductivity at a single location. Soundings are obtained by “expanding” the array along a straight line so that the spacing between the individual electrodes remains equal but takes on larger values after each measurement. The depth of the measurement is a function of the electrode spacing as well as the subsurface resistivity contrasts; as the electrode spacing is increased, the current penetrates to greater depths. Such curves have been used to delineate hydrogeological units as a function of depth. Sounding curves can be quite complex when the subsurface geology is not laterally homogeneous and when there are more than two subsurface layers and the interpretation of multilayer sounding curves is not unique. To interpret electrical resistivity sounding data, various curve-fitting or computer inversion schemes are used, or the measured data are compared with model computations (Zodhy et al., 1974).

Modern electrical equipment now includes multiplexing capabilities and automatic and autonomous computer acquisition, which greatly facilitate acquisition of a combined profiling and vertical soundings within acceptable timeframes. Such surface imaging is performed by acquiring profiles along transects using different electrode spacings. For each measurement, the apparent resistivity value is assigned to the midpoint of the four electrodes. By using a combination of electrodes, a pseudo-section is developed using the midpoint location technique. Modeling of electrical imaging data can then provide estimates of subsurface electrical resistivity values in their more “correct” subsurface positions. Surface imaging, which yields two- or three-dimensional information about subsurface electrical resistivity distribution, has become a particularly powerful geophysical technique for hydrogeological mapping, parameter estimation, and monitoring.

Electrical resistivity measurements are also collected between boreholes by inserting electrodes into a wellbore or installing electrodes on the outside of wellbore casing, collecting measurements, and processing of the measurements. Crosshole electrical imaging is often referred to as electrical resistance tomography (ERT), and ERT systems are now commercially available. With the recent development of advanced and automated acquisition systems, robust inversion routines and the capability of recording >3500 measurements per hour with only one technician, ERT methods have proven to be useful for dynamic process monitoring. For example, Ramirez et al. (1993) used ERT to map the subsurface distribution of a stream flood as a function of time, and Daily and Ramirez (1995) used ERT to monitor an *in situ* remediation process for removal of volatile organic compounds from subsurface water and soil. An excellent review of surface and crosshole electrical methods for hydrogeological applications is given by Binley and Kemna (2005), and several examples of the use of such methods for hydrogeological applications is given by Daniels et al. (2005).

As electrical resistivity measurements are influenced by lithology, pore fluid chemistry, and by water content, electrical measurements can be useful for hydrogeological mapping, property estimation, and monitoring as will be shown in Section 14.3. Electrical conduction usually takes place in fluids in connected pore spaces, along grain boundaries, and within fractures; the matrix of the material is generally considered nonconductive. Material texture, porosity, presence of clay minerals, moisture content, and the resistivity of the pore fluid all affect electrical resistivity. In general, all of the following reduce resistivity: increased water content, increased salinity of water, increased clay content, and decreased grain size (Burger, 1992). Assuming that water fills the voids, resistivity is also lowered by an increase in porosity, an increase in the number of fractures, and by increased weathering. Electrical resistivity values for common materials are shown in Table 14.1. From this table, it is evident that there are wide ranges of electrical resistivity for some materials and also that resistivity values of different materials overlap. Values commonly vary over 12 orders of magnitude and have a maximum range of 24 orders of magnitude (Telford et al., 1990). However, the following general statements can be made regarding electrical resistivity:

- Resistivity is sensitive to moisture content; thus unsaturated sediments usually have higher resistivity values than saturated sediments.
- Sandy materials generally have higher resistivity values than clayey materials.
- Granitic bedrock generally has a higher resistivity value than saturated sediments and frequently offers a large apparent resistivity contrast when overlain by these sediments.

TABLE 14.1 Approximate Electrical Resistivity and Dielectric Constant Values for Typical Near-Surface Materials

Material	Resistivity (Ohm-m)	Dielectric Constant
Sand (dry)	10^3 – 10^7	3–6
Sand (saturated)	10^2 – 10^4	20–30
Silts	10^2 – 10^3	5–30
Shales	10 – 10^3	5–15
Clays	1 – 10^3	5–40
Humid soil	50–100	30
Cultivated soil	200	15
Rocky soil	1000	7
Sandy soil (dry)	7100	3
Sandy soil (saturated)	150	25
Loamy soil (dry)	9100	3
Loamy soil (saturated)	500	19
Clayey soil (dry)	3700	2
Clayey soil (saturated)	20	15
Sandstone (saturated)	25	6
Limestone (dry)	10^9	7
Limestone (saturated)	40	4–8
Basalt	100	8
Granite	10^3 – 10^5	4–6
Coal	10^4	4–5
Fresh water	30 – 10^4	81
Permafrost	10^2 – 10^5	4–8
Dry snow	10^5 – 10^6	1
Ice	10^3 – 10^5	4–12

Source: Data from Davis, J. and Annan, A., 1989, *Geophys. Prospect.*, 37, 531–551; Telford, W. M., Gelard, L. P., and Sheriff, R. E., 1990, *Applied Geophysics*. Second Edition, Cambridge University Press, Cambridge; Ulriksen, P. F., 1982, Application of impulse radar to civil engineering, Ph.D. thesis, Lund University, Sweden.

Empirical formulae as well as theoretical models have been used to relate apparent resistivity measurements to hydrological parameters of interest. An excellent compendium of petrophysical relationships between electrical and hydrogeological properties is given by Lesmes and Friedman (2005). Archie's (1942) empirical relationship is perhaps the most widely used relationship to relate electrical conductivity to water-saturated hydrogeological properties. A modified version of Archie's relationship for variably saturated materials is given by:

$$\rho_{\text{eff}} = a\phi^{-m} S^{-n} \rho_w \quad (14.3)$$

where ρ_{eff} is the effective electrical resistivity of the sample, ρ_w is pore water resistivity, ϕ is fractional porosity, S is fractional water saturation, and n , a , and m are constants. The cementation index, m , ranges from approximately 1.3 for unconsolidated sands to 2 for consolidated sandstones; the saturation index, d , is about 2 for consolidated rocks and ranges from 1.3 to 2 for unconsolidated sands; and a ranges from 0.6 to 1 (Ward, 1990; Schon, 1996). Equation 14.3 is generally used for well log interpretation, but if ρ_{eff} , ρ_w , and ϕ can be independently measured, and m reasonably estimated, then the fractional water saturation can be found using electrical surveys (Ward, 1990). This concept was used by Pfeifer and Anderson (1995) to monitor the migration of tracer-spiked water through the subsurface using electrical measurements.

Site-specific petrophysical relationships have also been used to transform electrical measurements into estimates of hydrogeological properties or units. For example, site-specific petrophysical relationships have been developed between electrical measurements and hydraulic conductivity or transmissivity in porous aquifers (Kelly, 1977; Heigold et al., 1979; Kosinski and Kelly, 1981; Mazac et al., 1985; Huntley,

1986; Purvance and Andricevic, 2000). These studies suggest that the petrophysical relationships can be categorized into two groups. For coarse-textured, clay-free soils, permeability and electrical conductivity tend to be positively correlated, because they both increase with increasing porosity. With finer-textured soils, surface conductivity effects play a role, and permeability and electrical conductivity tend to be inversely correlated. Although the development of petrophysical relationships for near surface materials is still a topic of research, more petrophysical investigations for hydrogeological applications have been performed using electrical measurements than using other types of geophysical measurements.

The wide range of values of electrical resistivity in natural materials and the sensitivity of electrical resistivity to hydrogeological properties (such as moisture content, type of pore fluid, porosity, and lithology) have rendered electrical resistivity a useful tool for a wide range of hydrological investigations. Because of the myriad of influences on apparent electrical resistivity measurements, it is typically not possible to estimate hydrogeological properties, such as lithology or hydraulic conductivity, using electrical data without other constraining information. The increasing ease of use and reasonable cost of electrical equipment have greatly expanded the ability to use this method to survey large areas or to perform time-lapse imaging of dynamic processes.

14.2.2 Electromagnetic Induction

The use of electromagnetic techniques for groundwater and environmental site assessments has increased dramatically in recent years. Controlled-source inductive electromagnetic (EM) methods use a transmitter to pass a time-varying current through a coil or dipole on the earth's surface. This alternating current produces a time-varying magnetic field that interacts with the conductive subsurface to induce time-varying eddy currents. These eddy currents give rise to a secondary EM field (Figure 14.2). Attributes of this secondary magnetic field, such as amplitude, orientation, and phase shift, can be measured by the receiver coil and, by comparing these attributes with those of the primary field, information about the presence of subsurface conductors, or the subsurface electrical conductivity distribution, can be inferred.

As with electrical resistivity measurements, electrical conductivity measurements made using EM methods are also affected by moisture content, type of pore fluid, porosity, and lithology. EM methods

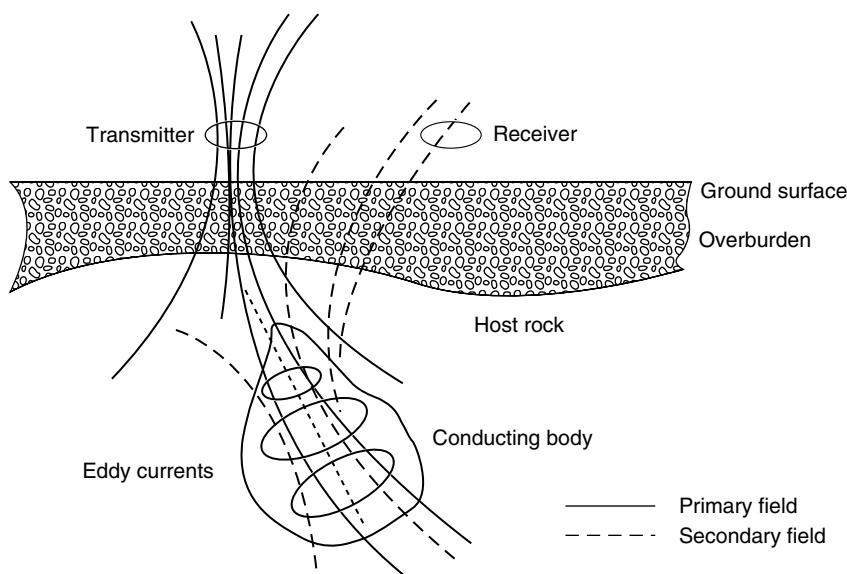


FIGURE 14.2 Electromagnetic induction technique. Changes in the primary magnetic field with time-induce eddy currents in the subsurface that give rise to a secondary EM field. (From Best, M. E., 1992, *Geological Association of Canada Short Course Notes Volume 10*. Wolfville, Nova Scotia, May 28–29. With permission.)

tend to require less time to acquire than DC resistivity methods and have the capability to image deeper subsurface regions. However, the interpretational methods necessary to extract qualitative information can be more complicated than those used with resistivity methods. As a conductive subsurface environment or target is required to set up the secondary field measured with inductive EM methods (Figure 14.2), EM methods are best suited for use when attempting to detect the presence of high-conductivity subsurface targets, such as salt water saturated sediments or clay layers. However, because coils do not require contact with the ground, EM methods are often more successful in electrically resistive or paved areas than DC resistivity method, which require electrode contact.

Electromagnetic instrumentation can take on many different forms. Most EM systems consist of a source that transmits a systematically-varying EM field, one or more receivers to measure the components of the total (primary and secondary) magnetic field and sometimes the components of the electric field, and the necessary electronic circuitry to process, store, and display the signals (Best, 1992). As with resistivity data, EM induction data can be collected in profile or sounding mode. The mode of acquisition and the resolution and penetration of the data are governed by the subsurface conditions, the domain (frequency or time) under which the system operates, and the coil configuration. For frequency domain systems, high transmitter frequencies permit high-resolution investigation of subsurface conductors at shallow depths while lower transmitter frequencies permit deeper observations but at a loss in resolution. Time domain systems measure the secondary magnetic field as a function of time, and early-time measurements yield information about the near-surface, while later-time measurements are increasingly influenced by the electrical properties of deeper layers. The depth of penetration and resolution are also governed by coil configuration; the measurements from larger coil separations are influenced by electrical properties at greater depths, while smaller coil spacings sample from the near-surface. Thus, to collect an EM profile, the coil spacing must be constant *and* the frequency or time of measurement fixed. To collect an EM sounding, measurements are collected at the same surface location using different frequencies, different observation times, *or* different coil spacings. Although still a research topic, EM crosshole geometries have been used for dynamic process monitoring (Wilt et al., 1995).

As electrical conductivity is the inverse of electrical resistivity, the discussion regarding the relation of electrical resistivity to lithology or hydrological parameters given in Section 14.2.1 can be applied in an inverse manner to electrical conductivity measurements. For example, electrical conductivity is higher for saturated sediments than for unsaturated sediments, and clayey materials generally have higher electrical conductivities than sandy materials. The introduction of contaminants into an aquifer may alter the electrical conductivity of the pore fluid. EM methods can be used to detect contaminants (Greenhouse and Harris, 1983) if:

- The contrast between the electrical conductivity of the contaminant and the original pore fluid is at least 10%.
- The plume or contaminated aquifer occupies a significant fraction of the zone of influence of the measurement.
- Variations in electrical conductivity due to other factors that affect electrical conductivity, such as lithology, are smaller than or separable from variations in electrical conductivity due to the contaminant.

Organic contaminants generally reduce electrical conductivity and are more difficult to detect over background variations than inorganic contaminants, which typically increase the electrical conductivity, but EM has been successful at detecting both organic and inorganic groundwater contamination plumes (Buselli et al., 1990).

Reviews of the instrumentation available for EM induction systems, their applicability for environmental site characterization, and EM interpretational methods are given by Frischknecht et al. (1991), Hoekstra and Blohm (1990), Goldstein (1994), Mac Lean (1992), and McNeill (1990). Examples of investigations that have used EM inductive methods to investigate hydrogeological parameters include those of Sheets and Hendricks (1995), who used EM induction methods to estimate soil water content, and Kachanoski et al. (1988), who used EM induction methods to investigate the spatial variations of soil

texture and pore fluid. An excellent review and discussion of the use of controlled source EM methods for hydrogeological investigations is given by Everett and Meju (2005).

14.2.3 Ground Penetrating Radar

Ground penetrating radar (GPR) methods use electromagnetic energy at frequencies of 10 to 1000 MHz to probe the subsurface. At these frequencies, the separation (polarization) of opposite electric charges within a material that has been subjected to an external electric field dominates the electrical response. The dielectric constant or relative permittivity, which is the ratio of the electric permittivity of the material relative to the electric permittivity of free space, is used to describe these high-frequency electric properties. GPR systems consist of an impulse generator that repeatedly sends a particular voltage and frequency source to a transmitting antenna. The GPR technique is similar in principle to seismic and sonar methods. In the case of the most commonly used bistatic systems, one antenna, the transmitter, radiates short pulses of high frequency (MHz to GHz) electromagnetic waves, and the other antenna, the receiver, measures the characteristics of the (modified) signal as a function of time. When the source antenna is placed on the ground surface, spherical waves are radiated both upward into the air and downward into the soil. Some of the propagating energy travels into the subsurface and is reflected by subsurface dielectric contrasts; these arrivals are known as GPR reflections. Some of the energy travels in the near subsurface along the air-ground interface; these arrivals are known as GPR groundwaves. In general, GPR performs better in unsaturated coarse or moderately coarse textured soils; GPR signal strength is attenuated in electrically conductive environments, such as systems dominated by the presence of clays or that have high ionic strength pore fluids.

Dielectric constants are affected by material saturation, porosity, material constituency, temperature, and pore fluid composition. Approximate dielectric constants for common near-surface materials are given in Table 14.1. This table shows that most dry materials have dielectric constants of 3 to 7, which are much lower than the dielectric constants of saturated materials.

The most common GPR acquisition mode is surface common-offset reflection, which involves the collection of one trace per surface location from a transmitter-receiver antenna pair (Figure 14.3a). With this acquisition mode, data are displayed as two-dimensional wiggle trace profiles. Such profiles can be displayed in real time during data collection and can be stored digitally for subsequent data processing. Resolution of GPR reflection data, defined as the ability of the system to distinguish two signals that are

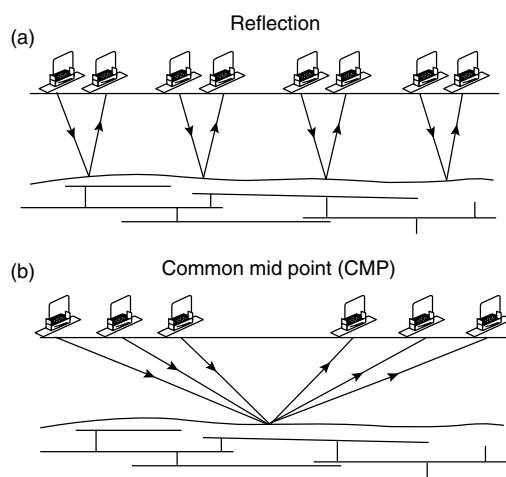


FIGURE 14.3 GPR acquisition geometry: (a) common-offset reflection mode and (b) common-midpoint mode. (From Annan, A. P. and S. Cosway, 1992, Ground penetrating radar survey design. Proc. of the Symposium on the Applications of Geophysics to Engineering and Environmental Problems [SAGEEP], Chicago. With permission.)

close together in time, varies as a function of antenna frequency and the electrical properties of the host material. Davis and Annan (1989) suggest that for a wet soil, vertical resolution on the order of 7 m is possible using 10 MHz radar antennas, 0.5 m resolution is possible using 100 MHz radar antennas, and less than a centimeter resolution is possible using 1000 MHz radar antennas. Increasing the frequency increases the resolution but decreases the penetration depth. Also affecting the radar signal penetration depths are the radar system performance, reflection properties at boundaries where electrical properties vary, and attenuation that is controlled by scattering, electrical properties of surface and subsurface materials, and geometrical spreading losses. Generally, one chooses a radar center frequency that yields both sufficient penetration and resolution; for field applications this is often between 50 and 200 MHz. It is, however, possible to collect radar data at the same location using several different frequency antennas. The use of multiple frequency antennas enables sampling over a greater depth range with optimal resolution for all depths (Smith and Jol, 1992).

Estimation of the dielectric constant is necessary to infer quantitative hydrogeological information from GPR data. For the high frequency range of interest used with GPR methods, the propagation phase velocity (V) in a low electrical conductivity material can be related to the dielectric constant (κ) as (Davis and Annan, 1989):

$$\kappa \approx \left(\frac{c}{V} \right)^2 \quad (14.4)$$

where c is the propagation velocity of electromagnetic waves in free space (3×10^8 m/sec).

Information about the dielectric structure of the shallow subsurface can be obtained via velocity analysis of reflection GPR data acquired with surface common-midpoint (CMP) or velocity analysis of transmitted signals acquired with crosshole acquisition geometry and the relationship shown in Equation 14.4. GPR CMP geometry results in gathers of traces where the midpoint between the transmitting and receiving antennas is common to several antenna pairs as shown in Figure 14.3b. GPR CMP data have been processed using a standard seismic data processing sequence to obtain estimates of dielectric constants at one location and as a function of depth (e.g., Fisher et al., 1992; Lunt et al., 2005).

Crosshole tomographic GPR techniques consist of measuring direct electromagnetic energy from a transmitter in one borehole to a receiving antenna in another borehole. A typical cross-hole tomographic geometry consists of sources and detectors located in separate boreholes. Direct energy from the transmitter in one borehole is received by several receivers that are connected by a cable in the other borehole. The vertical source position is changed, and the recording is repeated until the source has occupied all positions in the source borehole. The interwell area is discretized into a grid composed of cells or pixels, and inversion algorithms are used to transform the recorded travel time and amplitude information into estimations of velocity and sometimes attenuation at each pixel (Peterson et al., 1985; Peterson, 2001). Crosshole acquisition geometry has also been used with electrical and electromagnetic methods to produce high-resolution images of electrical conductivity between two boreholes.

Although both multi-offset GPR acquisition and tomographic methods yield information about electromagnetic velocity (and through Equation 14.4, dielectric constants), these datasets are time consuming to acquire, and thus are not well suited for rapid mapping of dielectric constants at the field scale. Dielectric constant can also be obtained using common-offset GPR acquisition approaches, provided some additional information is available. The use of GPR groundwave data, collected using common-offset acquisition geometry, for dielectric constant mapping was first suggested by Du and Rummel (1994). The groundwave is the part of the radiated energy that travels between the transmitter and receiver through the top of the soil. Recent experiments by Hubbard et al. (1997a), Huisman et al. (2002, 2003), and Grote et al. (2003, 2005) have confirmed that this is a viable approach for mapping dielectric constant, and that the information could be accurately transformed into estimates of shallow soil water content. An example of the use of GPR for water content estimation is provided in Section 14.3.

If the depth to a reflector is known or can be reasonably estimated, GPR reflection travel time data collected using common offset approaches can also be used to estimate the velocity, or with Equation 14.3,

the dielectric constant values of the subsurface. This method was used by Lunt et al. (2005) to map the dielectric structure of an agricultural site. The reflection method has an advantage over the groundwave method in that it can penetrate deeper into the subsurface. However, the reflection approach requires both a good reflecting horizon and information about the depth to the horizon, which can limit the applicability of this approach over large distances.

Petrophysical models can be used to link dielectric estimates obtained from GPR data to hydrogeologic parameters. Most commonly, dielectric constant estimates are used to obtain estimates of water content. For example, Topp et al. (1980) used a regression analysis to obtain an empirical relation between dielectric constant and volumetric moisture content (θ):

$$\theta = -5.3 \times 10^{-2} + 2.92 \times 10^{-2}\kappa - 5.5 \times 10^{-4}\kappa^2 + 4.3 \times 10^{-6}\kappa^3 \quad (14.5)$$

Greaves et al. (1996) used the Topp equation to transfer interval velocity estimates obtained from surface CMP GPR data into subsurface water content estimates, and Hubbard et al. (1997b) successfully used the relation to transfer dielectric constants obtained from crosshole radar measurements into water content estimates. For a sand-clay system, Wharton et al. (1980) used a mixture formula to express the bulk dielectric constant as a function of the dielectric constants of air (κ_a), water (κ_w), sand (κ_s), and clay (κ_{cl}), as well as the porosity (ϕ), fractional water saturation (S_w), and the volumetric clay fraction (V_{cl}):

$$\kappa = [(1 - \phi)V_{cl}\sqrt{\kappa_{cl}} + (1 - \phi)(1 - V_{cl})\sqrt{\kappa_s} + S_w\phi\sqrt{\kappa_w} + (1 - S_w)\phi\sqrt{\kappa_a}]^2 \quad (14.6)$$

Knoll et al. (1995) found under laboratory conditions that Equation 14.6 fit the observed dielectric behavior of a constructed unconsolidated sand–clay mixture. They converted the relation between effective dielectric constant, saturation, porosity, and clay fraction shown in Equation 14.6 to a relation between effective dielectric constant, saturation, and intrinsic permeability using petrophysical relationships. Their petrophysical formulas, which were developed theoretically and verified in the laboratory, are shown in Figure 14.4. These curves reveal that the dielectric constant is a function of both saturation and intrinsic permeability and that the dielectric constant is much more sensitive to saturation state than to the intrinsic permeability when the permeability values are high. Using the petrophysical relations given by Knoll et al. (1995), Hubbard et al. (1997b) developed a technique to estimate permeability and saturation in bimodal systems using CMP or crosshole GPR data; this example will be discussed in Section 14.4.

The petrophysical relations discussed above were all based on a pore fluid of water; however, contaminants may have a different dielectric constant than water and may be detectable in these cases using radar methods. GPR methods have been used for imaging perchloroethylene (PCE) contamination (Greenhouse et al., 1993) and trichloroethylene (TCE) contamination (Brewster and Annan, 1994). Because of the high frequencies employed, GPR offers perhaps the best resolution of any surface geophysical technique when operating under optimal conditions. An excellent overview of the GPR method is given by Annan (2005).

14.2.4 Seismic Methods

Seismic reflection, crosshole transmission, and refraction methods use artificially generated, high-frequency (100 to 5000 Hz) pulses of acoustic energy to probe the subsurface. These pulses can be delivered to the subsurface with hammer blows, weight drops, blasting caps, rifle or shotgun slugs, piezoelectric devices, or vibrators. These disturbances are produced at a point and propagate outward as a series of wavefronts. The passage of the wavefront creates a motion that can be detected by a sensitive geophone and recorded on the surface; the path that the energy follows is conceptualized by a ray drawn normal to the wavefront.

According to the theory of elasticity upon which seismic wave propagation is based, four different waves are produced by a disturbance, and these waves travel with different propagation velocities governed by the elastic constants and density of a material. One type of wave is transmitted by a back-and-forth particle movement in the direction of the propagating wave. This type has the greatest speed and is thus the first

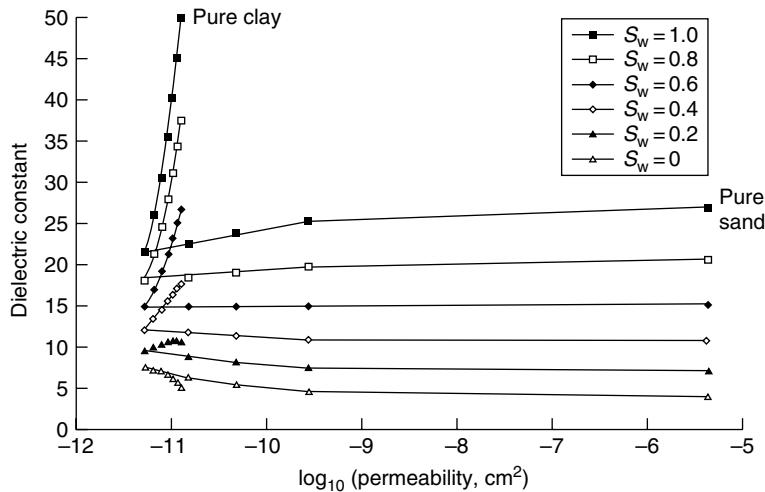


FIGURE 14.4 Petrophysical curves of Knoll et al. (1995) relating dielectric constant to fractional water saturation (S_w) and permeability. (From Knoll, M., R. Knight, and E. Brown, 1995, Can accurate estimates of permeability be obtained from measurement of dielectric properties? SAGEEP Annual Meeting Extended Abstracts. Orlando, FL, Apr. 23–26. With permission.)

to arrive at a geophone; for this reason it is called the P (primary)-wave. Other terms for this wave are the compressional or longitudinal wave. Approximate values for P-wave velocities in near surface materials are given in Table 14.2. The particle motion of transverse waves is perpendicular to the direction of the wavefront propagation. Transverse waves, also called S (secondary or shear)-waves, have lower velocities than the P-wave and thus arrive later in the recording. In contrast to both P- and S-waves, which are transmitted through the interior of a solid and are thus classified as body waves, surface waves, including both Raleigh and Love waves, are confined to a region near the free surface of a medium. Because it is easiest to detect the P-wave energy, most seismic exploration techniques have focused on the use of this wave, and in the following discussion, we also concentrate exclusively on information available from P-waves.

Figure 14.5 shows the major paths of P-wave energy upon disturbance in a horizontally layered, two-medium system where the *seismic velocity* of the upper layer is less than that of the lower layer. The first arrival at the detector or geophone is the (compressional) air wave, which is followed by the direct compressional wave that propagates along a path from the disturbance to the geophone. P-wave energy is reflected from the interface and also refracted from the head wave that travels along the interface. In addition to these compressional arrivals, ground motion from Raleigh surface waves, or ground roll, also contributes to the signal recorded at a geophone. The events commonly used for interpretation are the reflected P-waves, the direct P-waves, and the refracted P-waves. Different seismic methods focus on different seismic wave arrivals — reflection techniques focus on reflected arrivals, crosshole methods mainly utilize direct waves, and refraction methods are designed to detect refracted head waves. The principles of each of these different seismic methods are briefly described below, and an excellent review of shallow seismic acquisition and processing techniques is given by Steeples (2005).

14.2.4.1 Seismic Reflection

The surface reflection technique is based on the return of reflected P-waves from boundaries where velocity and density (or seismic impedance) contrasts exist. Seismic reflection data are usually collected as common-shot or common-receiver gathers that are sorted during processing into common-midpoint or CMP gathers. These gathers of traces represent reflections from a subsurface location (the midpoint) that has been sampled by several source–receiver pairs. Although processing of near surface seismic data often follows the procedures developed for petroleum reservoirs (e.g., Yilmaz, 1987), the acquisition parameters for shallow seismic campaigns are different from those used to image deeper reservoirs (e.g., Pullan

TABLE 14.2 Seismic P-Wave Velocities for Typical Near-Surface Materials

Material	Approximate P-wave Velocity (m/sec)
<i>Unconsolidated Materials</i>	
Weathered layer	300–900
Soil	250–600
Alluvium	500–2000
Clay	1100–2500
Unsaturated sand	200–1000
Saturated sand	800–2200
Unsaturated sand/gravel	400–500
Saturated sand/gravel	800–2200
Unsaturated glacial till	400–1000
Saturated glacial till	1700
<i>Consolidated Materials</i>	
Sandstone and shale	2000–4500
Chalk	2300–2600
Limestone	2000–6000
Coal	2200–2700
Basalt	5400–6400
Granite	5000–6400
Metamorphic	3500–7000
<i>Other</i>	
Water	1400–1600
Ice	3400–3800
Air	331.5

Source: Adapted from Press, F., 1966, Seismic velocities, in *Handbook of Physical Constants*, rev. ed., Clark, S.P., Jr., Ed. Geological Society of America Memoir, Vol. 97, pp. 97–173.

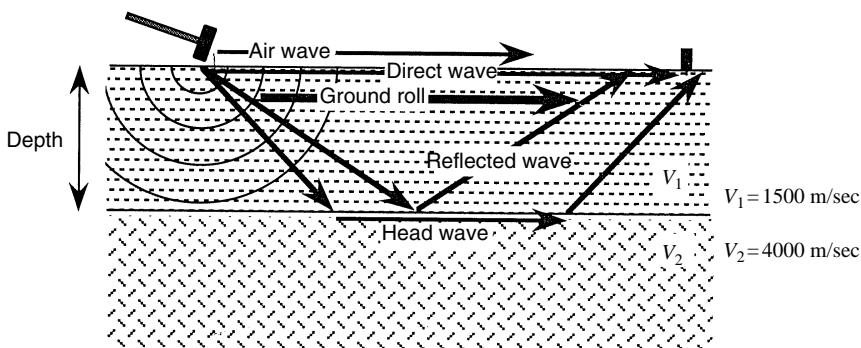


FIGURE 14.5 Major raypaths of P-wave energy. (From Burger, H. R., 1992, *Exploration of Geophysics of the Shallow Subsurface*. Prentice Hall, Englewood Cliffs, NJ. With permission.)

and Hunter, 1990). Due to the lack of well-defined velocity contrasts in unconsolidated and unsaturated materials (Table 14.2), seismic reflection data acquisition is often difficult in the near-surface. Hunter et al. (1984) demonstrated the importance of selecting the appropriate distance between the seismic source and the seismic geophones for near-surface seismic reflection imaging, and Doll et al. (1994) reviewed various types of sources for collecting near-surface seismic data.

Computer-based processing of seismic reflection data generally produces a wiggle-trace profile that resembles a geologic cross section; Figure 14.6 is an example of a seismic reflection section over a steep-sided valley that is floored by bedrock. An interpreted borehole log has been inserted into the section

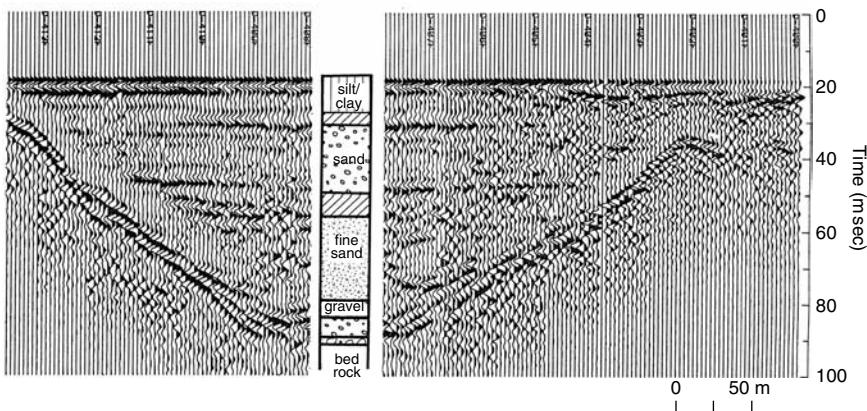


FIGURE 14.6 High resolution near-surface seismic section over a steep-sided valley. Interpreted borehole log inserted at the borehole location reveals that reflections imaged are interfaces between different hydrostratigraphic units. (From Pullan, S. E. and Hunter, J. A., 1990, Delineation of buried bedrock valleys using the optimum offset shallow seismic reflection technique, in *Geotechnical and Environmental Geophysics*, Vol. 3, Geotechnical, S.E.G. Investigations in Geophysics 5, Ed. Stanley Ward, pp. 75–87. With permission.)

at the borehole location and reveals that the seismic reflections correspond to interfaces between sands, clays, gravels, and bedrock. The vertical resolution of a few meters shown in this figure is an example of good quality surface seismic data collected in an area favorable for seismic data acquisition with careful attention to acquisition parameters. Seismic reflection methods have also been used to delineate the water table (Birkelo et al., 1987), to delineate shallow faults (Treadway et al., 1988), and to map shallow bedrock (Miller et al., 1989).

14.2.4.2 Seismic Refraction

With refraction methods, the incident ray is refracted along the target boundary before returning to the surface (Figure 14.4). The refracted energy arrival times are displayed as a function of distance from the source, and interpretation of this energy can be accomplished by using simple software or forward modeling techniques. The arrival times and distances can be used to obtain velocity information directly. Refraction techniques are most appropriate when there are only a few shallow (<50 m) targets of interest, or where one is interested in identifying gross lateral velocity variations or changes in interface dip (Lankston, 1990). Seismic refraction methods yield much lower resolution than seismic reflection and crosshole methods. However, because refraction methods are inexpensive and acquisition may be more successful in unsaturated and unconsolidated environments, it is sometimes chosen over reflection methods for applications such as determining the depth to the water table and to the top of bedrock, the gross velocity structure, or for locating significant faults. A review of the refraction method is given by Lankston (1990).

14.2.4.3 Tomography

Seismic tomography is commonly used in the petroleum industry, but has also been used for hydrogeological applications. As with radar tomography, the multiple sampling of the intrawell bore area permits very detailed estimation of the velocity structure. Although the majority of the crosshole tomographic seismic data sets have been collected for research projects, the extremely high resolution of up to 0.5 m offered by this method suggests that it offers extremely useful information for hydrogeologic characterization and monitoring. As seismic P-wave velocities can be related to lithological and hydrogeological parameters as discussed above, this extremely high resolution method is ideal for detailed stratigraphic and hydraulic characterization of interwell areas (e.g., Hyndman et al., 1994, 1996; Hubbard et al., 2001; Chen et al., 2005). Although not yet established as a field method, seismic imaging of organic contaminants in the

laboratory has provided a fundamental step toward the application of seismic tomographic imaging of interwell contamination (Geller and Myer, 1995). Recently, time-lapse seismic tomographic data have been used to image the evolution of products associated with field-scale remediation, such as the generation of gasses and precipitates (Hubbard and Williams, 2004). The processes of crosshole seismic data inversion and incorporation of these data with hydrological data are discussed in detail in Chapter 15 of this volume (Hyndman, 2006).

14.2.4.4 Seismic Petrophysics

Information about seismic velocity is available through seismic data processing of reflection CMP, refraction, or tomographic data. Seismic P-wave velocity is affected by porosity, permeability, pore fluid type, depth of burial, consolidation, and temperature. Table 14.2 reveals that unique relations between seismic P-wave velocities and lithology generally do not exist. However, some generalities can be made regarding seismic velocities and lithology (Burger, 1992):

- Velocities in unsaturated sediments are lower than in saturated sediments.
- Velocities in unconsolidated sediments are lower than in consolidated materials.
- Velocities are very similar in unconsolidated, saturated sediments.
- Velocities in weathered rocks are lower than in similar but unweathered rocks.

In general, then, it is possible using seismic velocity to distinguish dry from saturated sediments and sediments from consolidated rocks.

Empirical, mechanistic, and theoretical approaches have been used to relate seismic P-wave velocity to hydrogeological parameters. For example, a velocity (V)–porosity (ϕ) relationship is given by the Wyllie time-average equation using estimated matrix velocity (V_m) and pore fluid velocity (V_f) values:

$$\frac{1}{V} = \frac{1 - \phi}{V_m} + \frac{\phi}{V_f} \quad (14.7)$$

Marion et al. (1992) used a theoretical microgeometrical approach to relate seismic P-wave velocity to pressure and porosity within an unconsolidated, sand-clay system. This relation was verified empirically in the laboratory, and the model was extended by Rubin et al. (1992) to investigate the relationship between log permeability, effective pressure, and seismic P-wave velocity. Regression techniques are commonly used to relate site-specific measurements, such as velocity, porosity, and lithology. For example, Klimentos (1991) used regression analysis to obtain a relation between velocity, porosity, and clay content in a consolidated, sand-clay system. As is described in a very thorough review on seismic petrophysics given by Pride (2005), a universal relation between permeability and seismic velocity does not exist: the relationship depends on how the geologic materials were formed and have evolved over time.

14.2.5 Potential Field Methods

14.2.5.1 Gravitational Methods

Measurements of the changes in gravitational acceleration can be used to obtain information about subsurface density variations. As density is a bulk property of rocks and tends to be consistent throughout a geological formation, gravity methods are used to identify gross features based on density variations. Due to the lower resolution offered by this method, it is not commonly used for detailed site characterization. It does, however, provide an inexpensive way to detect or map targets such as the interface between sedimentary overburden and bedrock or locations of significant faults. Gravity methods have also been used to detect sinkholes, other subsurface voids, and landfill boundaries. The common measuring device is a gravimeter, an instrument that is portable and easy to use. A spring balance inside the gravimeter measures differences in the weight of a small internal object from location to location; the weight differences are attributed to changes in the acceleration of gravity due to lateral variations in subsurface density.

Measurements can be collected at a regional or local scale depending on the station spacing, which is usually less than half of the depth of interest.

Gravity measurements must be adjusted prior to interpretation by comparing the observed gravity values with the calculated response of a homogeneous formation. The theoretical response to the gravitational field due to such factors as the datum, latitude, tidal, altitude, terrain, drift, and regional gradient are calculated and subtracted from the raw data to yield what is called a gravity anomaly field (Zodhy et al., 1974). A Bouguer anomaly refers to the value that is obtained after latitude corrections, elevation corrections, and (sometimes) terrain corrections have been applied to the gravity data. Anomaly field interpretations are non-unique; a large number of hypothetical features with varying density, depth, and shape can produce the same gravity anomaly. Qualitative interpretation usually consists of constraining a profile or contoured anomaly map with other known geologic information. Forward modeling techniques and characteristic curve matching are also used for interpretation. Reviews of the gravity technique and applications to environmental studies are given by Hinze (1990) and by Butler (1991).

14.2.5.2 Magnetic Methods

Magnetic methods measure the direction, gradient, or intensity of the earth's magnetic field. The intensity of the magnetic field at the earth's surface is a function of the location of the observation point in the primary earth magnetic field as well as local or regional concentrations of magnetic material such as magnetite, the most common magnetic mineral. Magnetometers are used to measure the total geomagnetic field intensity or relative values of the vertical field intensity, and magnetic gradiometers measure the horizontal and vertical gradient of this magnetic field. After correcting for the effects of the earth's natural magnetic field, magnetic data can be presented as total intensity, relative intensity, and vertical gradient anomaly profiles or contour maps. Interpretation of magnetic surveys generally involves forward modeling or mapping of the anomalies and correlating them with other known geologic information. This interpretation is non-unique.

Magnetic methods are generally used to identify gross features at a resolution similar to that of seismic refraction and gravitational methods. As magnetic signatures depend to a large extent on magnetic mineral content, which is low in most sediments that comprise aquifers, magnetic field work is not commonly employed for near-surface investigations. It is used, however, to map the depth to basement rock, provided that it contains sufficient magnetic minerals. Magnetic methods with fine station spacing and high lateral resolution (a few meters) are now one of the most commonly used geophysical methods for site investigation due to their ability to locate metal objects such as drums and abandoned drill hole casings that are buried in near-surface sediments. A review of magnetic methods as applied to environmental problems is given by Hinze (1990).

14.2.6 Borehole Geophysical Methods

Borehole geophysics refers to the process of recording and analyzing of geophysical log measurements collected within wellbores. This information can be used together with direct measurements (such as core samples) or hydrological information (such as flowmeter or pumping tests) to develop site-specific petrophysical relationships. Log data are also useful to "tie" together more direct information collected at the wellbore location with geophysical signatures of property variations collected using surface geophysical data. With a combination of well log and surface geophysical data, hydrogeological information measured at the wellbore can be extrapolated between boreholes.

The volume of material sampled by the borehole log, or the volume of investigation, is related to the log type, source-detector spacing, and subsurface material, and thus varies with well site conditions and logging parameters employed. Interpretation of the recorded log data often involves comparing several different logs displayed side-by-side or by cross-plotting data from one log against data from other logs, core analyses, or tests. For hydrocarbon exploration, the decision to test and complete a well is based to a large extent on geophysical log information, and as a consequence, most of the interpretation guidelines for borehole geophysics have been developed for borehole and rock environments encountered in petroleum

exploration. Excellent references for borehole geophysics applied to hydrogeologic investigations are given by Keys (1989) and Kogr et al. (2005), which are the references for the following information where not otherwise cited. A review of the process of drilling, coring, and sampling water wells is given by Shuter and Teasdale (1981).

Borehole measurements are made by lowering a sonde into the borehole on the end of an electric cable. The sonde is a probe, generally 2.5 to 10.0 cm in diameter and 0.5 to 10.0 m in length, which encloses sources, sensors, and the electronics necessary for transmitting and recording signals. Measurements made in the borehole are recorded on the surface in analog form on chart paper or in digital form as shown in Figure 14.7, which is an example of several logs that have been collected in a shallow water well.

The following discussion focuses on the underlying physical principles of logging methods that are currently used in groundwater applications or that have potential for aiding these investigations, including caliper, electric, nuclear, and acoustic.

14.2.6.1 Caliper Logs

Caliper logs are mechanical or acoustic tools that measure the diameter of the borehole. The mechanical caliper tool is a single arm connecting between one and six caliper probes that are pressed against the borehole wall by spring pressure. As the tool is pulled up the borehole, the mechanical caliper probes move in response to changes in borehole diameter and the acoustic calipers measure the reflection transit time of an acoustic signal from the borehole wall. Changes in the diameter of the borehole affect the response of all geophysical tools, and thus a caliper log is generally collected in conjunction with all logging types to aid in interpretation and correction. As sloughing of the material is a function of the material cohesiveness and thus of the lithology, caliper logs are also used to indicate lithology or fracture zonation.

14.2.6.2 Electric Logs

Electric logs measure potential differences due to the flow of electric current in and adjacent to the well. There are many different types of electric logs, all of which must be run prior to casing and in saturated boreholes.

Single-point resistance is a very simple tool that is used to measure resistance by passing a constant current between an electrode in the well and another located on the surface. Electrical resistivity, includes the dimensions of the material being measured and is an intrinsic property of the material. Resistance depends on the material composition as well as on the cross-sectional area and length of the current path through the medium. Electrical resistance (r), measured in Ohms (Ω), is related to resistivity (ρ) by

$$r = \frac{\rho L}{S} \quad (14.8)$$

where S is the cross-sectional area normal to the flow of current (m^2) and L is the length of the current path in meters. The single-point resistance log shown in Figure 14.6 reveals the usefulness of this tool for high resolution lithological identification; the sand and gravel units characteristically give a higher resistance reading than the silts and clays. *Resistivity* measurements are made by sending a current through electrodes lowered down the borehole and recording the voltage across other electrodes located on the sonde. Resistivity calculations are performed using Ohm's law (Equation 14.1). Resistivity logs are commonly used to calculate water quality, as well as saturation and porosity (Equation 14.3). Short and long *normal, focused, and lateral* logs are all resistivity methods that differ by the electrode configuration on the sonde, and *microresistivity* logs employ electrodes that are forced to contact the borehole wall. All configurations measure the resistivity of the subsurface material, pore fluid, borehole fluid, or some combination of these. *Dipmeter* tools use three or four contact microresistivity probes connected to a single arm that press against the borehole walls to measure resistivity simultaneously at multiple orientations within the borehole. Interpretation involves correlating resistivity measurements made from each probe to estimate primary sedimentary structure orientation of the units penetrated by the wellbore. Computer analysis of dipmeter data is most effective. A *spontaneous potential* (SP) log measures the natural potential difference

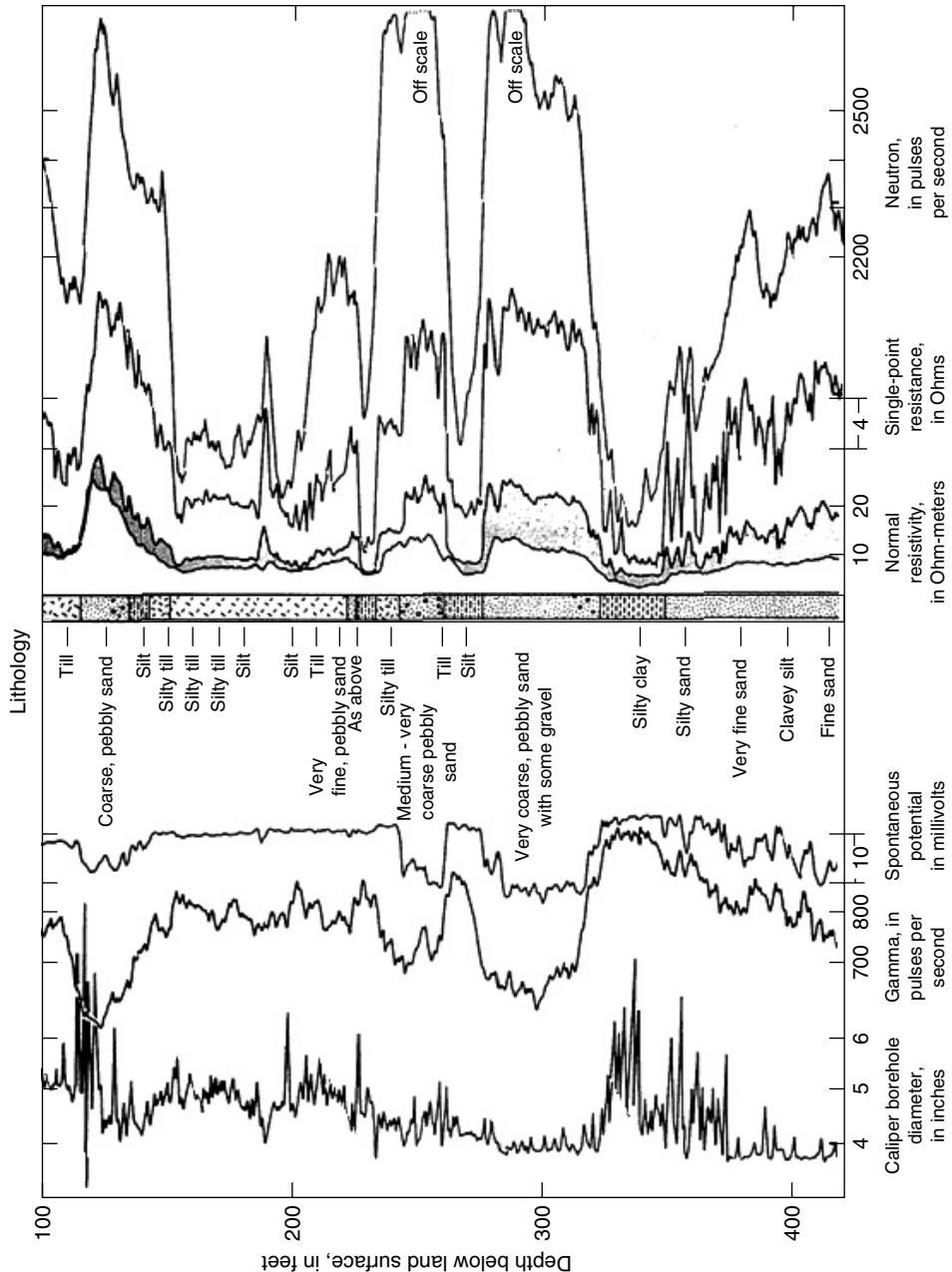


FIGURE 14.7 Borehole geophysical and lithological logs collected at a typical shallow water well. (From Keys, S.W., 1989. *Borehole Geophysics Applied to Ground-Water Investigations*. National Water Well Association, Dublin, OH. With permission.)

between a single electrode in the wellbore and a reference electrode at the surface. These potentials exist due to a combination of electrochemical effects created, for example, by migration of ions from concentrated to more dilute solutions. Additional effects may be caused by electrokinetic (streaming) potentials that result from fluid moving in or out of the hole or redox potentials caused by ionic charge accumulation between metallic mineral grains and the fluids adjacent to the grains.

14.2.6.3 Nuclear Logs

Nuclear logging entails the detection of unstable isotopes near the borehole. The considerable advantage of nuclear logs over electric logs is that they can be run after casing has been installed. As isotopes decay, they emit radiation such as gamma photons or rays from the nucleus. Gamma proton and neutron radiation sources are often used in borehole applications because of their ability to penetrate dense material such as rock and casing. Borehole geophysical tools that measure radioactivity of nearby formations may be classified as those that detect natural gamma radiation, those that employ controlled gamma rays to induce radiation, and those that use neutron sources to induce nuclear processes. The radioactivity is measured as electronic pulses, and the quantity and amplitude of the pulses yield information about the surrounding formation. Logging tools that use artificial radioisotopes as sources are regulated by governmental agencies and require a license for use.

The *gamma log* uses a scintillation detector to measure total naturally occurring gamma radiation of the material penetrated by the borehole within a selected energy range. The three most common naturally occurring radioactive materials that affect the gamma log are potassium-40, uranium-238, and thorium-232. The usefulness of this log lies in the fact that these isotopes are generally more abundant in shales and clays and less common in sands and calcareous materials, and that measurements can be reliably made above the water table. The gamma log shown in Figure 14.6 reveals the characteristically high gamma log count rate associated with silts and clays. In addition to the count rate that is measured with the gamma log, the *gamma-spectrometry* method records the amplitude of the pulses over a wide energy range. Analysis of this energy permits more diagnostic information on lithology and also permits estimation of the type and quantity of radioisotopes that may be contaminating the groundwater.

Gamma-gamma or *density* logs record gamma radiation that originates from an artificial gamma source in the well that is backscattered by the borehole and surrounding material. The count rate of the backscattered gamma rays can be related to the electron density of the material that is in turn proportional to the bulk density of the material. If the fluid and grain densities are known, the bulk density measured with the gamma-gamma log can be used to calculate porosity:

$$\phi = \frac{\text{grain density} - \text{bulk density}}{\text{grain density} - \text{fluid density}} \quad (14.9)$$

As moisture content affects the bulk density of materials, gamma-gamma logs can be used to record changes in moisture above the water table.

Neutron tools consist of an artificial low-energy wellbore neutron source and one or two neutron detectors. The neutrons emitted by the source lose energy upon collision with other elements in the vicinity of the borehole. As hydrogen has a mass similar to the neutron, it is the most effective element at slowing the neutrons. The quantity of slowed neutrons is thus interpreted to be proportional to the quantity of hydrogen present, which is in turn interpreted to be proportional to the moisture content or saturated porosity.

14.2.6.4 Acoustic Logs

Acoustic (sonic or velocity) tools transmit an acoustic pulse from a source, through the fluid and material in the vicinity of the borehole, and to the detector. All acoustic logs require saturated borehole conditions for signal transmission. These tools emit an acoustic source at frequencies of 10 to 35 KHz (0.01 to 0.035 MHz) which creates compressional or P-waves. As the waves travel, some of the energy is refracted back to two receivers located on the sonde. The difference in travel time between the receivers is used to

calculate interval velocity, which is recorded as a function of depth in the wellbore. *Acoustic waveform* logging entails the recording and interpretation of the entire waveform rather than just the time of the first arrival of energy as is recorded with the acoustic log. Analysis of the amplitude, frequency, and other components of the wave train may yield information about the quality of the bond between casing and cement and cement and borehole wall. Additionally, attributes of the signal can yield information about the formation penetrated by the borehole. For example, correlations have been found between attenuation and permeability of fractures.

The *acoustic televIEWer* (ATV) emits high frequency (1.3 MHz) acoustic energy toward the borehole wall. The transducer rotates clockwise at about three rotations per second. During rotation, the transducer sends signals toward and receives signals from the borehole wall. Fractures or other openings in the borehole wall force the signal to travel farther through the fluid and thus to be more attenuated. Attenuation causes amplitude variations that can be visually captured by adjusting the contrasts on the display. The display can be thought of as a cylinder that has been opened along the north side and flattened to provide a 360-degree ultrasonic picture of the borehole wall. Planar horizontal features that intersect the well will appear on this display as dark horizontal bands and planar dipping features that intersect the well will appear as dark sinusoids with the lowest point of the curve in the direction of dip.

14.3 Hydrogeophysical Concepts

In this section, we review the concepts associated with using geophysical data, such as those obtained using the methods described in Section 14.2, for hydrogeological characterization. Geophysical data can be collected from many different platforms, such as from satellites and aircraft, at the ground surface of the earth, and within and between wellbores (Table 14.3). Figure 14.8 illustrates relative length scales and resolutions associated with different geophysical (G) and hydrological (H) data acquisition categories. The scale category definitions defined in Figure 14.7 (point, local, regional) are approximate and are intended to convey the typical spatial extent associated with *investigations* rather than to imply specific notions of measurement support scales for individual instruments, averaging or dimensionality. The methods described in Section 14.2 were limited to borehole, crosshole, and surface geophysical techniques; readers interested in laboratory-scale geophysical methods for hydrogeological applications are referred to Ferre et al. (2005), and readers interested in the use of remote-sensing techniques for hydrogeological characterization are referred to Paine and Minty (2005). Figure 14.8 illustrates that tradeoffs typically exist between the resolution of measurements collected using a method having a certain acquisition geometry and the typical scale of investigation for that method. For example, core measurements provide information about a very small portion of typical field investigations at a very high resolution, while surface geophysical techniques often provide lower resolution information but over larger spatial scales. It is important to realize that the scale and resolution of a particular characterization technique may vary depending on the acquisition parameters and geometry employed. For example, electrical resistivity measurements are commonly collected from core samples, electrical logs, between wellbores, and from the ground surface; electrical measurements can thus be useful for point, local, and regional-scale investigations, and can have a range of resolutions. Additionally, varying the acquisition parameters of a particular method can change the resolution, even if the acquisition geometry stays the same. For example, collection of crosshole seismic travel-time measurements between wellbores that have a great separation distance and that have large source-receiver station spacing within the wellbores will result in a lower possible spatial resolution than seismic crosshole data collected using close wellbore spacing and source-receiver station spacing.

Tables 14.3 and 14.4 presents the surface and crosshole geophysical methods that were discussed in Section 14.2. The methods are classified according to their acquisition approach, and the table lists the geophysical attribute that is typically obtained from the method and examples of hydrogeological objectives for which each method is particularly well suited. The choice of which characterization or technique or acquisition approach to use for a particular investigation is made by considering many factors, such as: the objective of the investigation relative to the sensitivity of different geophysical methods to that

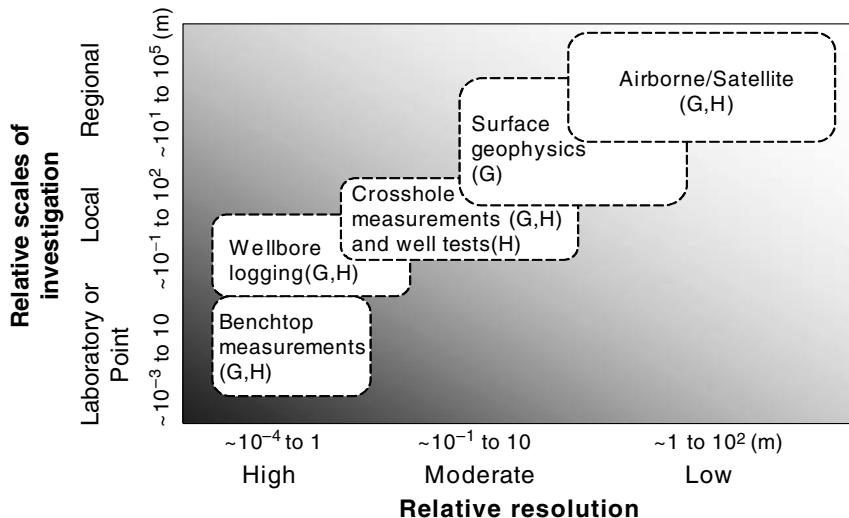


FIGURE 14.8 Chart showing trade-off between the relative resolution of the information obtained using different yet typical geophysical (G) and hydrological (H) measurement acquisition approaches and the relative scale of the investigations for which those acquisition geometries are typically used. Benchtop measurements refer to those such as core, column, and tank measurements collected within a laboratory.

objective; the desired level of resolution; conditions at the site; time, funds, and computational resources available for the investigation; experience of the investigator; and availability of other data. Some of the analyses and data-acquisition techniques described here can be carried out as reconnaissance studies, painting a picture of the aquifer using a wide paintbrush, while others can be applied at an advanced stage, resulting in a more detailed portrait. For example, in applications such as resource management and regional mass balance analyses, characterization of the aquifer through its geometry and effective properties may be all that is necessary; in fact, this may be all that can be realistically achieved given the severe data limitations usually encountered. However, studies of contaminant transport generally require detailed mapping of the spatial distribution of the hydraulic conductivity.

Data are often collected in a sequential fashion, whereby lower-resolution acquisition approaches are used for reconnaissance investigation, followed by higher-resolution approaches to provide more detailed information as needed. As there may be several plausible subsurface scenarios that could produce a given geophysical response, and because different geophysical techniques are sensitive to different properties, a combination of different geophysical techniques is often used to characterize the subsurface. Combinations of hydrogeological and geophysical data can be used to yield a better understanding of shallow subsurface parameters and processes over various scales and resolutions.

Examples of the relative value of the different methods for various surface and crosshole hydrogeological characterization objectives are given in Table 14.4. Table 14.3 and Table 14.4 suggest that geophysical data are used to assist with a wide range of hydrological objectives. These objectives can be broadly categorized into three key areas (1) Hydrogeological mapping, (2) Hydrological parameter estimation, and (3) Monitoring of hydrological processes, each of which is briefly described below.

14.3.1 Hydrogeological Mapping

Hydrogeological mapping using geophysical data is perhaps the most common hydrogeophysical characterization objective. Mapping techniques rely on the measurement of subsurface physical properties or contrasts of these properties. From these measurements, the nature and distribution of subsurface materials can often be deduced. For example, subsurface variations in elastic moduli and density cause seismic waves to travel at different speeds. Information about these changes, and thus about the nature and distribution of the subsurface materials, can be interpreted from analyzing the seismic arrival times.

TABLE 14.3 Common Geophysical Characterization Methods That are Used to Assist in Hydrogeological Investigations

Acquisition Approaches	Characterization Methods	Attributes Typically Obtained	Examples of Hydrogeological Objectives
Surface	Seismic refraction	P-wave velocity	Mapping of top of bedrock, water table, and faults
	Seismic reflection	P-wave reflectivity and velocity	Mapping of stratigraphy, top of bedrock and delineation of faults/fracture zones.
	DC resistivity	Electrical resistivity	Mapping of aquifer zonation, water table, top of bedrock, fresh-salt water interfaces and plume boundaries, estimation of hydraulic anisotropy, and estimation/monitoring of water content and quality.
	Electromagnetic	Electrical resistivity	Mapping of aquifer zonation, water table and fresh-salt water interface, estimation/monitoring of water content and quality.
	GPR	Dielectric constant values and dielectric contrasts	Mapping of stratigraphy and water table, estimation and monitoring of water content
	Gravity	Gravitational field	Mapping of sedimentary overburden, bedrock, and locations of landfill boundaries, significant faults, and sinkholes and other voids.
	Magnetics	Direction, gradient, and intensity of magnetic field	Mapping the depth to bedrock, sedimentary overburden, and locating buried metallic objects, such as waste drums and abandoned casings.
	Crosshole	Electrical resistivity (ERT)	Mapping aquifer zonation, estimation/monitoring of water content and quality
	Radar	Dielectric constant	Estimation/monitoring of water content and quality, mapping aquifer zonation
Wellbore	Seismic	P-wave velocity	Estimation of lithology, fracture zone detection
	Geophysical well logs	Many attributes, including electrical resistivity, seismic velocity and gamma activity	Lithology, water content, water quality, fracture imaging

Similarly, spatial variations in electrical and dielectric properties of subsurface materials produce spatial variations in electromagnetic and ground-penetrating radar responses, respectively. The contrasts in physical properties vary depending on which materials are juxtaposed, and the ability to detect these changes varies with the geophysical method employed. For example, consider the geologic section shown in Figure 14.9, where a saturated clay overlays a saturated sand layer, which in turn overlays a heavily fractured bedrock surface. The velocity contrast between the clay and sand layers may not be detectable using surface seismic methods, but the electrical resistivity contrast is likely sufficient to be detected using electrical methods. Depending on the method, “mapping” may pertain to the mapping of interfaces (as is

TABLE 14.4 Applications of Surface and Crosshole Geophysical Methods

Application	Surface/Crosshole Method	DC Resistivity	ERT	EM	Surface GPR	Crosshole GPR	Reflection Seismic	Seismic Crosshole	Refraction Seismic	Gravity	Magnetics
Depth to water table	4	1	4	3	1	2	0	4	1	1	0
Fresh/salt water interface	4	0	4	2	0	1	0	2	0	0	0
Depth to bedrock	4	0	4	3	0	4	0	4	4	4	2
Gross geometry of units or interfaces	4	0	4	3	0	4	1	4	1	1	1
Detailed geometry of units or interfaces	3	4	3	4	4	4	4	4	1	1	0
Significant fault detection	4	0	4	4	0	4	4	4	4	4	4
Cavity detection	2	2	1	3	3	2	3	0	2	2	1
Porosity or permeability estimation ^a	3	3	3	3	3	3	3	1	1	1	0
Water content estimation ^a	2	4	3	2	4	1	0	1	1	1	0
Water quality estimation ^a	4	4	4	3	3	1	1	1	1	1	1
Plume detection ^a	3	3	3	3	3	1	3	0	0	0	0
Landfill Delineation	4	2	4	4	2	2	2	2	2	2	4

0 = not considered applicable; 1 = limited use; 2 = used, or could be used, but not the best approach or has limitations; 3 = has potential but not fully developed; 4 = generally considered an excellent approach, techniques are well developed.

^aThese applications are current topics of research.

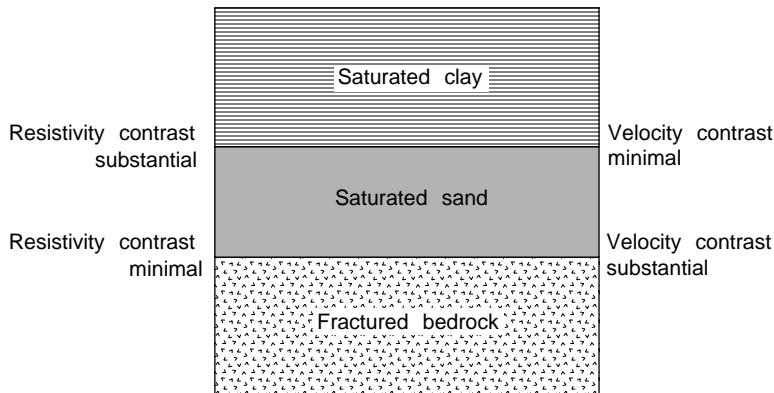


FIGURE 14.9 Schematic geological section and associated electrical resistivity and seismic velocity contrasts at the interfaces. (From Burger, H. R., 1992, *Exploration of the Shallow Subsurface*, Prentice Hall, Englewood Cliffs, NJ. With permission.)

typically the case with surface seismic and GPR methods), or the mapping of layers or units (as is the case with surface electrical or electromagnetic methods).

A common mapping approach with all surface-based geophysical methods is to compare sparse borehole data, which provide direct and detailed information about subsurface hydrogeology as a function of depth at a single location, with less invasive and more laterally continuous geophysical data. As geophysical responses can be non-unique and ambiguous, a comparison between direct measurements and geophysical attributes is typically necessary to interpret the geophysical data in a hydrogeologically meaningful manner. In some cases, this procedure is routine, whereas in other cases it can be quite challenging and demand frameworks for integration. Advanced stochastic techniques for handling more challenging integration problems are discussed in Section 14.4.

Geophysical methods are commonly used for hydrogeological mapping of the geometry of subsurface interfaces or units, the water table, top of bedrock, plumes, faults, and fresh–salt water interfaces. Figure 14.10 shows an example of hydrogeological mapping using 100 MHz GPR data within in an Atlantic Coastal Plain aquifer of Virginia (modified from Hubbard et al., 2001). The locations of the key reflectors visible on the GPR profile were compared with hydrogeological features detected using wellbore data. Once the GPR signatures were “matched” with the features identified using borehole data, the GPR data were then used to delineate key features laterally away from the boreholes. As shown in Figure 14.10, within the top 15 m of this sedimentary subsurface, the GPR profile reveals the water table, key unconformities, several stratigraphic packages, and the top of a subregional aquitard. As was illustrated in Figure 14.6, seismic profiles can also be useful for indicating the interfaces between hydrogeologically significant units.

An example of the use of both electrical profiling and sounding acquisition modes to map hydrogeological units is shown in Figure 14.11a and Figure 14.11b. The map shown in Figure 14.11a is a contour map produced from the resistivity measurements of several profiles collected near San Jose, CA, using an a-spacing (see Figure 14.1a) of 6.1 m (Zodhy et al., 1974; Burger, 1992). Contours of equal apparent resistivity delineate an approximately east-west trending apparent resistivity high. The geologic cross section (BA) shown in Figure 14.11b is based on four expanding-spread traverses (soundings), the apparent resistivity profile information, and information from three boreholes whose locations are indicated on the cross section. The cross section reveals that the high-resistivity area shown on the apparent resistivity map of Figure 14.11a is a zone of gravel and boulders that defines the location of a buried stream channel.

An example of the use of surface EM data for delineating a subsurface conductor is illustrated in Figure 14.12. Figure 14.12a shows an apparent conductivity contour map obtained from interpretation of frequency domain Geonics EM-34 profiling measurements over a brine evaporation pit (Hoekstra et al., 1992). Several measurements were collected, as shown by the dots in Figure 14.12a, using a single frequency and a fixed coil spacing of 20 m. An apparent electrical conductivity high trending from the

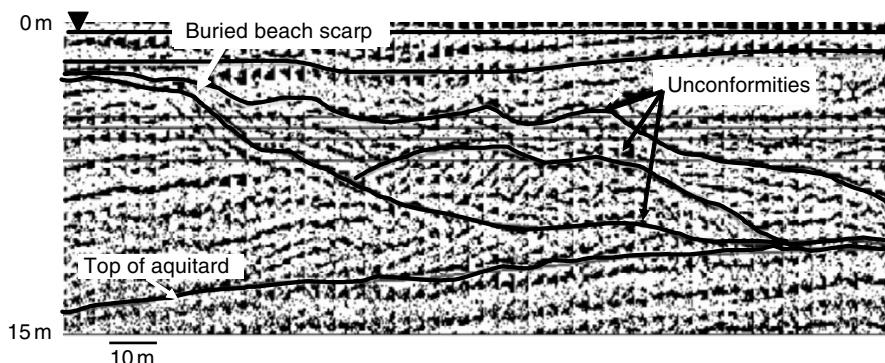


FIGURE 14.10 Example of the use of surface 100 MHz GPR data for mapping shallow subsurface features. (From Hubbard and Rubin, Chapter 1 in *Hydrogeophysics*, Kluwer, the Netherlands, 2005).

brine pond to the southwest is interpreted to be the main direction of contaminant migration. Five EM soundings at the locations indicated by the numbered boxes in Figure 14.12a were collected with the time domain Geonics EM-47 and inverted to yield electrical conductivity estimations as a function of depth. The geologic cross section shown in Figure 14.12b was obtained from lateral interpolation between the five soundings. This cross section suggests that there is a high-conductivity layer beneath stations 2 and 3, interpreted to be the area of brine contamination from the brine pond shown in Figure 14.12a.

An example of the use of gravity data for hydrogeological mapping is shown in Figure 14.13. In this study, gravity data were collected to determine the depths to bedrock in an urban area where the drilling of numerous wells was not feasible (Kick, 1985). Detailed gravity measurements were collected by siting gravity stations at 40-ft intervals and obtaining detailed elevation information that was necessary for the elevation correction at these stations. After corrections were applied to the gravity measurements, the Bouguer gravity values were plotted against distance (Figure 14.13a). A regional gravity trend, obtained from a smoothed version of the data in conjunction with the state gravity map, is also shown in Figure 14.13a. This regional trend was removed from the gravity curve to enhance the local effects such as those caused from soil thickness variations; the new data are called the residual gravity data. The interpretation of the residual gravity data is shown in Figure 14.13b. This figure demonstrates that, in all cases except one, the bedrock depths interpreted from the residual gravity data are close to those confirmed by borehole data.

Many more excellent examples of hydrogeological mapping using geophysical datasets are given at the regional scale by Paine and Minty (2005), at the local scale in the saturated sediments by Hyndman and Tronicke (2005), and at the local scale within the vadose zone by Daniels et al. (2005). It is useful to note that most hydrogeological mapping investigations involving geophysical data provide qualitative information about the location of key boundaries or units. Exceptions to this trend include the quantitative studies by Copty et al. (1993) and Hubbard et al. (2005), who used Bayesian techniques to integrate wellbore and surface-based geophysical information for the estimation of lithology and water average water content, respectively.

14.3.2 Hydrogeological Parameter Estimation

To understand, remediate, predict, or manage hydrological systems, information about the states, flowpaths, and fluxes of water through the system is often required, as is information about the controlling characteristics, such as the lithology, porosity, or hydraulic conductivity. For example, mathematical models are used to simulate contaminant plume transport over time for designing effective remediation plans, and to simulate aquifer depletion vs. recharge over time and for designing sustainable resource extraction programs. Such mathematical modeling requires accurate and sufficient hydrological parameter input. Given the difficulty of obtaining sufficient characterization data for model parameterization at

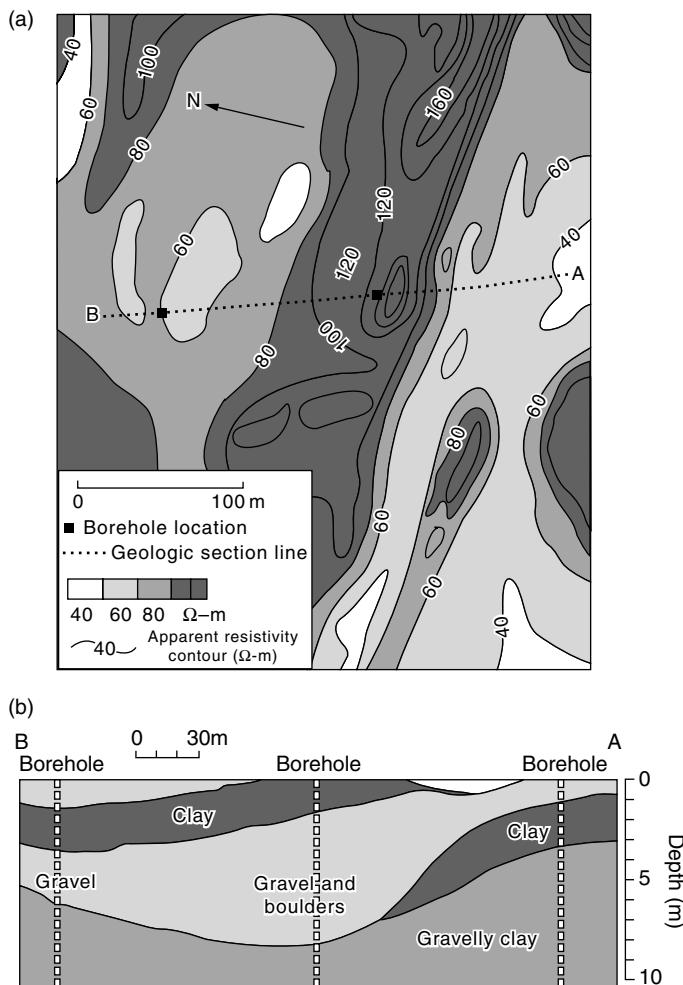


FIGURE 14.11 Resistivity survey used to delineate lateral and vertical variations in subsurface stratigraphy. (a) Contour map produced from resistivity measurements collected using many constant-spread Wenner profiles with an “a” spacing of 6.1 m. (From Zodhy A. A., G. P. Eaton, and D. R. Mayhey, 1974, Applications of surface geophysics to groundwater investigations. Techniques of Water Resources Investigations of the U.S. Geological Survey, Book 2, Chapter D1 and Burger, H. R. 1992, *Exploration Geophysics of the Shallow Subsurface*. Prentice Hall, Englewood Cliffs, NJ. With permission.) The contours of equal apparent resistivity delineate an approximately east–west trending apparent resistivity high. (b) A geologic cross section (BA) based on four electrical resistivity soundings, the apparent resistivity information from the profiles, and from three boreholes whose locations are shown above the cross section. The cross section reveals that the high-resistivity trend shown in (a) is a zone of gravel and boulders that define the location of a buried stream channel.

a reasonable spatial scale using conventional datasets, it is generally felt that computational flow and transport abilities have surpassed the ability to adequately parameterize the synthetic aquifer domains. As a result, the predictive capability of these models is limited, and a significant disconnect has developed between the state-of-the-art associated with modeling and the use of these models to manage practical problems on a regular basis.

As geophysical data offer the potential to provide high resolution information about the controlling hydrogeological parameters in a minimally invasive fashion, they have a large role to play in the interrogation and management of hydrological systems. The ability to estimate subsurface hydrogeological properties using a particular geophysical approach is a function of the relation between the property

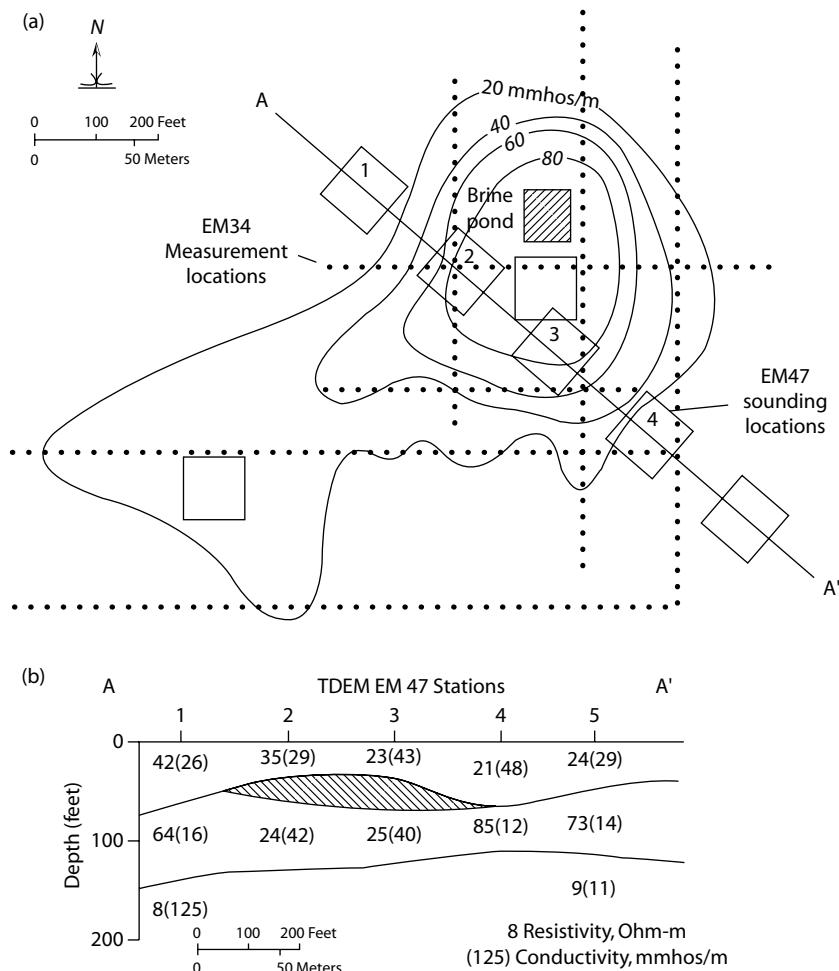


FIGURE 14.12 (a) Apparent conductivity contour map over a brine evaporation pit interpreted from EM profiles. (b) Geoelectric cross section along A-A' traverse interpreted from five EM soundings whose locations are indicated by the five numbered boxes shown in (a). (From Hoekstra, P., R. Lahti, J. Hild, R. C. Bates, and D. Phillips, 1992, *Groundwater Monitoring Rev.*, 12(4), 110–117. With permission.)

under consideration and geophysical attributes, the availability of direct measurements for calibration, the magnitude of the property variations and contrasts, and the spatial scale of the changes relative to the scale of the objective target. The majority of the hydrogeophysical studies that have focused on hydrogeological parameter estimation have been performed using crosshole acquisition geometries because they provide high-resolution information about geophysical attributes with a support scale that is often comparable to direct borehole measurements, thereby minimizing the scale matching problem.

Different approaches have been used for obtaining estimates of hydrogeological parameters using geophysical data. In some cases, the geophysical attributes can be directly “transferred” to estimates of hydrogeological parameters using a published petrophysical relationship or a relationship that has been developed specifically for the site under investigation. In recent years, several studies have developed stochastic frameworks for using geophysical data in hydrogeological site characterization as will be described in Section 14.4. Geostatistical approaches for fusing hydrogeological and geophysical data have been illustrated by many investigators (e.g., Poeter et al., 1997; Yeh et al., 2002). Bayesian approaches for estimating conductivities using head and conductivity measurements (as well as seismic data) have been reported by Rubin et al. (1992), Carty et al. (1993), Carty and Rubin (1995), Chen et al. (2001),

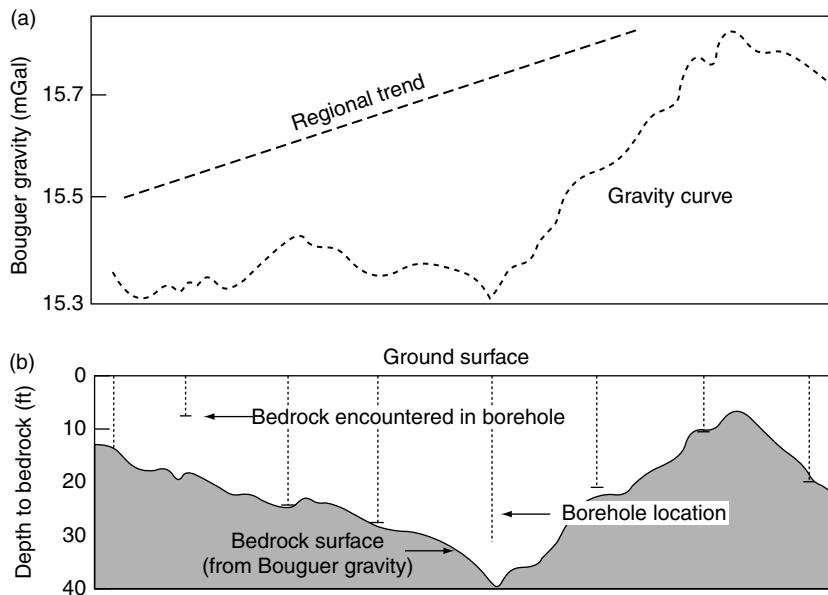


FIGURE 14.13 Interpretation of gravity data to estimate the depth to bedrock (Kick, 1985). (a) Bouguer gravity curve obtained from reduced gravity measurements and regional gravity trend. (b) Interpretation of the local gravity data after removal of the regional trend showing that the bedrock depths interpreted from the gravity data are similar to those confirmed by borehole information. (From Kick, J. F., 1985, Depth to bedrock using gravimetry, geophysics: the leading edge of exploration, *Soc. Exploration Geophysicists*, 4(4), 38–42. With permission.)

and Chen and Rubin (2003). Hubbard et al. (1997b) employed a Bayesian technique to incorporate GPR into mapping the spatial distribution of saturation and permeability as will be described in Section 14.4. Knight et al. (1997) and Hubbard et al. (1999) combined geophysical and hydrogeological data for estimating the spatial correlation structure of the log-permeability. An application of these concepts to estimate hydraulic conductivity was given by Hubbard et al. (2001) and will be described in Section 14.4. Purvance and Andricevic (2000) estimated the spatial distribution of the hydraulic conductivity and its spectral density using electrical and hydraulic conductivities. Day-Lewis et al. (2000) reported an application of the use of geophysical methods for investigating fractured terrains. Hyndman et al. (1994) and Hyndman and Harris (1996) used stochastic techniques and geophysical data to delineate hydrogeological units, and Hyndman and Gorelick (1996) used a stochastic approach to jointly invert geophysical and tracer test data. More recently, other investigators have experimented with the use of the sampling-based Markov chain Monte Carlo (MCMC) method to integrate often complex and nonlinear geophysical and hydrogeological data sets. For example, the MCMC approach was used by Chen et al. (2005) for estimating sediment geochemical parameters, such as Fe(II) and Fe(III), as well as lithology, using GPR crosshole attenuation data through the dependence of both radar attenuation and sediment geochemistry on lithology. Results from this study are illustrated in Figure 14.14. Estimates such as these can be very useful in parameterizing reactive flow and transport models used for predicting contaminant or bacterial transport, or for assessing redox conditions. Other hydrogeophysical estimation studies are described in Hyndman and Tronicke (2005), Daniels et al. (2005), and Linde et al. (2005). Although predominantly performed at the “local” scale, these recent studies highlight the promise of geophysical data for providing quantitative hydrogeological parameter estimates. Section 14.4 will further explore the potential of stochastic methods for providing the systematic framework needed to handle complexities that sometimes arise in fusing disparate hydrogeological–geophysical data sets, such as those associated with spatial variability, measurement error, model discrimination, and conceptual model uncertainty.

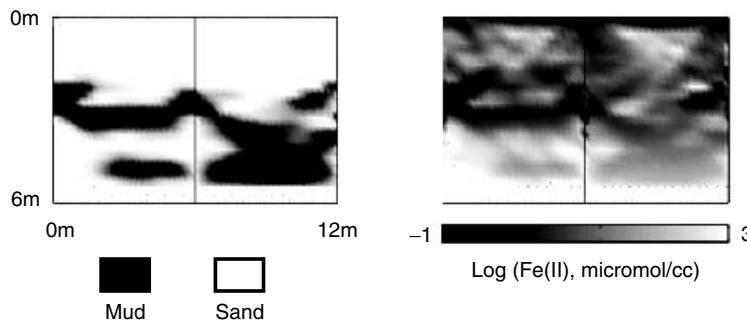


FIGURE 14.14 Estimates of lithology (left) and mean sediment Fe(II) distribution (right) obtained using tomographic GPR amplitude data within a Markov chain Monte Carlo approach (modified from Chen et al., 2005, *Water Resour. Res.*, 40, W12412).

14.3.3 Hydrological Process Monitoring

In addition to geophysical data collected at one point in time to characterize systems in a static sense, geophysical data can be collected at the same location as a function of time. Observing the data as “time difference” sections (measurements collected at an earlier time subtracted from measurements collected at a later time), or “time-lapse” sections, enhances the imaging of subtle geophysical attribute changes caused by natural or forced system perturbations. Time-lapse imaging also decreases the dependence of the geophysical measurements on both the static geological heterogeneities and on data inversion procedures and associated artifacts. Difference images are most commonly obtained from crosshole data sets, but have also been obtained using measurements acquired with surface, airborne, and single borehole acquisition modes. These data sets have been used to elucidate dynamic transformations and thus have the potential to assist in understanding hydrological or remediation processes, such as monitoring infiltration or hydraulic fracturing, imaging injected fluids such as steam or remediation amendments, and tracking the spread of contaminant plumes.

An example is given in Figure 14.15, which illustrates the use of time-lapse, crosshole 100 MHz GPR velocity data collected over a period of a year at an important aquifer in the United Kingdom (from Binley et al., 2002). These estimates show how high rainfall, which occurred in March 1999, increased the subsurface moisture content. This figure also shows the emergence of a drying front in the fall and winter of 1999, and the influence of an impeding layer (that exists between 6 and 7 m depth) on the infiltrating water front. Monitoring using such approaches is extremely useful for understanding complex hydrological processes. The use of noninvasive or minimally invasive geophysical measurements such as these, collected over a prolonged time period, will likely become a standard monitoring approach for improving our understanding of subsurface fluxes and our ability to manage water resources.

14.4 Stochastic Forward and Inverse Modeling Using Hydrogeophysical Datasets

Successful integration of geophysical and hydrogeological datasets represents a recent and major breakthrough in hydrogeological site characterization. As was previously discussed, the value of integrating these datasets for characterization lies in the extensive spatial coverage offered by geophysical techniques and in their ability to sample the subsurface in a minimally invasive manner. However, this breakthrough is associated with a few difficulties. One difficulty resides in the non-unique relationships that sometimes exist between hydrogeological and geophysical attributes; integration of hydrogeological and geophysical data under non-unique conditions has been investigated by Rubin et al. (1992), Copty et al. (1993), Hubbard et al. (1997b), Kowalsky et al. (2004b, 2004c, 2005), Chen et al. (2005), Hou and Rubin (2005), and Hubbard and Rubin (2000). This non-uniqueness can exist even under idealized conditions of error-free

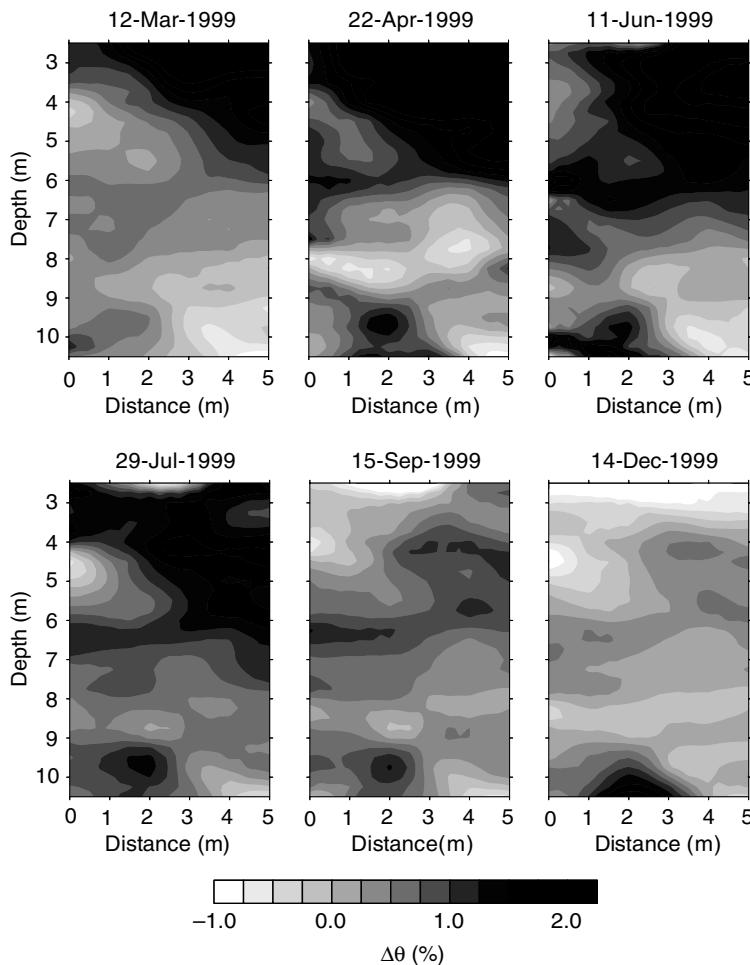


FIGURE 14.15 Percent change in volumetric water content obtained over time using time-lapse GPR velocity information associated with natural precipitation and drying. Changes in moisture content are shown relative to those estimates obtained using GPR data collected on February 1, 1999 (modified from Binley et al., 2002, *J. Hydrology*, 267(3–4), 160–172).

measurements in natural systems comprised of multiple hydrogeologically significant units (i.e., Prasad, 2003), and it is only exacerbated by measurement errors. In applications, the situation becomes even more difficult because the rock type at the location associated with the geophysical attribute is almost always unknown, and thus the applicable petrophysical model is also almost always unknown. Another difficulty stems from the disparity between the spatial resolution of the geophysical attributes and the scale that characterizes the hydrogeological attributes, collected for example, through boreholes (cf. Ezzedine et al., 1999). This scale disparity hinders efforts to develop unique and accurate relations between the two types of measurements, and introduces another source of uncertainty. It is obvious that the characterization techniques developed under such conditions must recognize these difficulties, and furthermore, that they must quantify their effects on estimation uncertainty. This leads us to formulate our data integration techniques under the very wide umbrella of stochastic techniques (Rubin, 2003). In this section, we shall present different concepts and ideas for dealing with these issues. The commonality among these concepts is in making uncertainty and spatial variability an integral part of the solution, in consistently attempting to address the challenge of estimation uncertainty, and in minimizing arbitrariness and subjectivity.

In this section, we start by setting a general framework for forward modeling under conditions of uncertainty. This is done by casting the hydrogeological prediction problem in a stochastic framework, and thus opening the door for estimating variables such as pressure or concentration in terms of their expected values and estimation variances, at a minimum, or through their probability distribution functions, for a more complete characterization of uncertainty. In hydrogeological applications, we can distinguish between dependent variables, such as concentration or soil moisture, and independent ones, such as hydraulic conductivity. The dependent variables are related, often through complex and nonlinear relationships, to the independent ones. Through these relationships, the uncertainty in estimating the independent variables can be propagated to the dependent variables. This error propagation is the essence of forward modeling: the goal is to express the uncertainty in the dependent variables as a function of the uncertainty in all the variables that affect them. Our discussion will continue by focusing on methods for estimating the uncertainty in the independent variables. The goal here is to model the effects of spatial variability and data scarcity, while utilizing data obtained from conventional hydrogeological testing as well as data from geophysical surveys. The goal is also to utilize data of different quality, starting with precise measurements and ending with information of a more qualitative nature, commonly referred to as “soft” data. We refer to this as hydrogeophysical inverse modeling. We shall discuss alternative approaches for inverse modeling, such as the maximum likelihood and maximum a posteriori methods. By emphasizing their origin in Bayes’ theorem, we will be able to present entropy-based concepts intended to improve their accuracy and generality.

14.4.1 Forward Modeling with Parameter Error

Let us consider a random variable Z . This variable can be an independent variable, such as permeability or porosity, or a dependent variable, such as pressure head, concentration, or soil moisture. To model its uncertainty, we opt to define Z as a random variable and strive to model it through a probability distribution function (pdf), or through its statistical moments. The variable Z is characterized through its pdf $f_Z(z|\underline{\theta})$, where $\underline{\theta}$ is a vector of the parameters. This vector may include the parameters of spatial correlation models (e.g., a semivariogram model) employed for conditioning the pdf on measurements, as well as parameters of petrophysical, correlation, or numerical models employed for relating between Z and other attributes or with primitive variables.

Defined as a random variable, Z can be characterized through its moments. The N th moment of Z is given by:

$$\langle Z^N | \underline{\theta} \rangle = \int_{-\infty}^{\infty} z^N f_Z(z|\underline{\theta}) dz \quad (14.10)$$

The moments of Z conditional to measurements can be obtained by replacing $f_Z(z|\underline{\theta})$ with the conditional pdf $f_Z(z|\underline{\theta}, \{M\})$, where $\{M\}$ is the vector of measurements. For a given vector $\underline{\theta}$, $f_Z(z|\underline{\theta})$ or $f_Z(z|\underline{\theta}, \{M\})$ can be obtained, for example, through Monte Carlo simulations (Rubin, 2003).

To account for parameter error, the parameter vector $\underline{\theta}$ is viewed as a realization of a random process characterized by the p -variate pdf $f_{\underline{\theta}}(\theta_1, \dots, \theta_p)$, where p is the number of parameters in $\underline{\theta}$. The expected value in Equation 14.10 must be taken over by all possible combinations of parameters, weighted by their probabilities:

$$\langle Z^N \rangle = \int_{\theta} \int_{-\infty}^{\infty} z^N f_{Z,\underline{\theta}}(z, \underline{\theta}) dz d^p \underline{\theta} = \int_{\theta} \int_{-\infty}^{\infty} z^N f_Z(z|\underline{\theta}) f_{\underline{\theta}}(\theta_1, \dots, \theta_p) dz d^p \underline{\theta} \quad (14.11)$$

where \int_{θ} implies integration over the entire vector space of $\underline{\theta}$, and $f_{Z,\underline{\theta}}$ is the joint pdf of Z and $\underline{\theta}$. The second equality in Equation 14.11 is based on Bayes’ theorem. Integrating over the parameter space yields

the marginal pdf of Z in the joint attribute-parameter space. This marginal is expected to be defined over a broader range than the one defined by the pdf of Z conditional on a given set of parameters, and it will lead to a larger variance, which reflects the effects of the parameter error. It is imperative to formulate inverse modeling strategies that can quantify the estimation errors of the parameters $\underline{\theta}$; an extensive discussion of various error types is provided in Rubin (2003, Chapter 13).

The integration in Equation 14.11 calls for a systematic screening of the multidimensional parameter space, which can be quite tedious. Some simplification is possible using the Monte Carlo integration. Applications of Equation 14.11 are discussed in Rubin and Dagan (1992), Maxwell et al. (1999), Rubin (2003), and Hou and Rubin (2005).

14.4.2 Inverse Modeling

This section discusses different strategies for determining the statistical distributions of hydrogeological parameters and their estimation errors. It is common to employ direct measurements of the attribute for this purpose, as well as measurements that can be related to the attribute using forward models, and hence the name “inverse modeling.” For example, inverse modeling of the spatial distribution of hydraulic conductivity may involve hydraulic conductivity and head measurements, and in that case will rely on the flow equation to relate between the two. Or it may employ a geophysical signal of some sort, and in this case will employ an equation for modeling the propagation of that signal and possibly petrophysical models. As our understanding of inverse modeling has evolved over time, it has become obvious that its success depends on the ability to incorporate information of all types — either “hard” information, such as direct measurements, or “soft” information, such as expert opinions given possibly in the form of pdfs, geophysical measurements, or information borrowed from geologically similar formations, used as preliminary (or prior) information.

Stochastic inverse modeling is a formulation of inverse modeling in a probabilistic framework. Employing a stochastic paradigm eliminates the problem of non-uniqueness (Carrera and Neuman, 1986a, 1986c) and introduces a rational approach for modeling uncertainty (Rubin, 2003). Stochastic inverse models are commonly based on Bayes’ theorem. Bayesian-related methods commonly include a statement of a prior pdf of the parameters, and a likelihood function that relates the attribute to the output of the forward model. Using the prior and the likelihood function, Bayes’ theorem is then employed to define a posterior pdf (or a posteriori pdf). Bayes’ theorem serves to update the plausibility of a proposition as the state of information changes because of the availability of new data. The Bayesian framework is broad enough to allow for a remarkably large number of different interpretations.

As prior information is an important component of Bayes’ theorem, we shall start our discussion by considering methods for defining minimally subjective prior pdfs. This will help us in understanding the significance and implications of choosing a model for the prior pdf. Armed with this knowledge, we shall proceed to discuss the maximum likelihood (ML) and maximum a posteriori (MAP) methods, which offer alternative interpretation of Bayes’ theorem.

A complementary classification of the various methods that we shall present is as follows (1) The sequential approach, where the focus is on defining a pdf structure and identifying its parameters, and subsequently applying it for point estimation; (2) methods that assume a pdf structure and focus on identifying its parameters, sometimes simultaneously with spatial (point) estimation.

14.4.2.1 Bayesian Estimators

A general statement of Bayes’ theorem must include the prior and posterior pdfs. The starting point is of course the prior pdf. In this section, we present several different approaches for estimating the prior pdf.

Let the pdf $g_{\underline{\theta}}(\underline{\theta}|I)$ denote the prior pdf of the vector of parameters $\underline{\theta}$, which summarizes, in a probabilistic manner, our understanding of the values $\underline{\theta}$, and can assume direct or indirect measurements of these parameters based on other sources of information. Such information may include expert opinions or information borrowed from geologically similar sites (cf. Chapters 2, 3, and 13 of Rubin, 2003). As soon as measurements at the site become available, they can be used for updating the prior pdf (Hou and

Rubin, 2005). This can be done by viewing the measurements as realizations of a space random function (SRF) X . The pdf of $\underline{\theta}$ conditioned on the vector of N measurements, \mathbf{X} , where $X_1 = x_1, \dots, X_N = x_N$, denoted by $f_{\underline{\Theta}|\mathbf{X}_1, \dots, \mathbf{X}_N}(\underline{\theta} | x_1, \dots, x_N, I)$, is called the posterior distribution of $\underline{\Theta}$, and it is given by:

$$f_{\underline{\Theta}|\mathbf{X}_1, \dots, \mathbf{X}_N, I}(\underline{\theta} | x_1, \dots, x_N, I) = \frac{f_{X_1, \dots, X_N|\underline{\Theta}, I}(x_1, \dots, x_N | \underline{\theta}, I) g_{\underline{\Theta}|I}(\underline{\theta} | I)}{f_{X_1, \dots, X_N|I}(x_1, \dots, x_N | I)} \quad (14.12)$$

where $f_{X_1, \dots, X_N}(x_1, \dots, x_N | I)$ is the marginal of the joint pdf of \mathbf{X} and $\underline{\Theta}$. In cases where several types of data are available, the vector \mathbf{X} should be redefined accordingly, for example X_i , $i = 1, \dots, M$, for the M measurements of one attribute, and X_i , $i = 1, \dots, M + 1, \dots, N$, for the measurements of the other type.

The posterior distribution incorporates all information available, including measurements and prior knowledge. The function $f_{X_1, \dots, X_N|\underline{\Theta}, I}(x_1, \dots, x_N | \underline{\theta}, I)$ is the probability of making the observations for a given $\underline{\Theta} = \underline{\theta}$, and it is commonly referred to as the likelihood function. If we want to estimate $\underline{\theta}$, we need to adopt a decision rule, for example, we may take for estimator that $\underline{\theta}$ that maximizes the posterior distribution, or we could just estimate $\underline{\theta}$ with the mean or median of the posterior distribution. None of these choices reflects the width of the distribution. Rather, the pdf of the parameters should be considered in its entirety, and not just through a subset.

14.4.2.2 A Systematic Approach to Defining a Prior Probability

How do we estimate the prior pdf of unknown model parameters, $g_{\underline{\Theta}|I}(\underline{\theta} | I)$? The determination of the statistical properties of the unknown model parameters, and at a more fundamental level, the statistical model itself, are the major impediments to the effective use of probabilistic models. It is common to assume a model for the prior, and then to update it as additional information becomes available. But one can argue that the choice of the prior model is arbitrary. How significant is the choice of a prior? Jaynes (1968) stated, “Since the time of Laplace, applications of probability theory have been hampered by the difficulties in the treatment of prior information,” and “Therefore, the logical foundations of decision theory cannot be put in fully satisfactory form until the old problem of arbitrariness (sometimes called ‘subjectiveness’) in assigning prior probabilities is resolved.”

Instead of postulating a prior pdf, we can infer it from measurements based on assumptions or axioms. One such approach is based on the principle of minimum relative entropy (MRE), the application of which to hydrogeology was pioneered by A. Woodbury and coworkers. The MRE approach (Woodbury and Rubin, 2000; Woodbury and Ulrych, 2000; Rubin, 2003) couples between Bayes’ theorem and entropy measures of information. It represents the general approach of inferring a probability distribution from constraints that incompletely characterize that distribution (Hou and Rubin, 2005). These constraints can be (but are not limited to be) in the form of averages or higher-order moments, following Equation 14.10. In the context of groundwater, such constraints can be, for example, the average and variance of the hydraulic conductivity obtained from a limited number of tests, from empirical determination, or from an informed guess.

Entropy is a measure of the amount of uncertainty in a pdf. The principle of MRE states that of all the probabilities that satisfy the given constraints, choose the one that has the highest entropy with respect to a known prior, which is the one that is the most uncommitted, or the least prejudiced, with respect to unknown information. The model identified based on MRE can be interpreted as the one which fits all the constraints and the one which has the greatest “spreadout” that can be realized in the greatest number of ways. The pdf thus identified can be used as priors for any of the inverse methods described below.

14.4.2.3 The Maximum Likelihood and Maximum A Posteriori Concepts in Inverse Modeling

In this section, we examine and formulate the inverse problem based on ML and MAP concepts. Both methods are based on Bayes’ theorem, but with different interpretations assigned to the prior and to the inferred parameter values.

14.4.2.3.1 The Maximum Likelihood Method

The ML formulation discussed below employs only information available from measurements and ignores prior information. In the absence of any prior information I , the likelihood function defined in Equation 14.12 becomes $f_{X_1, \dots, X_N | \underline{\Theta}}(x_1, \dots, x_N | \underline{\theta})$. Let us denote it by $L(\underline{\theta} | x_1, \dots, x_N)$. L gives the likelihood (or in fact, the pdf) that the random variables X_1, \dots, X_N , assume the particular values x_1, \dots, x_n . X_i can denote, for example, the log-conductivity (with $X_i = X(\mathbf{b}_i)$ being the log-conductivity at location \mathbf{b}_i) and $\underline{\Theta}$, which will include the parameters of the X SRF (such as its mean, variance, and the parameters of its correlation function). For a given vector $\underline{\theta}$, the particular values of the random variables most likely to occur are the values x_1, \dots, x_N that maximize $f_{X_1, \dots, X_N | \underline{\Theta}}(x_1, \dots, x_N | \underline{\theta})$, or in fact, those values that define its mode.

Let us now consider a particular vector of observations x_1, \dots, x_N . We wish to find the particular combination of parameters of $\underline{\Theta}$, denoted by $\hat{\underline{\theta}}$, which maximizes the likelihood function $L(\underline{\theta} | x_1, \dots, x_N)$ or, in other words, maximizes the probability of observing X_1, \dots, X_N , equal to x_1, \dots, x_N , respectively. This vector, $\hat{\underline{\theta}}$, which is a function of the measurement vector x_1, \dots, x_N , is called the maximum likelihood estimator of $\underline{\Theta}$. The maximum likelihood estimator is the solution for the set of equations:

$$\frac{\partial L(\underline{\theta})}{\partial \theta_j} = 0, \quad j = 1, \dots, p \quad (14.13)$$

where p is the number of parameters in $\underline{\Theta}$. Both the logarithm of L and L have their maxima at the same value of $\underline{\Theta}$, and it is sometimes easier to find the maximum of the logarithm of the likelihood (or the minimum of the negative log-likelihood).

The pioneers in bringing the concept of maximum likelihood to hydrogeology are Peter Kitanidis and his coworkers. The foundations for their work were laid in Kitanidis and Vomvoris (1983), followed by Hoeksema and Kitanidis (1984, 1985), where they considered the problem of identifying the spatial distribution of the log conductivity.

One application of these ideas in hydrogeophysics is provided by Hubbard et al. (1997b). It demonstrates the power of using Bayesian concepts in the case where non-unique petrophysical relationships exist. In that study, numerical experiments were performed to investigate the utility of ground-penetrating radar (GPR) data for saturation and permeability estimation under a range of hydrogeological conditions. As described in Section 14.2.3, GPR is a noninvasive, high-resolution geophysical method, and estimates of dielectric constants can be extracted from both surface and crosshole GPR data. Petrophysical relations can then be used to transfer these dielectric constants into saturation or permeability estimates.

Petrophysical curves that illustrate a plausible relationship of saturation and dielectric constant as a function of facies (and corresponding permeability) for a synthetic sand–clay system are shown in Figure 14.16. These curves illustrate the non-uniqueness that sometimes exists when trying to map a geophysical attribute (such as dielectric constant obtained from radar data) into a hydrological estimate (such as saturation or facies permeability). For example, in the case study shown in Figure 14.16, for each dielectric

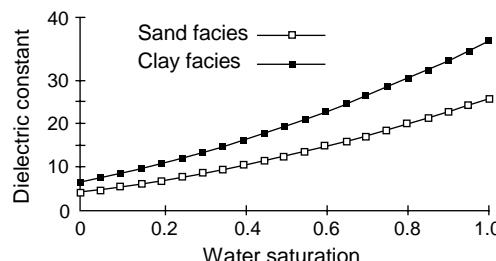


FIGURE 14.16 Example of a non-unique petrophysical relationship between dielectric constant and water content in a sand–clay system (modified from Hubbard et al., 1997, and reduced from Knoll et al., 1995 shown in Figure 14.4.)

constant value obtained from the radar data in the bimodal system under consideration, two values of saturation (S_1 and S_2) are possible, depending on whether the facies at that subsurface location is a clay or a sand, both with associated distinct permeability values. With two possible values for the saturation, the pdf of the saturation S can be expressed as follows:

$$f_S(s | S_1, S_2) = p\delta(s - S_1) + (1 - p)\delta(s - S_2) \quad (14.14)$$

where p denotes the probability to observe $s = S_1$, and δ is the Dirac delta function. The probability is related to the expected value of S , $\langle S \rangle$, through the relationship (Hubbard et al., 1997b, Equation 9):

$$p = \frac{\langle S \rangle - S_2}{S_1 - S_2} \quad (14.15)$$

Hence, once $\langle S \rangle$ is defined, the pdf of s is also defined. The saturation S that maximizes the probability to observe the measured dielectric constant is equal to S_1 if p is larger than 0.5, and to S_2 otherwise. The S thus chosen is the ML estimator, because we choose for $\hat{\theta}$ the value of S that maximizes the probability to observe the measured dielectric constant. Hubbard et al. (1997) suggested a couple of alternatives for computing $\langle S \rangle$: the first was based on numerical infiltration simulations, and the second was based on kriging borehole saturation measurements.

The vector $\hat{\theta}$ that minimizes $-\ln L$ is the ML estimator. Once it is determined, we can estimate the variances and covariances of the parameters. We can define the matrix $\underline{\underline{B}}$ of order p , where

$$B_{ij} = \frac{\partial}{\partial \theta_i} \left(\frac{\partial}{\partial \theta_j} \ln L \right) \quad (14.16)$$

which is known as the Fisher information matrix. The matrix $\underline{\underline{\Sigma}} = \underline{\underline{B}}^{-1}$ is referred to as the Cramer–Rao lower bound of the estimation error covariance (Schweppe, 1973), and its (i, j) term provides an estimate for the estimation error's variance–covariance matrix for the vector $\underline{\theta}$. Once the parameters of the X SRF are identified, the spatial distribution of X , conditioned on the measurements, can be determined using geostatistical methods such as kriging (when only X measurements are available) and co-kriging (when other types of measurements are also available).

14.4.2.3.2 The Maximum A Posteriori Approach

MAP extends ML to situations where prior information does exist. This extension provides for a wide range of alternative interpretations and approaches. The MAP concept is illustrated in Figure 14.17. Figure 14.17a schematically shows an example of a prior log-conductivity pdf at a single subsurface location. Figure 14.17b illustrates the joint distribution as well as the likelihood function. The geophysical data are given in the form of a precise measurement (solid line) and, alternatively, in the form of a distribution, representing imprecise data (dashed line). Uncertainty in the geophysical attribute can stem from many factors, such as the choice of tomographic data inversion approach and parameters, acquisition and data reduction errors, and variable coverage/sensitivities (cf. Peterson, 2001; Alumbaugh et al., 2002; Binley et al., 2002). In the first case, the likelihood function is defined over a much narrower range of conductivity values, which will limit the range of physically plausible conductivity values.

Using the likelihood function for the precise information, the prior pdf can be updated using Bayes' theorem (Equation 14.12), to obtain the posterior pdf, as illustrated by the solid line in Figure 14.17c. We note that the high quality of the geophysical data renders this posterior pdf considerably sharper than the prior pdf (compare solid-line log-conductivity distributions in Figure 14.17a and 14.17c), because Bayes' theorem indicates that reconciliation between the prior pdf and the geophysical attribute is possible by assigning a larger probability to the higher log conductivities. In the case of imprecise geophysical data, as illustrated by the geophysical attribute distribution in Figure 14.17b, the improvement in the posterior pdf (shown by the dashed line in Figure 14.17c) is not as significant as when precise geophysical data are used.

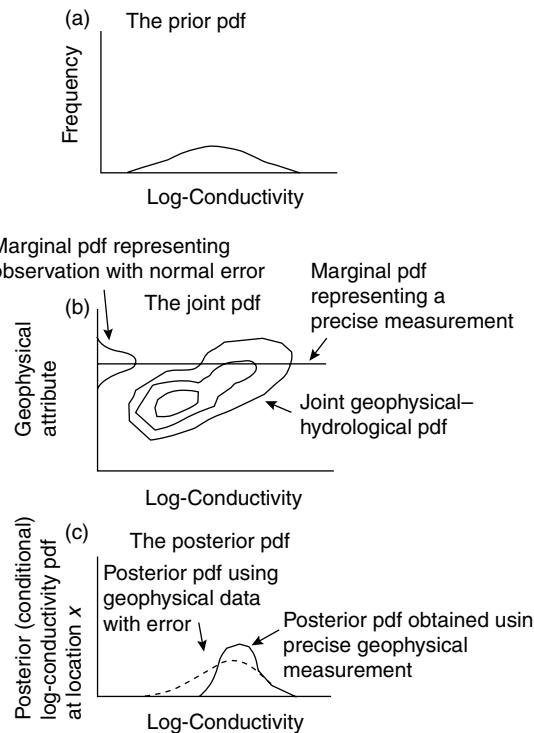


FIGURE 14.17 Example of the MAP approach to hydrogeophysical data fusion: (a) example of a prior log-conductivity pdf at a single subsurface location; (b) example of the joint geophysical–hydrological pdf, as well as marginal pdf representing a geophysical observation with normal error (dashed line) and marginal pdf representing a precise measurement (solid line); (c) posterior pdfs at a single subsurface location obtained by fusing the prior log-conductivity pdf, joint pdf, and marginal geophysical pdf within the MAP formalism. The solid posterior pdf was obtained using the precise geophysical measurement, while the dashed posterior pdf was obtained using the geophysical observation with normal error (shown in Figure 14.17b).

The posterior pdf is not as sharp because the improvement of the posterior compared to the prior depends on the amount of information provided by the geophysical data. This example illustrates the essence of MAP as a “reconciliation” between prior information, usually in the form of direct measurements of the attribute of interest, and additional information of varying quality. The information used for updating the prior pdf can be either empirically correlated or functionally related to the attribute of interest.

For discussion of the mathematical aspects of MAP, let us consider the case of identifying the spatial distribution for hydraulic conductivity based on measurements of conductivity as well as auxiliary information denoted by h . Such information can obviously include geophysical information collected through various types of geophysical surveys. Let us denote a finite-sized vector of conductivities (or log conductivities, which is more commonly employed) and h measurements distributed over the aquifer by $z = (z_h, z_a)$, with z_a denoting the vector of conductivity measurements and z_h denoting the h measurements. Let us now denote, by the vector of order p , the conductivities that we wish to estimate, for example, at the nodes of a numerical grid. Let us define the functional relationship between z_a and a in the form $z_a = \mathbf{F}_a(a) + \mathbf{v}_a$, where $z_{a,i}$ and $v_{a,i}$ denote the i th conductivity measurement and the measurement error, respectively. \mathbf{F}_a represents the large-scale trend of the conductivity, given for example in the form of a spatial distribution of the conductivities over a grid or by a set of polynomials. Similarly, let us define the measurements vector $z_h = \mathbf{F}_h(a) + \mathbf{v}_h$, where \mathbf{F}_h represents the large-scale trend of the h , with $v_{h,i}$ denoting the measurement error at x , owing to instrumentation and recording error, data reduction procedures, and the small-scale fluctuations of the actual point values from the corresponding large-scale trends. Note that \mathbf{F}_a and \mathbf{F}_h may include parameters whose values are unknown and need to be

identified. The formulation below can be used to identify \mathbf{a} at unknown locations, as well as the unknown parameter values, although in our discussion below we shall concentrate on \mathbf{a} . It is common to assume that the measurement errors are zero mean, and that \mathbf{v} and \mathbf{a} are uncorrelated. These two relationships can be summarized by the following measurement equation (cf. McLaughlin and Townley, 1996):

$$\begin{bmatrix} z_h \\ z_a \end{bmatrix} = \begin{bmatrix} \mathbf{F}_h(\mathbf{a}) \\ \mathbf{F}_a(\mathbf{a}) \end{bmatrix} + \begin{bmatrix} \mathbf{v}_h \\ \mathbf{v}_a \end{bmatrix} = \mathbf{F}(\mathbf{a}) + \mathbf{v} \quad (14.17)$$

The MAP estimator of \mathbf{a} , $\hat{\mathbf{a}}$, is the one most likely given the measurements (unlike ML, which seeks \mathbf{a} that maximizes the probability of observing the measurements). It is defined based on Bayes' theorem by the posterior pdf:

$$f_{a|z}(\mathbf{a}|z) = \frac{f_{z|\mathbf{a}}(z|\mathbf{a})f_a(\mathbf{a})}{f_z(z)} = \frac{f_v(z - \mathbf{F}(\mathbf{a}))f_a(\mathbf{a})}{f_z(z)} \quad (14.18)$$

where f_a and $f_{a|z}$ are the prior and posterior pdfs of \mathbf{a} , $f_{z|\mathbf{a}}$ is the conditional pdf of \mathbf{a} given z , and f_z and f_v are the measurement and measurement-error pdfs, respectively. The second equality in Equation 14.18 follows from Equation 14.17 and from the independence of \mathbf{a} and \mathbf{v} . The MAP estimate $\hat{\mathbf{a}}$ corresponds to the mode of $f_{a|z}$.

If f_a and f_v are multinormal with covariance matrices $\underline{\underline{c}_a}$ and $\underline{\underline{c}_v}$, the posterior pdf becomes:

$$f_{a|z}(\mathbf{a}|z) = c(z) \exp \left\{ -\frac{1}{2} [\mathbf{z} - \mathbf{F}(\mathbf{a})]^T \underline{\underline{c}_v}^{-1} [\mathbf{z} - \mathbf{F}(\mathbf{a})] \right\} \exp \left\{ -\frac{1}{2} [\mathbf{a} - \bar{\mathbf{a}}]^T \underline{\underline{c}_a}^{-1} [\mathbf{a} - \bar{\mathbf{a}}] \right\} \quad (14.19)$$

where $c(z)$ is a normalization factor not depending on \mathbf{a} , and $\bar{\mathbf{a}}$ is the prior mean of \mathbf{a} .

Assuming \mathbf{v} to be normal is quite common, because it is the simplest model, requiring to state only mean and variance. However, it turns out that it is also the safest assumption, if all that is known about \mathbf{v} is its mean and variance, since as can be shown using MRE considerations, normal distribution in that case is the least informative, and hence is minimally subjective. If the variance is not known precisely, Equation 14.19 can be written as a pdf conditional on a given variance σ_v : $f_{a|z,\sigma_v}(\mathbf{a}|z, \sigma_v)$, and the dependence on σ_v can be eliminated by taking the expected value over the range of σ_v (see Woodbury and Rubin, 2000; Woodbury and Ulrych, 2000 and Hou and Rubin, 2005).

It is convenient to find the MAP estimate by minimizing $-2\ln[f_{a|z}]$. The MAP estimate $\hat{\mathbf{a}}$ is then the \mathbf{a} that minimizes the weighted-least-squares criterion (Schwepppe, 1973; Carrera and Neuman, 1986a, 1986b; Tarantola, 1987):

$$[\mathbf{z} - \mathbf{F}(\mathbf{a})]^T \underline{\underline{c}_v}^{-1} [\mathbf{z} - \mathbf{F}(\mathbf{a})] + [\mathbf{a} - \bar{\mathbf{a}}]^T \underline{\underline{c}_a}^{-1} [\mathbf{a} - \bar{\mathbf{a}}] \quad (14.20)$$

which yields a compromise between the prior estimate (represented here by the term on the right) and the weighted-least-squares fit (which is summarized by the first term). Kawalsky et al. (2004b, 2005) showed how to minimize (14.20) and how to generate multiple realizations conditional to measurements.

Numerical solutions for the above minimization problems are generally found with iterative search algorithms such as the Gauss–Newton method. Note that the number of unknown parameters may be large, on the order of the number of grid blocks. This number may be reduced by zonation, which calls for identifying areas of uniform conductivity (thus potentially reducing the number of unknowns), but this procedure is not without its problems, as we shall discuss below.

The concept underlying Equation 14.20 attaches a probabilistic interpretation to \mathbf{v} , which is viewed as a measurement error vector. If instead, we view the \mathbf{z} as SRFs, with \mathbf{F} representing their expected values, then $\mathbf{z} - \mathbf{F}(\mathbf{a})$ contains the local fluctuations of the SRFs from their expected values, and it represents the effects of natural heterogeneity rather than measurement error. Furthermore, the matrices in Equation 14.20 describe now the covariances between SRFs, which can be developed from physical

principles (Rubin, 2003; Chapters 4 and 6). This is the concept underlying the inverse method of Kitanidis and Vomvoris (1983), and Rubin and Dagan (1987a, 1987b, 1988).

Equation 14.20 can be used to identify other formulations of the inverse problem. One formulation is the least squares method. Least squares estimates minimize:

$$[z - \mathbf{F}(\mathbf{a})]^T \underline{\underline{c}_v^{-1}} [z - \mathbf{F}(\mathbf{a})] \quad (14.21)$$

and ignore the prior information term. The least squares estimator obtained by minimizing Equation 14.21 is equivalent to a Gaussian ML estimator with known statistical properties. We note from Equation 17.9 that the likelihood function is:

$$f_{z|a}(z|\mathbf{a}) = f_v[z - \mathbf{F}(\mathbf{a})] \quad (14.22)$$

and that the ML estimator is obtained by minimizing $-2\ln f_{z|a}$. For a Gaussian f_v and a known covariance function, the ML estimator will be obtained by minimizing Equation 14.22.

McLaughlin and Townley (1996) note that ML estimators are equivalent to MAP estimators when the prior covariance matrix $\underline{\underline{c}_a}$ (Equation 14.21) is arbitrarily large, implying that the prior information is uninformative. They note further that an uninformative prior is pessimistic, as we can usually put some bounds on the range of values that parameters such as conductivity can assume. Most regression and ML algorithms provide for such constraints in their search algorithm. Bounded least squares or ML estimators, which assume that f_v is Gaussian and that the prior probability is uniform between the specified bounds, are in fact equivalent to a MAP estimator, where the performance criteria to be minimized is given by:

$$[z - \mathbf{F}(\mathbf{a})]^T \underline{\underline{c}_v^{-1}} [z - \mathbf{F}(\mathbf{a})] - \ln f_a(\mathbf{a}) \quad (14.23)$$

and where the prior pdf f_a is a positive constant when \mathbf{a} lies within the bounds, and 0 otherwise. As this performance criterion imposes an infinite penalty on estimates lying outside the bounds, it will enforce the constraints. Hence, MAP offers flexibility by being able to handle prior information. A MAP generalization of the Fisher information matrix $\underline{\underline{B}}$ (see Equation 14.16) is provided in McLaughlin and Townley (1996) and Rubin (2003).

The error matrices $\underline{\underline{c}_a}$ and $\underline{\underline{c}_v}$ of Equation 14.20 pose a difficult challenge because they are generally not known, at least not a priori. MRE offers ideas on how to address this issue (see our discussion following Equation 14.19). Another way to approach this issue is by defining $\underline{\underline{c}_a}$ and $\underline{\underline{c}_v}$ as diagonal matrices, in the form $\underline{\underline{c}_a} = \sigma_a^2 \underline{\underline{I}}$ and $\underline{\underline{c}_v} = \sigma_v^2 \underline{\underline{I}}$, where σ_a^2 and σ_v^2 are the error variances and $\underline{\underline{I}}$ is the identity matrix. With these definitions, Equation 14.20 becomes proportional to:

$$[z - \mathbf{F}(\mathbf{a})]^T [z - \mathbf{F}(\mathbf{a})] + \beta [\mathbf{a} - \bar{\mathbf{a}}]^T [\mathbf{a} - \bar{\mathbf{a}}] \quad (14.24)$$

where β is the ratio σ_v^2/σ_a^2 , and it can be used to change the relative weights of the two terms in Equation 14.24. A small σ_a^2 indicates a greater confidence in the prior, and hence a larger weighting in the form of a larger β . The parameter β can be determined by numerical experimentation. For example, we can use a training set of data to identify an optimal β . Equation 14.24 can be viewed as the weighted least squares representation of MAP, with the term on the right representing the prior, acting to regularize the solution, and the term on the left representing updating through the acquisition of additional measurements.

Hydrogeophysical inverse modeling based directly or indirectly on Equation 14.18 was included in the work of Ezzedine et al. (1999), who reported on a Bayesian method for using borehole and surface resistivity data for lithology estimation, and in Day-Lewis et al. (2000), who reported an approach for conditioning geostatistical simulation of high-permeability fracture zones in multiple dimensions to hydraulic-connection data inferred from wellbore tests, and in Chen et al. (2001) and Chen and Rubin (2003).

14.4.3 Point Estimation with Known Parameters

The formulations of Section 14.4.2 are commonly applied in two different modes. The first mode calls for identification of nodal values (e.g., conductivity values at nodes over a grid). The second mode calls for defining the parameters of interest as SRFs as a first step and their identification using ML or MAP, which is then followed by mapping the spatial distributions of the parameters using methods such as kriging or co-kriging (see Chapter 16 of this volume or Chapter 3 of Rubin, 2003). The second mode originated in the works of Kitanidis and co-workers and was subsequently pursued by Rubin and Dagan (1987a, 1987b, 1988). One immediate advantage of the second mode is the reduction in the number of target parameters at the inversion stage. In other words, high resolution on the grid does not translate into a large number of target parameters, with its associated difficulties, including non-uniqueness, instability, and non-identifiability (see Carrera and Neuman, 1986a, 1986b, 1986c).

For example, if we are interested in mapping the spatial distributions of the conductivity, then the second mode will include the following two steps: (1) Identifying the parameters of the conductivity (or log-conductivity) SRF, such as its mean and spatial covariance parameters, using all available data; and (2) Mapping the spatial distribution of the conductivity, conditional to all available data, using a statistical interpolation or simulation techniques (Kowalsky et al. 2004b, 2005) technique. Such mapping also includes characterization of uncertainty, which can be accomplished to varying degrees of completeness and accuracy.

As discussed in the previous section, MAP and ML formulations applied in the first mode produce MAP or ML estimates and estimation variances of the target parameters. As we shall show below, separating the identification problem from point estimation, as done in the second-mode applications, allows us to obtain more complete characterization of estimation uncertainty and better flexibility in conditioning the estimates on the various types of information available. Below, we discuss the second mode, and while its first stage was discussed in Section 14.4.2, we focus here on the stage of deriving spatial distributions over the computational grid, with the SRF parameters already known.

14.4.3.1 Bayesian Point Estimators and Conditioning on “Hard” and “Soft” Information

The most straightforward approach to mapping spatial distributions is via kriging or co-kriging. Kriging provides an estimate and a kriging variance (Chapter 3 of Rubin, 2003). We should note the following: (1) The kriging variance can be viewed as a measure of uncertainty, but cannot be used to compute confidence intervals unless the estimation error is normally distributed. While this is often assumed, it is not part of the kriging algorithm. Hence, as kriging estimates are not characterized by any underlying statistical distributions, their characterization of uncertainty is lacking. (2) Kriging methods rely on two-point correlations, in the form of covariances or semi-variograms, to project measurements onto unsampled locations. Such forms of correlation are efficient in capturing linear correlations, but fail in the presence of nonlinearity. Measures to amend for this limitation can be found in nonlinear transformations of the target variable, such as log-transforms or indicator transforms (Chapter 3 in Rubin, 2003), but such transforms are not always efficient in removing nonlinearity. This is particularly true for the case of inherent non-uniqueness in the correlation structure (i.e., stemming from physical principles, not from measurement errors). For example, Rubin et al. (1992) and Copty et al. (1993) considered the ramifications of incorporating information from a non-unique relationship between permeability and seismic velocity in sandy clays for hydrogeological parameter estimation, where a given value of seismic velocity can be interpreted as an indication of high or low permeability, depending on the porosity.

To address these challenges, we choose to concentrate here on Bayesian conditioning. Let us consider an estimate of Z at \mathbf{x} , conditional to a measurement of Z at \mathbf{x}' , $Z(\mathbf{x}')$. The conditional pdf of $Z(\mathbf{x})$ is obtained from Bayes' theorem as follows:

$$f_{Z(\mathbf{x}), Z(\mathbf{x}')} (z|z') = \frac{f_{Z(\mathbf{x}), Z(\mathbf{x}')} (z, z')}{f_{Z(\mathbf{x}')}(z')} \quad (14.25)$$

Thus, $f_{Z(\mathbf{x}), Z(\mathbf{x}')} (z|z') dz$ is the probability to observe $Z(\mathbf{x})$ in the vicinity dz of z , given that $Z(\mathbf{x}') = z'$. When the bivariate distribution of $Z(\mathbf{x})$ and $Z(\mathbf{x}')$ is normal, the conditional pdf is also normal. Furthermore, in that case, the mean of the conditional distribution is equal to the kriging estimate, and the variance is equal to the kriging variance (with one important difference: as now we do have an underlying statistical distribution, the variance can be employed to determine confidence intervals; see Chapter 2 in Rubin, 2003). All this is also true when conditioning on a larger number of measurements. Finally, as Equation 14.25 models the joint distribution of Z at \mathbf{x} and \mathbf{x}' through a bivariate pdf and not through the two-point correlation, such as used by kriging estimators, it opens the door for dealing more effectively with nonlinear correlations (cf. Rubin et al., 1992; Copty et al., 1993; Ezzedine et al., 1999; Hubbard et al., 1997b).

Heretofore, we have assumed that $Z(\mathbf{x}')$ is measured without error, and hence we refer to it as “hard” data. Equation 14.25 provides the means for conditioning on hard data. In applications, however, it is very likely that we would want to condition on a range of values for $Z(\mathbf{x}')$. Such a range of values may represent a distribution of values resulting from measurement errors, or a situation where $Z(\mathbf{x}')$ was not measured, but rather estimated (e.g., by expert opinion, from geophysical data, or by analogy to a set of measurements collected under similar conditions). Several case studies illustrating different methods of extracting hydrogeological property information using geophysical data were given in Section 14.3. We refer to such imprecise information as soft data, and identify the need for conditioning on soft information. Rubin (2003, Chapter 3) proposed to condition on the soft information by modeling it as a probabilistic distribution of \bar{Z} , the estimator of $Z(\mathbf{x}')$, in the form of a pdf $f_{\bar{Z}}(\bar{z})$, defined over a range of values, A . The distribution of Z conditional to \bar{Z} is thus (Chapter 3 in Rubin, 2003):

$$f_{Z(\mathbf{x})|Z(\mathbf{x}')}(z|z' \in A) = \int_A f_{Z(\mathbf{x})|Z(\mathbf{x}')}(z|z') f_{\bar{Z}}(z') dz' \quad (14.26)$$

which implies taking the expected value of the conditional estimate over the range A . In the limiting case of precise measurements at \mathbf{x}' , the soft data pdf $f_{\bar{Z}}(\bar{z})$ is equal to the Dirac delta, $f_{\bar{Z}}(\bar{z}) = \delta(\bar{z} - z')$, and Equation 14.26 simplifies to Equation 14.25. The estimator in Equation 14.26 is exact in the sense that as \mathbf{x} approaches \mathbf{x}' , the conditional pdf approaches the pdf of the soft data, $f_{\bar{Z}}(\bar{z})$. In the limiting case of a Dirac, $f_{\bar{Z}}(\bar{z}) = \delta(\bar{z} - z')$, $Z(\mathbf{x}')$ will approach z' , which is the better-documented case of the exactitude property.

14.4.3.1.1 An Example Based on the Normal Linear Methods.

As was earlier discussed, one of the challenges in using geophysical data for estimating hydrogeological parameters is understanding or developing a meaningful relationship between the two different types of measurements. Relationships between hydrogeological and geophysical parameters in sediments typical of near-surface environments have been a topic of research in recent years (e.g., Marion et al., 1992; Knoll et al., 1995; Mavko et al., 1998; Slater and Lesmes, 2002; Prasad, 2003). Chen et al. (2001) presented a method to formulate a site-specific petrophysical model in a probabilistic fashion using likelihood functions. Their formulation was based on co-located geophysical and hydrogeological data collected at a site, and they used normal linear regression models to relate the disparate parameters. A review of their approach is given here.

Following Chen et al. (2001), let Y denote the SRF of the log conductivity, and let v_g , α , and v_s denote GPR velocity, GPR attenuation, and seismic velocity, respectively. The posterior pdf of $Y(\mathbf{x})$ is obtained using Bayes’ theorem as follows:

$$\underbrace{f'_Y[y(\mathbf{x})|v_g(\mathbf{x}), \alpha(\mathbf{x}), v_s(\mathbf{x})]}_{\text{posterior}} = \underbrace{\text{CL}[y(\mathbf{x})|v_g(\mathbf{x}), \alpha(\mathbf{x}), v_s(\mathbf{x})]}_{\text{Likelihood function}} \underbrace{f_Y[y(\mathbf{x})]}_{\text{prior}} \quad (14.27)$$

Rubin et al. (1992) considered a simplified version of Equation 14.27, whereby the only data available for updating was the seismic velocity. Similar formulations can be written for different types of data combinations.

Note that the likelihood function in Equation 14.27 is a trivariate pdf. It can be expressed as a product of three univariate conditional pdfs:

$$L[y(\mathbf{x})|v_g(\mathbf{x}), \alpha(\mathbf{x}), v_s(\mathbf{x})] = f_{v_g}[v_g(\mathbf{x})|y(\mathbf{x})] f_\alpha[\alpha(\mathbf{x})|y(\mathbf{x}), v_g(\mathbf{x})] f_{v_s}[v_s(\mathbf{x})|y(\mathbf{x}), v_g(\mathbf{x}), \alpha(\mathbf{x})] \quad (14.28)$$

With this representation, the problem of identifying a trivariate pdf is replaced with that of identifying three univariate pdfs.

The normal linear regression model reported in Chen et al. (2001) provides a systematic approach to the inference of the conditional pdfs shown in Equation 14.28. In this approach, each of the univariate pdfs is modeled as a normal, nonstationary pdf. Consider, for example, one of the univariate pdfs forming the likelihood function:

$$f_{v_s}[v_s(\mathbf{x})|y(\mathbf{x}), v_g(\mathbf{x}), \alpha(\mathbf{x})] \quad (14.29)$$

It is assumed to be normal with a nonstationary mean $\mu(\mathbf{x})$, which is further assumed to be a member of the linear function space G , whose basis functions consist of m monomials, g_i , $i = 1, \dots, m$, formed from a combination of powers and products of $y(\mathbf{x})$, $v_g(\mathbf{x})$, and $\alpha(\mathbf{x})$, such as: 1, $y(\mathbf{x})$, $v_g(\mathbf{x})$, $\alpha(\mathbf{x})$, $y^2(\mathbf{x})$, $v_g^2(\mathbf{x})$, $\alpha^2(\mathbf{x})$, $y(\mathbf{x})v_g(\mathbf{x})$, $v_g(\mathbf{x})\alpha(\mathbf{x})$, and so on. It is modeled as follows:

$$\mu(\mathbf{x}) = \sum_{i=1}^m \beta_i g_i(\mathbf{x}) \quad (14.30)$$

The variance of the distribution is taken to be stationary. The final set of functions used to model the mean is determined by a model selection procedure, which will be described below.

The coefficients β_i are estimated by minimizing the residuals sum of squares:

$$\sum_{j=1}^n [v_s(x_j) - \mu(x_j)]^2 \quad (14.31)$$

Let $\underline{\beta} = (\beta_1, \beta_2, \dots, \beta_m)^T$ and $\underline{Z} = [v_s(\mathbf{x}_1), v_s(\mathbf{x}_2), \dots, v_s(\mathbf{x}_n)]^T$, where the exponent T denotes the transpose operator. The estimate $\hat{\underline{\beta}}$ of $\underline{\beta}$ that minimizes Equation 14.31 is given by:

$$\hat{\underline{\beta}} = (\underline{\underline{D}}^T \underline{\underline{D}})^{-1} \underline{\underline{D}}^T \underline{Z} \quad (14.32)$$

where $\underline{\underline{D}}$ is a design matrix, given by:

$$\begin{pmatrix} g_1(\mathbf{x}_1), g_2(\mathbf{x}_1), g_3(\mathbf{x}_1), \dots, g_m(\mathbf{x}_1) \\ g_1(\mathbf{x}_2), g_2(\mathbf{x}_2), g_3(\mathbf{x}_2), \dots, g_m(\mathbf{x}_2) \\ \vdots \\ g_1(\mathbf{x}_n), g_2(\mathbf{x}_n), g_3(\mathbf{x}_n), \dots, g_m(\mathbf{x}_n) \end{pmatrix} \quad (14.33)$$

Once $\hat{\underline{\beta}}$ is obtained, the mean and variance of the pdf $f_{v_s}[v_s(\mathbf{x})|y(\mathbf{x}), v_g(\mathbf{x}), \alpha(\mathbf{x})]$ are defined by:

$$\hat{\mu}(\mathbf{x}) = \sum_{i=1}^m \hat{\beta}_i g_i(\mathbf{x}) \quad (14.34)$$

and

$$\hat{\sigma}^2 = \frac{1}{n-m} \sum_{j=1}^n [\nu_s(\mathbf{x}_j) - \mu(\mathbf{x}_j)]^2 \quad (14.35)$$

where n is the number of measurements and m the number of basis functions g_i .

Selecting and eliminating basis functions is the key to the normal linear regression model. Chen et al. (2001) includes all possible distinct monomials of powers up to 4 in the initial set of basis functions. By deleting some of the initial basis functions based on t -test and p -value selection criteria (Stone, 1995) and in an iterative manner, Chen et al. (2001) obtained the final set of basis functions.

In principle, an empirical approach such as we present here is at a disadvantage compared to methods that use physically based forward models to relate between the various attributes. However, in many cases, accurate and reliable mathematical models are difficult to construct, or are subject to uncertain boundary conditions.

An application of the Bayesian approach is discussed in great detail in Hubbard et al. (2001), and is briefly summarized here. Prior estimates of hydraulic conductivity were obtained from kriging flowmeter measurements available from wellbores at a U.S. Department of Energy Bacterial Transport Site in Virginia. Updating of the prior estimates was performed using 100 MHz radar velocity data available from crosshole radar tomograms collected between the wellbores as well as a likelihood relationship developed using co-located measurements near the wellbore. Figure 14.18a (modified from Hubbard et al., 2001) illustrates the estimates of the mean hydraulic conductivity posterior pdfs along a vertical cross section at the study site, obtained using radar tomographic data and wellbore flowmeter data within a Bayesian approach. As the sediments within the saturated section at this site are fairly homogeneous (consisting predominantly of sands with different levels of sorting), the range of hydraulic conductivity is only about an order of magnitude.

Even under the relatively homogeneous conditions and with extensive wellbore flowmeter control available at this site, it was difficult to capture the variability of hydraulic conductivity using wellbore data alone. Figure 14.18b illustrates the mean values of hydraulic conductivity at each point in space obtained through kriging measurements obtained from the five wellbores whose locations are shown on the figure, or the prior pdfs at each point in space. Comparison of Figure 14.18a and 14.18b highlights how incorporation of geophysical data can lead to dramatically different estimates of hydraulic conductivity than if only wellbore data were used in the estimation procedure, especially at distances of a few meters from direct wellbore control (such as the area between wellbores T2 and M3). Recent studies have attested to the value of the geophysically obtained posterior estimates for understanding subsurface bacterial transport (Johnson et al., 2001; Mailloux et al., 2003) and for improving the prediction of chemical transport (Scheibe and Chien, 2003) at this site.

14.4.4 Model Selection

In our earlier discussion, reference was made to models (e.g., F in Equation 14.17), and we expand upon our model discussion in this section. Model development is a subjective process (see Chapter 1 in Rubin, 2003), often subject to large errors. Such errors, caused by data scarcity and errors in judgment, are unavoidable, but they should be minimized. Model selection is an attempt to address model error in a systematic manner. It requires formulation and comparison of alternative models, and elimination of all but the model that maximizes some optimality criteria. The idea (and hope) is that performing model selection will minimize modeling error. The most commonly used model selection methods include hypothesis testing (Mood et al., 1963; Luis and McLaughlin, 1992), cross-validation (Deutsch and Journel, 1998; Christensen and Cooley, 1999), Akaike's Information Criterion (AIC, see Akaike, 1974; Carrera and Neuman, 1986b), and Bayesian Information Criterion (BIC).

Cross-validation is the most intuitive approach, and it consists of dividing the data into two non-overlapping subsets. The first subset, called the training subset, is used for model development and

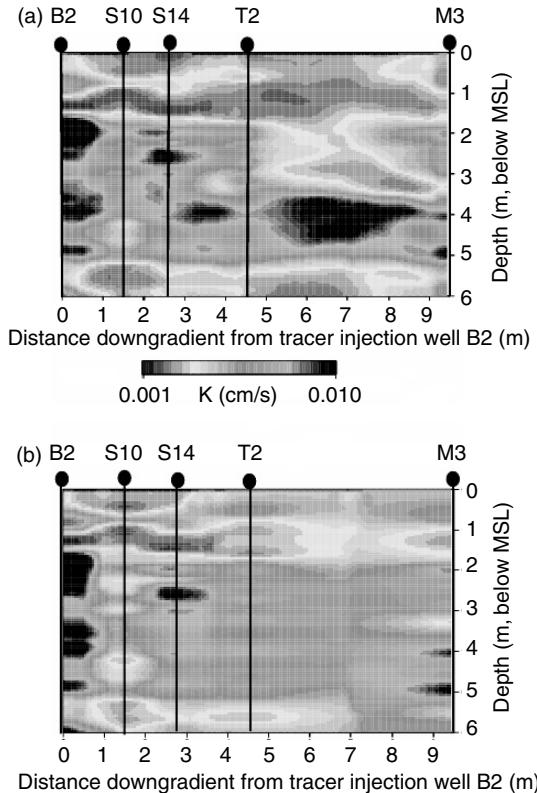


FIGURE 14.18 (a) Contoured mean values of posterior pdfs obtained using radar tomographic data conditioned to flowmeter data within a Bayesian estimation procedure; (b) Contoured mean values of prior pdfs obtained from kriging dense wellbore flowmeter measurements (modified from Hubbard et al., 2001, *Water Resour. Res.*, 37(10), 243–2456). Comparison of these figures shows that inclusion of the geophysical data significantly changes the estimates of hydraulic conductivity, especially at distances over ~ 1 m from direct wellbore control.

calibration, and a second, complementary, testing subset is used to evaluate the predictive accuracy of the calibrated model. This is done by comparing the model-predicted values at the locations where the testing data are, with the actual, measured values. Metrics are then developed based on model predictions and actual measurements, and are used to compare between models.

To demonstrate this idea, let us consider $z(\mathbf{x}_i)$, the measurement of the SRF Z at \mathbf{x}_i . A geostatistical model, identified using the training subset, is employed to obtain a kriging estimate at \mathbf{x}_i , z_i^* , and a kriging variance, σ_i^2 (see Chapter 3 of this volume). This is done repeatedly for all \mathbf{x}_i , $i = 1, \dots, N$, where N denotes the number of data in the testing subset. At the end of this exercise, the following series of analysis can be performed:

- A correlation analysis between $z(\mathbf{x}_i)$ and z_i^* , $i = 1, \dots, N$. This analysis can identify a systematic bias in the model. The absence of correlation is a positive indication.
- A correlation analysis between σ_i and $|z_i^* - z(\mathbf{x}_i)|$ for $i = 1, \dots, N$. This analysis aims at evaluating the model's ability to predict correctly the estimation error. A high degree of correlation is a positive indication.
- A correlation analysis between σ_i and $z(\mathbf{x}_i)$ for $i = 1, \dots, N$. This analysis aims to reveal heteroskedasticity. Any significant correlation here is an indication of a systematic bias (i.e., an unmodeled deterministic model component).
- Correlation analysis between the residuals $|z_i^* - z(\mathbf{x}_i)|$ and $|z_j^* - z(\mathbf{x}_j)|$ for $i, j = 1, \dots, N$. In an ideal situation, there will not be any correlation.

Additional analyses are suggested in Kitanidis (1997). These analyses provide metrics that can be used to compare between models. None of them has the power to indicate optimality, but they can definitely indicate poor performance, and upon comparison between models, they can reveal which performs better.

Can we use the ML criterion to compare between models? ML leads to optimal parameters, in a maximum likelihood sense, for a given model structure without regard to how well the model represents the real system (Carrera and Neuman, 1986a, 1986b, 1986c). The estimated parameters may be meaningless if the model fails to capture the important characteristics of the aquifer. To compensate for poor model performance, the modeler may be tempted to increase the number of model parameters (e.g., number of gridblocks used in the numerical model of the aquifer), and hence the model complexity. Obviously, a larger number of parameters will produce a better fit, but not necessarily a better model, as an increase in the number of parameters leads to larger parameter variance and to instability (in the sense that the slightest changes in data will lead to significant changes in parameter values. Hence, the ML criterion cannot be used for model selection, as it accounts for goodness of fit but not for model simplicity.

The common model selection methods attempt to balance between simplicity, usually associated with the number of adjustable parameters, and goodness of fit (Forster, 2000). The main difference between these methods is in the relative weights they give to simplicity versus goodness of fit. The AIC intends to maximize entropy and simplicity, and it selects the model that has the smallest AIC:

$$AIC(\hat{\theta}) = -2 \ln L(\hat{\theta}) + 2p \quad (14.36)$$

where $\hat{\theta}$ is the maximum likelihood estimate of the parameter vector, $L(\hat{\theta})$ is the corresponding maximum value of the likelihood criterion, and p is the number of parameters in $\hat{\theta}$. The first term on the right-hand side of Equation 14.36 is the negative logarithm of the maximum value for the likelihood criterion, while the second term penalizes complexity, expressed as the number of parameters p .

The BIC method selects the a posteriori most probable model among the proposed alternatives. That model is the one that minimizes the BIC (Forster, 2000; Gaganis and Smith, 2001):

$$BIC(\hat{\theta}) = -2 \ln L(\hat{\theta}) + p \ln(n) \quad (14.37)$$

where n is the number of measurements. Compared to AIC, BIC gives a greater weight, by a factor of $\ln(n)/2$, to simplicity. Despite the difference, both AIC and BIC support the principle of parsimony in that, all other things being equal, the model with smallest p is chosen. The relative performance of BIC vs. AIC depends on the nature of the problem, the size of the parameter vector, and the amount of available data on the dependent variables (Ye et al., 2004; Neuman, 2003).

14.4.5 Summary of Stochastic Techniques for Hydrogeophysical Investigations

The stochastic paradigm presented here offers a rational framework for analyzing spatial variability and the consequences of data scarcity, and hence for subsurface characterization. In this context, Bayes' theorem offers a flexible platform, allowing for alternative formulations, and for incorporation of complementary data of various types. There are several elements of Bayes' theorem that are needed for a successful application (which we reviewed here), including a systematic approach for modeling prior information and the likelihood function, and a rigorous model selection. For a systematic modeling of prior information, we presented the Minimum Relative Entropy (MRE) concept. The idea here is to select for the prior pdf the one with the largest entropy among all pdfs (relative to a benchmark prior) that satisfy a given set of data constraints, and in doing that, to minimize the infusion of information that is not data-based and to eliminate potential bias. This is a critical element in integration between hydrological and geophysical data. MRE is a relatively new concept in hydrogeological applications. For model selection, we reviewed a few classical, well-documented methods such as the Akaike's Information Criterion and the Bayesian Information Criterion.

14.5 Summary

In this chapter, we have provided a review of common geophysical methods and hydrogeophysical concepts, and have reviewed advanced stochastic approaches for integrating hydrogeological and geophysical data. By combining measurements made from hydrologic instruments with geophysical methods, accurate characterization and monitoring of subsurface properties can potentially be achieved with high temporal and spatial resolution over a range of spatial scales. The existence of such potential is intuitively appealing: geophysics offers a view of the subsurface from a perspective other than hydrogeology, and thus, when reconciled with the hydrogeological perspective, can help obtain a more coherent image.

What are the keys for unleashing this potential? First and foremost, in our opinion, is the availability of a systematic approach for estimation and data integration; the development of true joint hydrogeological-geophysical inversion approaches, which are distinguished from the sequential approaches used now, whereby geophysical surveys are inverted, and the outcome is then viewed as a starting point for hydrogeological inversion. True joint inversion approaches that are now in development (cf. Kowalsky et al., 2005; Linde et al., 2005) introduce geological constraints on the geophysical inversion, which will lead to better accuracy. Additional discussion is provided in Hubbard and Rubin (2002). Second is a substantial improvement in our understanding of the petrophysical models that relate between geophysical and hydrogeological attributes. An improved understanding will allow better usage of the geophysical data and will permit us to benefit from various geophysical surveys, exploring each for its proven strength. Finally, there is a need to gain more hydrogeophysical experience through practice and collaboration, and to extend the quantitative approaches that have been successfully applied at the local scale to more regional investigations of hydrological watersheds. Given that there are many nuances associated with the quantitative use of geophysical data for subsurface investigations, and that improper use of geophysical methods may reduce the overall credibility of such approaches (as discussed by Greenhouse, 1991), the use of geophysical data for quantitative parameter estimation must be approached with caution. As suggested by the National Research Council (2000), a dedicated and long-term effort is needed to investigate and document the utility of noninvasive characterization methods for a wide variety of applications, with an emphasis on research performed by multidisciplinary teams, increased accurate automation of data acquisition, data processing, and decision-making for real-time field interpretation and improved methods for transferring advanced tools and approaches to field practitioners.

Increasing population and pollution, decreasing natural resources, and changing climatic conditions all contribute to an urgency associated with improving our understanding of the shallow subsurface. Although investigators are just starting to explore the benefits of fusing geophysical and hydrogeological datasets for an improved understanding, there is already a substantial body of evidence that illustrates the potential of such endeavors.

Glossary

Conditional Monte Carlo Simulation Modeling uncertainty through generation of alternative images of the flow domain using the space random function model and, in addition, requiring that the generated images fit field data at specified locations.

Crosshole Geophysical Methods Acquisition geometry used by seismic, radar, electrical resistivity, and electromagnetic induction methods whereby energy from sources located in one borehole are recorded by sensors in another.

Dielectric Constant A measure of the separation (polarization) of opposite electrical charges within a material that has been subjected to an external electrical field.

Dynamic Process Monitoring Collection of data at the same location as a function of time. Processing of these data as “time difference” cross sections can reveal information about processes such as water infiltration, stream flooding, or contaminant plume migration.

Effective Conductivity The coefficient multiplying the expected value of the head gradient to yield the expected value of the flux. Must not depend on the parameters defining the flow problem or

the boundary conditions. In a stationary domain, the expected value of the flux is also equal to its space average.

Electrical Profiling An electrical survey made at several surface locations using a constant electrode separation distance. Electrical profiles, also called constant spread profiles or resistivity profiles, provide information about lateral changes in apparent resistivity.

Electrical Resistivity A measure of the ability of electrical current to flow through materials, measured in Ohm-m. Electrical resistivity is an intrinsic property of a material and is the inverse of electrical conductivity.

Electrical Sounding An electrical survey made at a single surface location by moving the electrodes progressively farther apart. Electrical soundings, also called expanding spread profiles, provide information about apparent electrical resistivity as a function of depth.

Monte Carlo Simulation Modeling uncertainty through generation of alternative images of the flow domain using the space random function model.

Seismic Velocity The propagation velocity of a seismic disturbance through a material. As seismic velocity is a function of the elastic parameters of a material, seismic velocity has been used to estimate lithology and porosity.

Space Random Function A probabilistic model for describing the variations of attributes such as hydraulic conductivity in heterogeneous media.

Spatial Correlation A statistical measure of similarity between attributes in a heterogeneous domain, usually as a function of separation distance between two locations. Spatial correlation is modeled through the spatial covariance or the semi-variogram.

Stationarity Invariance of the space random function model to translation in space.

Statistical Anisotropy Implies that the spatial correlations are direction dependent.

Statistical Isotropy Invariance of the spatial correlation to rotation of the separation distance.

Upscaling Representing the heterogeneous domain for modeling purposes through elements larger than the scales of heterogeneity but not large enough to warrant the use of effective properties. The effects of the sub-element heterogeneity are represented using coefficients which depend on the element size.

Acknowledgments

Funding for this study was provided by the Environmental Management Science Program, Office of Biological and Environmental Research, U.S. Department of Energy (DOE) Grant DE-FG07-96ER14726 to Susan Hubbard, and by NSF-EAR-0439649 to Yoram Rubin.

References

- Akaike, H., 1974, A new look at the statistical model identification, *IEEE Trans. Autom. Control*, AC-19, 716–723.
- Alumbaugh, D., P. Change, L. Paprocki, J. Brainard, R. J. Glass, and C. A. Rautman, 2002, Estimating moisture contents in the vadose zone using cross-borehole ground penetrating radar: a study of accuracy and repeatability, *Water Resour. Res.*, 38, doi: 2001WRR001028.
- Annan, 2005, Shallow seismic methods, Chapter 8 in *Hydrogeophysics*, Eds. Y. Rubin and S. Hubbard, Springer, Netherlands, pp. 185–214.
- Annan, A. P. and Cosway, S. 1992. Ground penetrating radar survey design. *Proc. of the Symposium on the Applications of Geophysics to Engineering and Environmental Problems (SEGEEP)*. Chicago.
- Archie, G. E. 1942. The electric resistivity log as an aid in determining some reservoir characteristics. *Trans. AIME*, 146, 54–62.
- Best, M. E. 1992. *Geological Association of Canada Short Course Notes Volume 10*. Wolfville, Nova Scotia, May 28–29.

- Binley, A., P. Winship, L. J. West, M. Pokar, and R. Middleton, 2002, Seasonal variation of moisture content in unsaturated sandstone inferred from borehole radar and resistivity profiles, *J. Hydrology*, 267(3–4), 160–172.
- Binley, A. and A. Kemna, 2005, DC resistivity and induced polarization methods, Chapter 5 in *Hydrogeophysics*, Eds. Y. Rubin and S. Hubbard, Springer, Netherlands, pp. 129–156.
- Birkelo, Steeples, Miller, and Sophocleous, 1987, Seismic reflection study of a shallow aquifer during a pumping test, *Ground Water*, 25, 703.
- Brewster, M. L. and Annan, A. P. 1994. Ground-penetrating monitoring of a controlled DNAPL release: 200 MHZ radar, *Geophysics*, 59(8), 1211–1221.
- Burger, H. R. 1992. *Exploration Geophysics of the Shallow Subsurface*. Prentice Hall, Englewood Cliffs, NJ.
- Buselli, G., Barber, C., Davis, G. B., and Salama, R. B. 1990. Detection of groundwater contamination near waste disposal sites with transient electromagnetic and electrical methods, in *Geotechnical and Environmental Geophysics Vol. 2: Environmental and Groundwater, S.E.G. Investigations in Geophysics* 5. Ed. Stanley Ward. Society of Exploration Geophysicists. pp. 27–39.
- Butler, D. K. 1991. Tutorial — engineering and environmental applications of microgravity. *Proc. of the Symposium on the Application of Geophysics to Engineering and Environmental Problems*. Knoxville, TN, pp. 139–177.
- Carrera, J. and S. P. Neuman, 1986a, Estimation of aquifer parameters under transient and steady state conditions, 1. Maximum likelihood method incorporating prior information, *Water Resour. Res.*, 22(2), 199–210.
- Carrera, J. and S.P. Neuman, 1986b, Estimation of aquifer parameters under transient and steady state conditions, 2. Uniqueness, stability and solution algorithms, *Water Resour. Res.*, 22(2), 211–227.
- Carrera, J. and S.P. Neuman, 1986c, Estimation of aquifer parameters under transient and steady state conditions, 3. Application to synthetic and field data, *Water Resour. Res.*, 22(2), 228–242.
- Chen, J., S. Hubbard, and Y. Rubin, 2001, Estimating the hydraulic conductivity at the South Oyster site based on the normal linear regression model, *Water Resour. Res.*, 37(6), 1603–1613.
- Chen, J., and Y. Rubin, 2003, An effective Bayesian model for lithofacies estimation using geophysical data, *Water Resour. Res.*, 39(5), 4–1 to 4–11.
- Chen, J., S. Hubbard, Y. Rubin, C. Murray, E. Roden, and E. Majer, 2005, Geochemical characterization using geophysical data: a case study at the South Oyster bacterial transport site in Virginia, *Water Resour. Res.*, 40, W12412.
- Christensen, S. and R.L. Cooley, 1999, Evaluation of prediction intervals for expressing uncertainties in groundwater flow model predictions, *Water Resour. Res.*, 35(9), 2627–2639.
- Cpty, N., Y. Rubin, and G. Mavko, 1993, Geophysical–hydrological identification of field permeabilities through Bayesian updating, *Water Resour. Res.*, 29(8), 2813–2876.
- Cpty, N. and Y. Rubin, 1995, A stochastic approach to the characterization of lithofacies from surface seismic and well data, *Water Resour. Res.*, 31(7), 1673–1686.
- Daniels, J.J., B. Allred, A. Binley, D. LaBrecque, and D. Alumbaugh, 2005, Hydrogeophysical case studies in the vadose zone, Chapter 14 in *Hydrogeophysics*, Eds. Y. Rubin and S. Hubbard, Springer, Netherlands.
- Davis, J. and A. Annan, 1989, Ground-penetrating radar for high-resolution mapping of soil and rock stratigraphy. *Geophys. Prospect.* 37, 531–551.
- Day-Lewis, F., P.A. Hsieh, and S.M. Gorelick, 2000, Identifying fracture-zone geometry using simulated annealing and hydraulic connection data, *Water Resour. Res.*, 36(7), 1707–1721.
- Degan, G., 1989, *Flow and Transport in Porous Formations*, Springer-Verlag, New York.
- Deutsch, C. and A.G. Journel, 1998, *GSLIB: Geostatistical Software Library and Users Guide*, Oxford University Press, New York.
- Doll, W.E., R.D. Miller and J. Xia, 1994, Noninvasive shallow source comparison for hazardous waste site investigations, *SEG Technical Program Expanded Abstracts*, pp. 591–594.
- Du, S. and P. Rummel, 1994, Reconnaissance studies of moisture in the subsurface with GPR, paper presented at the 5th International Conference on GPR, University of Waterloo, Kitchener, Ontario, Canada.

- Everett, M.E. and M.A. Meju, 2005, Near-surface controlled source electromagnetic induction: background and recent advances, Chapter 6 in *Hydrogeophysics*, Eds. Y. Rubin and S. Hubbard, Springer, Netherlands.
- Ezzedine, S., Y. Rubin, and J. Chen, 1999, Bayesian method for hydrogeological site characterization using borehole and geophysical survey data: theory and application to the Lawrence Livermore National Laboratory Superfund site, *Water Resour. Res.*, 35(9), 2671–2683.
- Ferre, T.A., A. Binley, J. Geller, E. Hill, and T. Illangasarkare, 2005, Hydrogeophysical methods at the laboratory scale, Chapter 15 in *Hydrogeophysics*, Eds. Y. Rubin and S. Hubbard, Springer, Netherlands.
- Fisher, E., G. McMechan, and A. P. Annan, 1992, Acquisition and processing of wide-aperture ground-penetrating radar data, *Geophysics*, 57(3), 495–504.
- Forster, M.R., 2000, Key concepts in model selection: performance and generalizability, *J. Math. Psychol.*, 44, 205–231.
- Frischknecht, F. C., V. F. Labson, B. R. Speis, and W. L. Anderson, 1991, Profiling methods using small sources, in *Electromagnetic Methods in Applied Geophysics*. Vol. 2, S.E.G. *Investigations in Geophysics* 3. Ed. M. Nabighian, pp. 105–270.
- Gaganis, P. and L. Smith, 2001, A Bayesian approach to the quantification of the effect of model error on the predictions of groundwater models, *Water Resour. Res.*, 37(9), 2309–2322.
- Gelhar, L., 1993, *Stochastic Subsurface Hydrology*, Prentice Hall, Englewood Cliffs, NJ, 290p.
- Geller, J. T. and L. R. Myer, 1995, Ultrasonic imaging of organic liquid contaminants in unconsolidated porous media. *J. Contam. Hydrology*, 19, 85–104.
- Goldstein, N. E., 1994, Expedited site characterization geophysics: geophysical methods and tools for site characterization. *Lawrence Berkeley Laboratory LBL-35384*.
- Greaves, R. J., D. P. Lesmes, J. M. Le, and M. N. Toksoz, 1996, Velocity variations and water content estimated from multi-offset, ground-penetrating radar, *Geophysics*, 61(3), 683–695.
- Greenhouse, J. P. and R. D. Harris, 1983, Migration of contaminants in groundwater at a landfill: a case study. *J. Hydrology*, 63, 177–197.
- Greenhouse, J., M. Brewster, G. Schneider, J. D. Redman, P. Annan, G. Olhoeft, J. Lucius, K. Sander, and A. Mazzella, 1993, Geophysics and solvents: the Borden experiment. *The Leading Edge*, 12(4), 261–267.
- Grote, K., S. Hubbard, and Y. Rubin, 2003, Field-scale estimation of volumetric water content using ground-penetrating radar ground wave techniques. *Water Resour. Res.*, 39(11), doi: 1029/2003WRR002045.
- Grote, K., S. Hubbard, J. Harvey, and Y. Rubin, 2005, Evaluation of Infiltration in Layered Pavements using surface GPR reflection techniques. *J. Appl. Geophys.*, 57(2), 129–153.
- Heigold, P. C., R. H. Gildenson, K. Cartwright, and P. C. Reed, 1979, Aquifer transmissivity from surficial electrical methods. *Ground Water*, 17, 338–345.
- Hinze, W. J., 1990, The role of gravity and magnetic methods in engineering and environmental studies, in *Geotechnical and Environmental Geophysics*, Vol. 1: *Review and Tutorial*, SEG *Investigations in Geophysics* No. 5. Ed. Stanley Ward, pp. 75–126.
- Hoeksema, R. J. and P. K. Kitanidis, 1984, An application of the geostatistical approach to the inverse problem in two-dimensional groundwater modeling. *Water Resour. Res.*, 20(7), 1003–1020.
- Hoeksema, R. J. and P. K. Kitanidis, 1985, Analysis of the spatial structure of properties of selected aquifers, *Water Resour. Res.*, 21(4), 563–572.
- Hoekstra, P. and M. Blohm, 1990, Case histories of time-domain electromagnetic soundings in environmental geophysics, in *Geotechnical and Environmental Geophysics*, Vol. 2: *Environmental and Groundwater*, S.E.G. *Investigations in Geophysics* 5. Ed. Stanley Ward, pp. 1–17.
- Hubbard, S. S., J. E. Peterson, Jr., E. Major, P. Zawislanski, K. Williams, J. Roberts, and F. Wobber, 1997a, Estimation of permeable pathways and water content using tomographic radar data. *The Leading Edge*, 16(11), 1623–1628.

- Hubbard, S., Y. Rubin, and E. Majer, 1997b, Ground penetrating radar-assisted saturation and permeability estimation in bimodal systems, *Water Resour. Res.*, 33(5), 971–990.
- Hubbard, S., Y. Rubin, and E. Majer, 1999, Spatial correlation structure estimation using geophysical and hydrological data, *Water Resour. Res.*, 35(6), 1709–1725.
- Hubbard, S. and Y. Rubin, 2000, Integrated hydrogeological–geophysical site characterization techniques, *J. Contam. Hydrology*, 45, 3–34.
- Hubbard, S., J. Chen, J. Peterson, E.L. Majer, K.H. Williams, D.J. Swift, B. Mailloux, and Y. Rubin, 2001, Hydrogeological characterization of the South Oyster Bacterial Transport Site using geophysical data, *Water Resour. Res.*, 37(10), 2431–2456.
- Hubbard, S. and Y. Rubin, 2002, Hydrogeophysics: state-of-the-discipline, *EOS*, 83(51), 602–606.
- Hubbard, S. and K. Williams, 2004, Geophysical monitoring of gas production during biostimulation, WM04 Conference Proceedings, Tuscon Arizona, Feb. 10, 2004.
- Hubbard, S., I. Lunt, K. Grote and Y. Rubin, 2005, Vineyard soil water content: mapping small scale variability using ground penetrating radar, *Gesocience Canada Geology and Wine series*, Eds. R. W. Macqueen and L. D. Meinert, in press.
- Hubbard, S. and Y. Rubin, 2005, Introduction to hydrogeophysics, Chapter 1 in *Hydrogeophysics*, Eds. Y. Rubin and S. Hubbard, Springer, Netherlands, pp. 3–22.
- Huisman, J.A., J.J.J. C. Snepvangers, W. Bouting, and G.B.M. Heuvelink, 2002, Mapping spatial variation in surface soil water content: comparison of ground-penetrating radar and time domain reflectrometry, *J. Hydrology*, 269, 194–207.
- Huisman, J. A., S. S. Hubbard, J. D. Redman, and A. P. Annan, 2003, Measuring soil water content with ground penetrating radar: a review, *Vadose Zone J.*, 4(2), 476–491.
- Hunter, J. A., S. E. Pullan, R. A. Burns, R. M. Gagne and R. L. Good, 1984, Shallow seismic reflection mapping of the overburden-bedrock interface with the engineering seismograph — some simple techniques, *Geophysics*, 49, 1381–1385.
- Huntley, D., 1986, Relations between electrical resistivity in granular aquifers. *Ground Water*, 24(4), 466–474.
- Hyndman, D. W., J. M. Harris, and S. M. Gorelick, 1994, Coupled seismic and tracer test inversion for aquifer property characterization, *Water Resour. Res.*, 30(7), 1965–1977.
- Hyndman, D. W. and S. M. Gorelick, 1996, Estimating lithologic and transport properties in three dimensions using seismic and tracer data: the Kesterson aquifer, *Water Resour. Res.*, 32(9), 2659–2670.
- Hyndman, D. W. and J. M. Harris, 1996, Traveltime inversion for the geometry of aquifer lithologies, *Geophysics*, 61(6).
- Hyndman, D. and J. Tronicke, 2005, Hydrogeophysical studies at the local scale: the saturated zone, Chapter 13 in *Hydrogeophysics*, Eds. Y. Rubin and S. Hubbard, Springer, Netherlands, pp. 391–412.
- Johnson, W. P., P. Zhang, M. E. Fuller, T. D. Scheibe, B. J. Mailloux, T. C. Onstott, M. F. DeFlaun, S. S. Hubbard, J. Radtke, W. P. Kovacik, and W. Holbin, 2001, Ferrographic tracking of bacterial transport in the field at the narrow channel focus area, Oyster, VA., *Environ. Sci. Technol.*, 35, 182–191.
- Kachanoski, R. G., E. G. Gregorich, and I. J. Van Wesenbeeck, 1988, Estimating spatial variations of soil water content using noncontacting electromagnetic inductive methods, *Can. J. Soil Sci.*, 68, 715–722.
- Kelly, W. E., 1977, Geoelectric sounding for estimating aquifer hydraulic conductivity, *Ground Water*, 15(6), 420–425.
- Keys, S. W., 1989, *Borehole Geophysics Applied to Ground-Water Investigations*. National Water Well Association, Dublin, OH.
- Kick, J. F., 1985, Depth to bedrock using gravimetry. *Geophysics: The Leading Edge of Exploration*, Soc. Exploration Geophysicists, 4(4), 38–42.
- Kitanidis, P. K. and E. G. Vomvoris, 1983, A geostatistical approach to the inverse problem in groundwater modeling, *Water Resour. Res.*, 19, 677–690.
- Kitanidis, P. K., 1997, *Introduction to Geostatistics*, Cambridge University Press.
- Klimentos, T., 1991, The effects of porosity-permeability-clay content on the velocity of compressional waves, *Geophysics*, 56, 1930–1939.

- Knight, R., P. Tercier, and H. Jol, 1997, The role of ground penetrating radar and geostatistics in reservoir description, *Leading Edge*, 16(11), 1576–1583.
- Knoll, M., R. Knight, and E. Brown, 1995, Can accurate estimates of permeability be obtained from measurements of dielectric properties? SAGEEP Annual Meeting *Extended Abstracts, Environ. Eng. Geophys. Soc.*, Englewood, CO.
- Kobr, M., S. Mares, and F. Paillet, 2005, Geophysical well logging: borehole geophysics for hydrogeological studies: principles and applications, Chapter 10 in *Hydrogeophysics*, Eds. Y. Rubin and S. Hubbard, Springer, Netherlands, pp. 291–332.
- Kosinski, W. K. and W. E. Kelly, 1981, Geoelectric soundings for predicting aquifer properties, *Ground Water*, 19, 163–171.
- Lankston, R. W., 1990, High-resolution refraction seismic data acquisition and interpretation, in *Geotechnical and Environmental Geophysics Vol. 1: Environmental and Groundwater*, S.E.G. *Investigations in Geophysics* 5. Ed. Stanley Ward, pp. 45–73.
- Lesmes, D. and S. Friedman, 2005, Relationships between the electrical and hydrogeological properties of rocks and soils, Chapter 4 in *Hydrogeophysics*, Eds. Y. Rubin and S. Hubbard, Springer, Netherlands, pp. 87–128.
- Linde, H., S. Finsterle, and S. Hubbard, 2006, Inversion of hydrological tracer test data using tomographic constraints, 42, WO4410, doi: 10.1029/2004WR003806.
- Luis, S. J. and D. McLaughlin, 1992, A stochastic approach to model validation *Adv. Water Resour.*, 15, 15–32.
- Lunt, I., S. Hubbard, and Y. Rubin, 2005, Soil moisture content estimation using ground-penetrating radar reflection data, *J. Hydrology*, doi:10.1016/j.jhydrol.2004.10.014, 307/1–4, 254–269.
- Mac Lean, H. D., 1992, *Commercially Available Broadband Electromagnetic Systems for U.S. Department of Energy Waste-Site Characterization*. Chem-Nuclear Geotech, Inc., U.S. DOE, Grand Junction Projects Office, Grand Junction, CO, GJPO-GP-2.
- Mailloux, B. J., M. E. Fuller, T. C. Onstott, J. Hall, H. Dong, M. F. DeFlaun, S. H. Streger, R. K. Rothmel, M. Green, D. J. P. Swift, and J. Radke, 2003, The role of physical, chemical and microbial heterogeneity on the field-scale transport and attachment of bacteria, *Water Resources Research*, 39(6), 2–1 to 2–17.
- Marion, D., A. Nur, H. Yin, and D. Han, 1992, Compressional velocity and porosity in sand-clay mixtures, *Geophysics*, 57, 554–563.
- Mavko, G., T. Mukerji, and J. Dvorkin, 1998, *The Rock Physics Handbook: Tools for Seismic Analysis in Porous Media*, Cambridge University Press, New York.
- Maxwell, R., W. Kastenberg, and Y. Rubin, 1999, Hydrogeological site characterization and its implication on human risk assessment, *Water Resour. Res.*, 35(9), 2841–2855.
- Mazac, O., W. E. Kelly, and I. Landa, 1985, A hydrogeophysical model for relations between electrical and hydraulic properties of aquifers, *J. Hydrology*, 79, 1–19.
- McLaughlin, D. and L. R. Townley, 1996, A reassessment of the groundwater inverse problem, *Water Resour. Res.*, 32(5), 1131–1161.
- McNeill, J. D. 1990. Use of electromagnetic methods for groundwater studies, in *Geotechnical and Environmental Geophysics Vol. 1: Review and Tutorial*, SEG *Investigations in Geophysics* No. 5. Ed. Stanley Ward, pp. 191–218.
- Miller, R.D., D.W. Steeples, and M. Brannan, 1989, Mapping a bedrock surface under dry alluvium with shallow seismic reflections, *Geophysics*, 54, 1528–1534.
- Mood, A.M., F.A. Graybill, and D.C. Boes, 1963, *Introduction to the Theory of Statistics*, 3rd Edition, McGraw-Hill, New York.
- Paine, J.G. and B.R. Minty, 2005, Airborne hydrogeophysics, Chapter 11 in *Hydrogeophysics*, Eds. Y. Rubin and S. Hubbard, Springer, Netherlands, pp. 333–360.
- Peterson, J. E., B. N. Paulsson, and T. V. McEvilly, 1985, Applications of algebraic reconstruction techniques to crosshole seismic data, *Geophysics*, 50, 1566–1580.
- Peterson, J.E., Jr., 2001, Pre-inversion corrections and analysis of radar tomographic data, *J. Env. Eng. Geophys.*, 6, 1–17.

- Pfeifer, M. C. and H. T. Anderson, 1995. CD-resistivity array to monitor fluid flow at the INEL-infiltration test. Proceedings of the Symposium on the Application of Geophysics to Engineering and Environmental Problems. Orlando, FL, April 23–26, pp. 709–718.
- Poeter, EP, S.A. McKenna, and W.L. Wingle, 1997, Improving groundwater project analysis with geophysical data, *The Leading Edge*, 16(11), 1675–1681.
- Prasad, M., 2003, Velocity-permeability relations within hydraulic units, *Geophysics*, 68(1), 108–117.
- Press, F., 1966, Seismic velocities, in *Handbook of Physical Constants*, rev. ed., Clark, S. P., Jr., Ed. Geologic Society of America Memoir, vol. 97, pp. 97–173.
- Pride, S., 2005, Relationships between seismic and hydrological properties, Chapter 9 in *Hydrogeophysics*, Eds. Y. Rubin and S. Hubbard, Springer, Netherlands, pp. 253–290.
- Pullan, S. E. and J. A. Hunter, 1990. Delineation of buried bedrock valleys using the optimum offset shallow seismic reflection technique, in *Geotechnical and Environmental Geophysics: Vol. 3, Geotechnical, S.E.G. Investigations in Geophysics 5*. Ed. Stanley Ward, pp. 75–87.
- Purvance, D. T. and R. Andricevic, 2000, Geoelectric characterization of the hydraulic conductivity field and its spatial structure at variable scales, *Water Resour. Res.*, 36(10), 2915–1924.
- Ramirez, A., W. Daily, D. LaBrecque, E. Owen, and D. Chesnut, 1993, Monitoring an underground steam injection process using electrical resistance tomography. *Water Resour. Res.*, 29(1), 73–87.
- Robinson, E. S. and C. Coruh, 1988. *Basic Exploration Geophysics*, John Wiley & Sons, New York.
- Rubin, Y. and G. Dagan, 1987a, Stochastic identification of transmissivity and effective recharge in steady groundwater flow, 1. Theory, *Water Resour. Res.*, 23(7), 1175–1192.
- Rubin, Y. and G. Dagan, 1987b, Stochastic identification of transmissivity and effective recharge in steady groundwater flow, 2. Case study, *Water Resour. Res.*, 23(7), 1193–1200.
- Rubin, Y. and G. Dagan, 1988, Stochastic analysis of the effects of boundaries on spatial variability in groundwater flows: 1. Constant head boundary, *Water Resour. Res.*, 24(10), 1689–1697.
- Rubin, Y. and G. Dagan, 1992, Conditional estimation of solute travel time in heterogeneous formations: impact of the transmissivity measurements, *Water Resour. Res.*, 28(4), 1033–1040.
- Rubin, Y., G. Mavko, and J. Harris, 1992, Mapping permeability in heterogeneous aquifers using hydrological and seismic data, *Water Resour. Res.*, 28(7), 1192–1700.
- Rubin, Y., 2003, *Applied Stochastic Hydrogeology*, Oxford University Press, Oxford.
- Rubin, Y. and S. Hubbard, 2005, Stochastic forward and inverse modeling: the “Hydrogeophysical Challenge,” Chapter 17 in *Hydrogeophysics*, Eds. Y. Rubin and S. Hubbard, Springer, Netherlands, pp. 487–512.
- Scheibe, T. D. and Y. Chien, 2003, An evaluation of conditioning data for solute transport prediction, *Ground Water*, 41(2), 128–141.
- Schon, J.H., 1996, Physical properties of rocks — fundamentals and principles of petrophysics, Elsevier Science Ltd., Handbook of Geophysical Exploration, Seismic Exploration, Vol. 18, pp. 379–478.
- Schwepppe, F.C., 1973, *Uncertain Dynamic Systems*, Prentice-Hall, Englewood Cliffs, NJ.
- Sheets, K. R. and J. M. H. Hendricks, 1995, Noninvasive soil water content measurement using electromagnetic induction. *Water Resour. Res.*, 31(10), 2401–2409.
- Shuter, E. and W. E. Teasdale, 1981, Application of drilling, coring, and sampling techniques to test holes and wells, in *Techniques of Water Resources Investigations TWI 02-F1*.
- Slater, L. and D.P. Lesmes, 2002, Electrical-hydraulic relationships observed for unconsolidated sediments, *Water Resour. Res.*, 38(10), 1213, doi: 10.1029/2001WR001075.
- Smith, D. G. and H. M. Jol, 1992, Ground-penetrating radar investigation of a Lake Bonneville Delta, Provo level, Bingham City, Utah. *Geology*, 20, 1083–1086.
- Sorensen, K., 1996, Pulled array continuous electrical profiling, *First Break*, 14, 85–90.
- Steeple, D., 2005, Shallow seismic methods, Chapter 7 in *Hydrogeophysics*, Eds. Y. Rubin and S. Hubbard, Springer, Netherlands, pp. 215–252.
- Stone, C., 1995, *A Course in Probability and Statistics*, Duxbury, Boston, MA.
- Tarantola, A., 1987, *Inverse Problem Theory*, Elsevier, New York.

- Telford, W. M., Geldart, L. P., and Sheriff, R. E. 1990. *Applied Geophysics*. Second Edition, Cambridge University Press, Cambridge.
- Topp, G. C., J. L. Davis, and A. P. Annan, 1980. Electromagnetic determination of soil-water content: measurements in co-axial transmission lines. *Water Resour. Res.*, 16, 574–582.
- Treadway, J.A., D.W. Steeples, and R.D. Miller, 1988, Shallow seismic study of a fault scarp near Borah Peak, Idaho, *J. Geophys. Res.*, 93(B6), 6325–6337.
- Ulriksen, P. F., 1982, Application of impulse radar to civil engineering. Ph.D. thesis, Lund University, Sweden.
- VanNostrand, R. G. and K. L. Cook, 1966, Interpretation of Resistivity Data, *U.S. Geological Survey Professional Paper*, PO499, 310pp.
- Ward, S., 1990, Geotechnical and environmental geophysics, in Geotechnical and Environmental Geophysics Environmental and Groundwater, *S.E.G. Investigations in Geophysics* 5.
- Wilt, M., H. F. Morrison, A. Becker, H. Tseng, K. Lee, C. Torres-Verdin, and D. Alumbaugh, 1995, Crosshole electromagnetic tomography: a new technology for oil field characterization, *The Leading Edge*, 14(3), 173–177.
- Woodbury, W. and Y. Rubin, 2000, A full-Bayesian approach to parameter inference from tracer travel time moments and investigation of scale-effects at the Cape Cod experimental site, *Water Resour. Res.*, 36(1), 159–171.
- Woodbury, A. D. and T. J. Ulrych, 2000, A full-Bayesian approach to the groundwater inverse problem for steady state flow, *Water Resour. Res.*, 36(8), 2081–2093.
- Yeh, T.-C. J., S. Liu, R.J. Glass, K. Baker, J. R. Brainard, D. L. Alumbaugh, and D. LaBrecque, 2002, A geostatistically based inverse model for electrical resistivity surveys and its application to vadose zone hydrology, *Water Resour. Res.*, 38(12), 1278.
- Yilmaz, O., 1997, *Seismic Data Processing: Investigations in Geophysics*, Vol 2. Society of Exploration Geophysics, Tulsa, OK.
- Zodhy A. A., G. P. Eaton, and D. R. Mayhey, 1974. Applications of surface geophysics to groundwater investigations. Techniques of Water Resources Investigations of the U.S. Geological Survey, Book 2, Chapter D1.

Further Information

Rubin and Hubbard (2005) provide a compilation of the fundamentals of hydrogeophysics for both the hydrogeological and geophysical perspectives. This book emphasizes the theory, assumptions, approaches, and interpretations that are particularly important for hydrogeological applications, with a focus on electrical resistivity, induced polarization, electromagnetic, GPR, logging, and seismic methods; petrophysical relationships; and case studies. Also geared toward near-surface systems, Ward (1990) provides a compilation of environmental and engineering case studies using geophysical data and Keys (1989) provides a review of well logging methods for groundwater applications. More information about the acquisition, reduction, and interpretation of geophysical methods in general is given by Robinson and Coruh (1988), Burger (1992), and Telford et al. (1990). Additional background on stochastic groundwater hydrology is provided by Degan (1989), Gelhar (1993), and Rubin (2003).

15

Geophysical and Tracer Characterization Methods

15.1	Introduction.....	15-1
15.2	Simulation/Optimization Methods for Aquifer Property Estimation	15-2
	Tracer Tests	
15.3	Application of Geophysical Methods to Environmental Problems.....	15-7
	Surface Seismic Reflection and Refraction Surveys	
15.4	Ground-Penetrating Radar.....	15-10
15.5	Cross-Well Tomography.....	15-14
	Traveltime Tomography • Attenuation Tomography	
15.6	Combining Geophysical and Tracer Data	15-18
	Geostatistical Methods • Zonal Inversion Methods •	
	Estimating the Relation between Geophysical and	
	Hydrogeologic Properties	
15.7	Dynamic Imaging of Tracer and Contaminant Plumes	15-24
	Glossary.....	15-25
	References	15-26
	Further Reading	15-30

David W. Hyndman
Michigan State University

15.1 Introduction

Detailed aquifer properties are needed to predict the fate of groundwater solutes and to design efficient remediation strategies for contaminated aquifers, yet the hydrogeologic data used to estimate these properties are generally local and sparse. Low-resolution aquifer property estimates generally fail to describe the dominant flow and transport paths through an aquifer, resulting in significant errors in groundwater flow and solute transport predictions. One way to improve the resolution of heterogeneous aquifer properties involves combining densely sampled geophysical information with sparse hydrogeologic data in numerical inversions. This chapter provides: a synopsis of simulation-optimization methods; a description of surface

seismic, ground-penetrating radar (GPR), and cross-well seismic tomography methods; and a discussion of approaches to combine geophysical and hydrogeologic data with an emphasis on applications that have incorporated tracer test data.

15.2 Simulation/Optimization Methods for Aquifer Property Estimation

A powerful approach for estimating aquifer properties involves coupling nonlinear optimization methods with numerical groundwater flow and solute transport simulations. Such numerical simulations overcome the difficulties associated with incorporating highly heterogeneous aquifer properties into analytic solutions to the equations that govern groundwater flow and solute transport. The groundwater flow equation is used to solve for hydraulic head (h) given estimates of hydraulic conductivity (K) (and specific storage (S_s) for transient simulations)

$$S_s \frac{\partial h}{\partial t} = \nabla \cdot (K \cdot \nabla h) - W \quad (15.1)$$

where t is time; ∇ is the gradient operator ($\partial/\partial x, \partial/\partial y, \partial/\partial z$), x, y , and z are Cartesian coordinates, and W incorporates the source fluid fluxes. Darcy's law can then be used to approximate the average linear velocity of the fluid (V) given the calculated hydraulic head values as well as estimates of hydraulic conductivity and effective porosity (θ).

$$V = -\frac{K}{\theta} \nabla h \quad (15.2)$$

The advection-dispersion equation is used to approximate the spatial and temporal distribution of solute concentrations (C) using the calculated velocities and given estimates of the hydrodynamic dispersion tensor (D)

$$R \frac{\partial C}{\partial t} = \nabla(D \cdot \nabla C) - V \nabla C + \frac{W(C - C')}{\theta} - B \quad (15.3)$$

where $D = D(\alpha, \nu)$, α is the dispersivity, R is the retardation factor, C' is the solute concentration in the source fluid, and B is the reaction term (i.e., biodegradation). Numerical approximations of these equations, however, can easily incorporate realistic aquifer heterogeneity in three dimensions, while optimization methods provide an efficient method of adjusting aquifer properties to best match all available data.

Despite the power of optimization methods, parameters for hydrogeologic models are still often determined using trial and error calibration, which is inefficient and highly dependent on the knowledge and experience of the model operator. A more efficient and mathematically rigorous approach can be developed using weighted nonlinear least squares *regression*, which can incorporate measurement uncertainty and operator judgment through the assigned sets of weights. Regression approaches provide a consistent statistical framework for estimating aquifer properties.

The goal (or objective function) of least squares regression for aquifer property estimation is generally to minimize the weighted sum-of-squared residuals between measured and simulated values of hydraulic heads, tracer concentrations, and other measured states of the system. Other objective functions include minimizing the sum of the absolute value of residuals, or the squared residuals between some index of the data, such as spatial or temporal concentration *moments* or concentration arrival time quantiles. A well-posed problem will rapidly converge toward the *global minimum* objective value (i.e., no combination of parameters will provide a lower objective value) from a variety of starting parameter sets. Optimization methods work best when the available data (hydraulic head, tracer concentrations, advective front location,

water temperature, velocity, etc.) are highly sensitive to the estimated *aquifer properties*, such as hydraulic conductivity, dispersivity, and effective porosity. These concepts are discussed in more detail along with a discussion of other parameter estimation methods in Aster et al. (2005).

15.2.1 Tracer Tests

The main limiting factor for inferring the detailed heterogeneous structure of aquifers is the sparse nature of hydrogeologic data. Readily available hydrogeologic data include local hydraulic conductivity values from well cores, *pump tests and slug tests*, as well as hydraulic head and tracer concentration measurements at observation wells. Concentration and hydraulic head measurements provide spatially averaged information about aquifer properties, while the limited core and pump test data provide local information about some irregularly shaped region near observation wells. Unfortunately, the size and geometry of the region characterized by single well pump tests or slug tests is unknown because water preferentially flows through high conductivity regions. The hydraulic response during such tests thus provides parameter estimates for the high hydraulic conductivity regions near the well screen. Early time pump test data provide information about properties very close to the well, while later measurements provide information about more distant regions. Borehole flowmeters can also be used to measure flow rates across the screened region of a well due to natural or imposed flow conditions. This approach can provide high resolution information about heterogeneous aquifer properties as discussed by Molz (1994), however such estimates may need to be corrected for well skin effects in some environments (Barrash et al., 2006).

Tracer tests are valuable for estimating aquifer properties, such as effective porosity (θ) and dispersivity (α), yet hydraulic head should be measured at all possible locations during such tests to estimate the hydraulic conductivity field. Tracer concentration histories from a densely sampled well network provide a direct indication of transport velocities, thus providing information about the ratio of hydraulic conductivity to effective porosity for a known hydraulic gradient. Hydraulic head measurements are sensitive to the regional variations in hydraulic conductivity and provide information about the gradients. These measurements are inexpensive to obtain, and densely sampled hydraulic head data may be used in conjunction with seismic data to develop detailed hydraulic conductivity estimates, while head measurements alone provide limited information about the spatial patterns of hydraulic conductivity. In addition, head data can be used as observations for flow simulations, which are less computationally intensive than tracer transport simulations. When geophysical information is also available, it can be jointly analyzed with tracer and hydraulic data to provide high-resolution aquifer property estimates, as will be described in Section 15.6.

15.2.1.1 Natural Gradient Methods

The two main categories of tracer tests are natural gradient tests and forced gradient tests. Natural gradient methods are often preferred because they do not disturb the natural flow system and the measured concentrations are more sensitive to the aquifer dispersivity. The disadvantages of natural gradient tests include their time-consuming nature and the difficulty in designing a downgradient sampling system that adequately measures the tracer plume through time. Classic examples of natural gradient tracer experiments were conducted at the Borden Aquifer in Ontario (relatively homogeneous glacial deposit), the Cape Cod Aquifer in Massachusetts (heterogeneous glacial outwash), and the MAcroDispersion Experiment Site (MADE) in Mississippi (highly heterogeneous alluvial deposit). These tests illustrated that heterogeneous aquifer properties strongly influence groundwater flow and solute transport.

Dispersivity at the Borden site was analyzed using *spatial moments* (M) of the measured concentrations (Freyberg, 1986)

$$M_{ijk}(t) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \theta C(x, y, z, t) x^i y^j z^k \, dx \, dy \, dz \quad (15.4)$$

where θ is effective porosity, C is the concentration of a specific tracer at time t and location (x, y, z) , and the indices (i, j , and k) define the components of each moment; M_{000} is the zeroth spatial moment, which is the total mass; $M_{100}, M_{010}, M_{001}$, are the x, y , and z components of the first spatial moment; and $M_{200}, M_{020}, M_{002}$, are the x, y , and z components of the second spatial moment.

The concentration covariance tensor (σ) can be defined about the center of mass as described by Freyberg (1986):

$$\begin{aligned}\sigma_{xx} &= (M_{200}/M_{000}) - x_c^2 \\ \sigma_{yy} &= (M_{020}/M_{000}) - y_c^2 \\ \sigma_{zz} &= (M_{002}/M_{000}) - z_c^2 \\ \sigma_{xy} = \sigma_{yx} &= (M_{110}/M_{000}) - x_c y_c \\ \sigma_{yz} = \sigma_{zy} &= (M_{011}/M_{000}) - y_c z_c \\ \sigma_{xz} = \sigma_{zx} &= (M_{101}/M_{000}) - x_c z_c\end{aligned}\quad (15.5a)$$

where the center of mass can be defined as:

$$\begin{aligned}x_c &= M_{100}/M_{000} \\ y_c &= M_{010}/M_{000} \\ z_c &= M_{001}/M_{000}\end{aligned}\quad (15.5b)$$

The dispersivity tensor (A , with components $\alpha_l, \alpha_t, \dots$) can then be approximated using

$$A = \frac{1}{2|V|} \frac{d}{dt}(\sigma) \quad (15.6)$$

for cases with negligible molecular diffusion, which is a common assumption for groundwater transport studies due to the larger magnitude of hydrodynamic dispersion. This analysis of the Borden tracer moments showed a significant temporal increase in both the longitudinal and transverse components of the covariance tensor (related to dispersivity with Equation 15.6), which may indicate the effects of the plume encountering more heterogeneity at larger scales (Figure 15.1). In addition, the center of mass sinks through time, especially early in the experiment, and the original plume appears to split into two distinct plumes later despite the relatively homogeneous material at this site. Sudicky (1986) developed detailed estimates of log hydraulic conductivity along two transects at this site using permeameter measurements on cores. These were used to estimate dispersivity coefficients, using the theory of Gelhar and Axness (1983), which were consistent with those obtained from the tracer test analysis.

The tracer tests at the Cape Cod site were used to analyze the heterogeneous hydraulic and chemical properties of the shallow aquifer (Garabedian et al., 1991; LeBlanc et al., 1991; Hess et al., 1992). Each tracer plume was irregular in shape and appeared to sink as it moved downgradient (Figure 15.2). Two of the likely causes for this effect are recharge-induced downward flow, and the slight density contrast between the tracer plume and the regional groundwater. The concentration moment analysis indicated an order of magnitude larger longitudinal dispersion for a reactive tracer (lithium) as compared to a nonreactive tracer (bromide) (Gelhar, 1993), which indicates the need to properly describe both the physical and chemical transport processes of aquifers.

At the MADE site in Mississippi, only about half of the injected tracer could be accounted for at late times as the tracer plume advected down an unidentified channel deposit in advance of the expanding sampling network (Adams and Gelhar, 1992; Boggs et al., 1992). This result indicates the need for a detailed tracer test design that accounts for the likely heterogeneous aquifer conditions. The tracer data clearly illustrated the effect of large-scale heterogeneities in hydraulic conductivity. This was considered

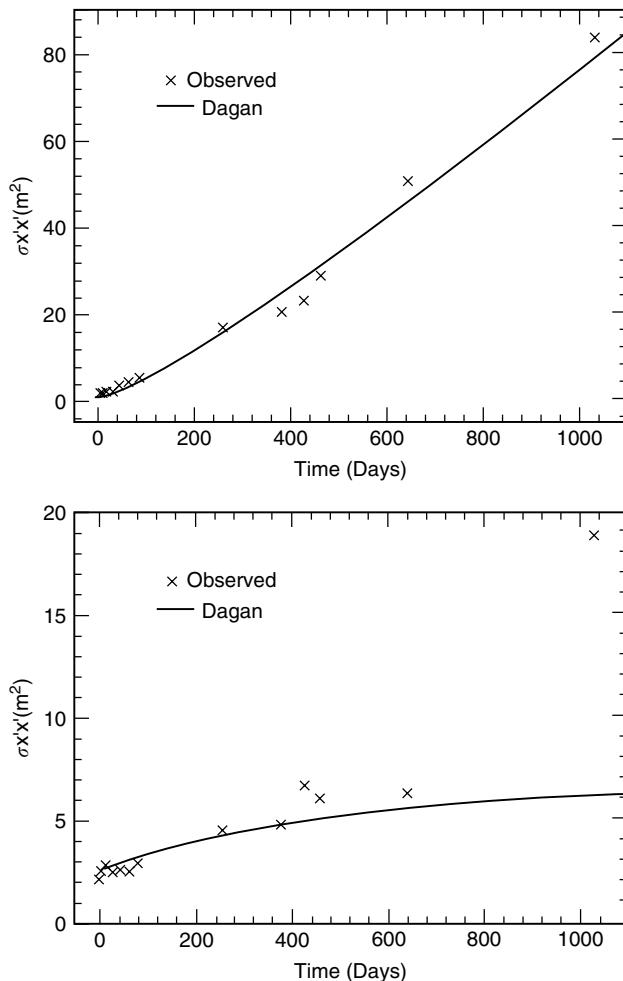


FIGURE 15.1 Plots of estimated longitudinal and transverse components of the spatial covariance tensor illustrating the temporal dependence of these statistics. These values can be used to approximate the dispersivities in these directions. (From Freyberg, D.L. 1986. *Water Resour. Res.* 22(13):2031–2046. With permission.)

a trend in the spatial covariance analysis (Rehfeldt et al., 1992), with the primary hydraulic conductivity estimates obtained using a borehole flowmeter. In addition, this experiment demonstrated the need for high-resolution noninvasive geophysical techniques to identify such large-scale heterogeneities. Barlebo et al. (2004) modeled the tracer transport from the MADE-2 tritium transport experiment by using a finite element model and optimizing hydraulic conductivity values for a small number of zones. They also evaluated detailed flowmeter estimates of hydraulic conductivity, and found that these values had significant random error and bias.

A natural gradient tracer test conducted at the Schoolcraft Plume-A Site near the town of Schoolcraft MI, was reasonably predicted using a three-dimensional finite difference flow and transport model (Biteman et al., 2004; Phanikumar et al., 2005). The hydraulic conductivity dataset for this model was developed using permeameter analysis of repacked core segments (Zhao et al., 2005), and a hydrostratigraphic analysis of minimally disturbed cores to characterize the nonstationary nature of the hydraulic conductivity at the site (Biteman et al., 2004). Unlike most sites where detailed tracer tests are conducted, the Schoolcraft Plume-A site was chosen because of its groundwater contamination. The system used to inject tracer at this site was designed to optimally mix injected solutes and microbes across a vertical transect through the aquifer (Hyndman et al., 2000), and was then used to develop and maintain

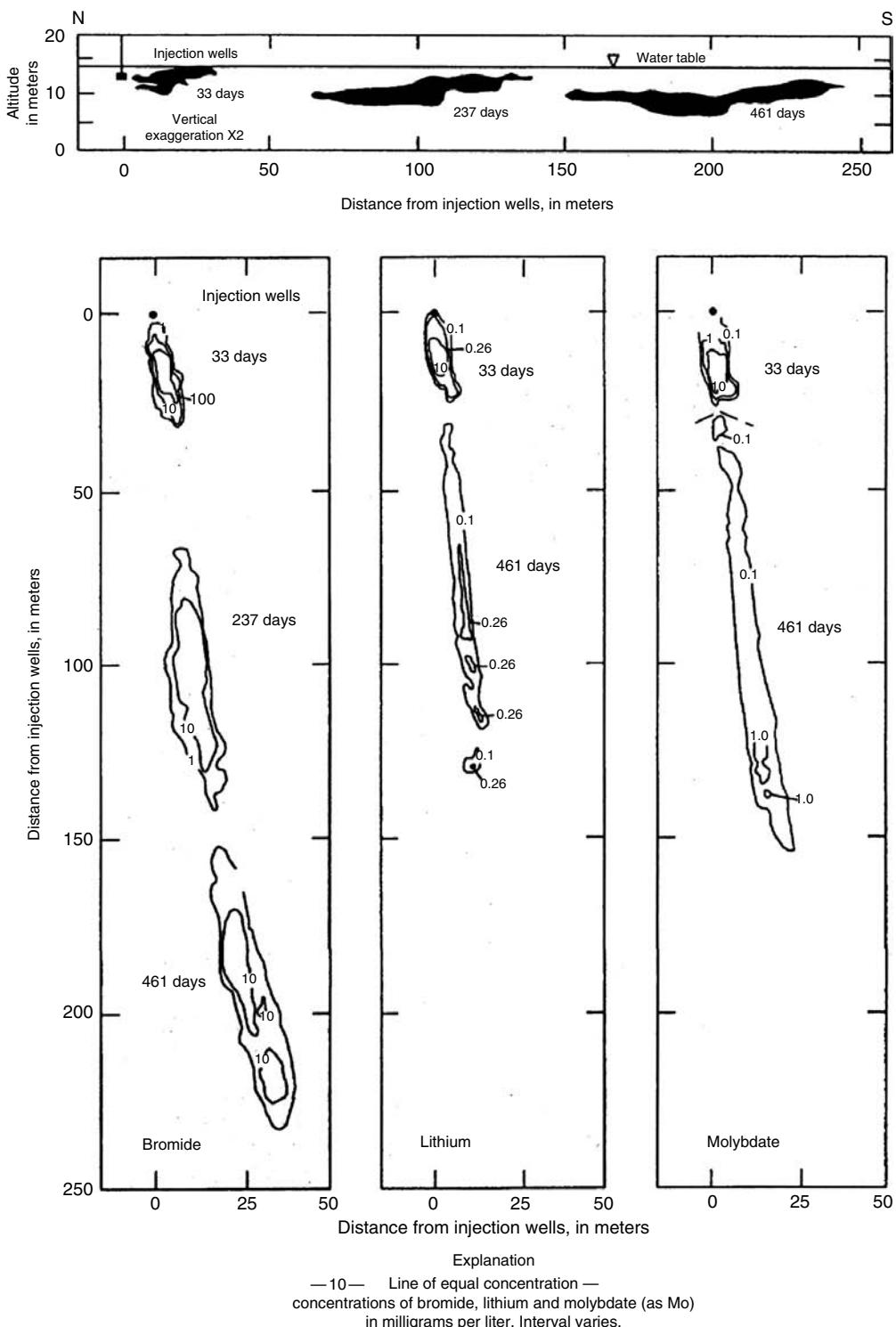


FIGURE 15.2 Vertical and horizontal plots of tracer concentration for the Cape Cod natural gradient tracer test. The variable plume geometry demonstrates the influence of physical heterogeneities on solute transport, while the difference in shapes of the three tracers illustrates the influence of geochemistry on solute transport. (From LeBlanc, D.R. et al. 1991. *Water Resour. Res.* 27(5):895–910. With permission.)

a highly effective bioremediation system (Hyndman et al., 2000; Dybas et al., 2002). Although it was possible to predict tracer transport at this site, the level of characterization required was far more than would be possible at most remediation sites.

15.2.1.2 Forced Gradient Methods

The long duration of natural gradient methods is often inconsistent with the need for rapid site assessment, thus forced gradient tests are often used in practice. The first stage of a forced gradient test is pumping for a period of time to approach steady-state hydraulic heads in the region of the test. In divergent flow tracer tests, water is pumped into a well while concentrations are measured at several different observation wells at radius (r). This type of test can be used to estimate effective porosity (θ) in a homogeneous hydraulic conductivity field using (Sanchez-Villa et al., 1992)

$$\theta = \frac{Qt_r}{\pi r^2 b} \quad (15.7)$$

where Q is the injection flow rate over the entire saturated thickness of the aquifer (b), and t_r is the concentration travel time from the injection well to the observation well (i.e., can be approximated as the peak concentration arrival time or center of mass for a pulse injection of a conservative tracer).

Convergent flow tracer tests can be used to explore anisotropy in hydraulic conductivity and dispersivity. In a convergent tracer test, water is pumped out of one well which is monitored for tracer concentrations derived from small injected tracer pulses from other wells. Each test has to be conducted at different times, or multiple tracers with similar transport characteristics must be injected into the different wells. This type of analysis showed significant directional dependence in tracer concentration histories in an anisotropic fractured gneiss and the Culebra dolomite near the Waste Isolation Pilot Plant (Sanchez-Villa et al., 1992). Simulation experiments performed by Tiedeman and Hsieh (2004) demonstrated that longitudinal dispersivity estimates from forced gradient tracer tests depend on the geometry of wells used in the test, and that in many cases such tests will underestimate the dispersivity for natural gradient tests.

At the Kesterson Aquifer in California's central valley, a forced gradient tracer test was conducted using both an injection and a withdrawal well to assess the heterogeneous properties of a shallow aquifer between these wells (Benson et al., 1991). Benson (1988) performed a series of one-dimensional transport simulations to approximate the longitudinal dispersivities of this site. This site was also used to test several approaches that combine seismic, tracer, and hydraulic data for aquifer property estimation, as described later in this chapter.

15.3 Application of Geophysical Methods to Environmental Problems

Spatially averaged hydrogeologic data can be supplemented with higher resolution geophysical information to develop detailed estimates of aquifer structure and properties. Geophysical techniques, such as seismic and radar methods, have successfully imaged a variety of subsurface features including heterogeneous lithologies in oil reservoirs, aquifers, and mines; the geometry of subsurface cavities and buried objects; variations in fluid properties and saturation; and the structure of glaciers. The following sections discuss (1) an overview of surface seismic and ground-penetrating radar methods; (2) cross-well tomography; and (3) approaches to combining geophysical and tracer data.

15.3.1 Surface Seismic Reflection and Refraction Surveys

Seismic *reflection* and *refraction* surveys have been widely used to characterize subsurface environments for decades. Such surveys provide information about the geometry and nature of subsurface lithologies for both petroleum exploration and environmental remediation work. Reflection and refraction seismic surveys generate sound waves at a near-surface source (e.g., hammer on a steel plate, explosion)

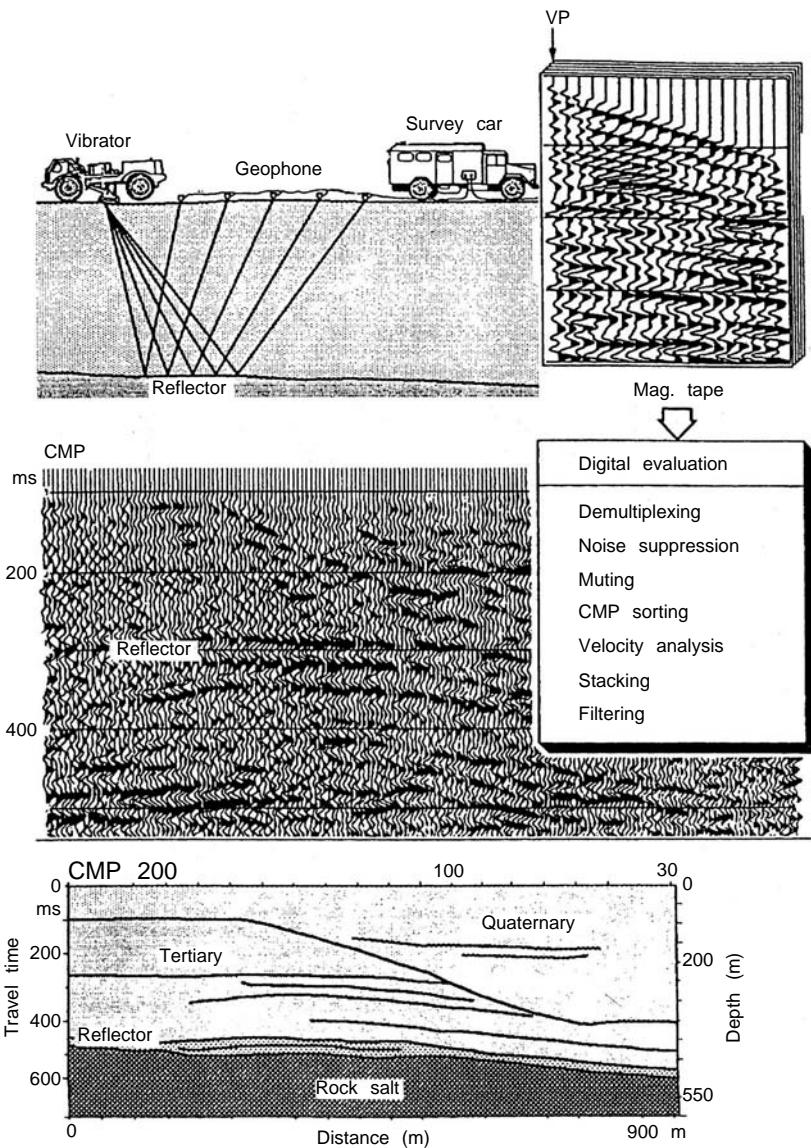


FIGURE 15.3 A diagram of seismic reflection data collection, processing, and interpretation. (From Vogelsang, D. 1995. *Environmental Geophysics: A Practical Guide*. Springer-Verlag, Berlin, 173 p. With permission.)

and measure the resulting seismic signal at multiple receivers called geophones (Figure 15.3). The measured signals are then processed for amplitudes and travel times for refracted and reflected seismic energy.

A portion of the incident seismic energy is reflected when it encounters a contrast in seismic impedance, which equals density (ρ) times seismic velocity (V), and the amount of reflected energy is proportional to the degree of contrast. The time (t) for reflected seismic energy to reach the surface from a single flat interface is:

$$t = \sqrt{\frac{4d^2}{V^2} + \frac{x^2}{V^2}} \quad (15.8)$$

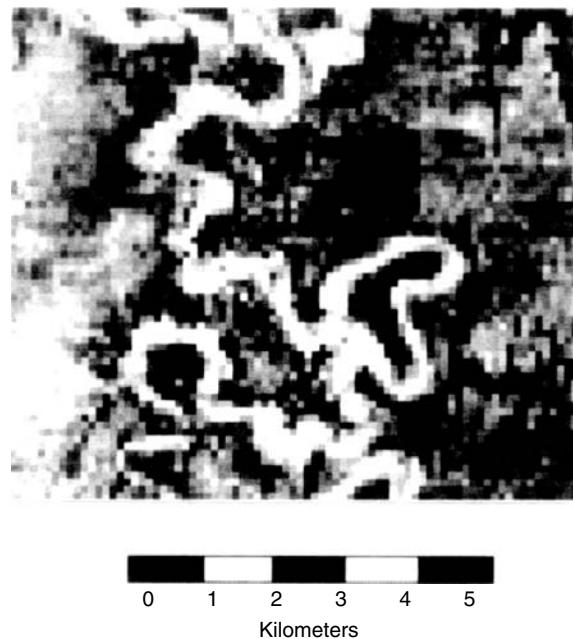


FIGURE 15.4 A horizontal slice through a 3D processed seismic reflection survey from the Gulf of Thailand that illustrates a buried meandering river channel. (Data courtesy of the Stanford Exploration Project.)

where d is the depth to the interface. The first term under the square root represents the time for a normal incidence reflection (source and receiver at same location), while the second term represents parabolic moveout due to offset x between source and receiver. More complex formulations for multiple dipping and curved interfaces can be found in introductory geophysics texts (e.g., Telford et al., 1990). Surface seismic reflection data can be processed by adjusting the velocity for a *common midpoint gather* to flatten the observed hyperbolic shape of the reflector, which is called a *normal moveout correction*.

Through this correction, the measured times are corrected from their offset positions back to a normal incidence position (no offset). The data are then stacked across each corrected common midpoint gather to estimate the seismic velocity structure at the midpoint of the plane between source and receiver locations (see Figure 15.3). This method does not account for interfaces that dip out of the plane, so the energy reflected from out of the image plane will appear at incorrect locations. As a result, the best approach is to collect and process surface seismic data in three dimensions. An excellent example of the type of information that can be obtained by 3D analysis of surface seismic data is the clear evidence of an old meandering channel that can be observed in the processed data from the Gulf of Thailand (Figure 15.4). The fundamentals of seismic reflection methods are described in more detail in Steeples and Miller (1990).

Seismic refraction surveys are collected with far offset receivers to measure energy that travels along an interface with a higher velocity at depth (Figure 15.5a). Refracted arrivals can be discerned from directly reflected arrivals by the slope of the distance versus time curve (Figure 15.5b). Direct reflected arrivals for a two-layer geologic setting will fall on a line that intersects the origin, while a line connecting refracted arrivals will have a positive zero-offset intercept. To ensure refracted energy is the first to arrive at the receivers, a refraction survey needs to be several times longer than the depth of investigation and velocity must increase with depth. The slope of the time versus distance line at the appropriate offset gives the reciprocal of the velocity of the deeper layer (Figure 15.5). In complex heterogeneous systems, detailed velocity analysis and ray tracing may be necessary to infer velocities from measured seismic data (Claerbout, 1992). More detailed descriptions of surface seismic surveys are covered in texts such as Telford et al. (1990).

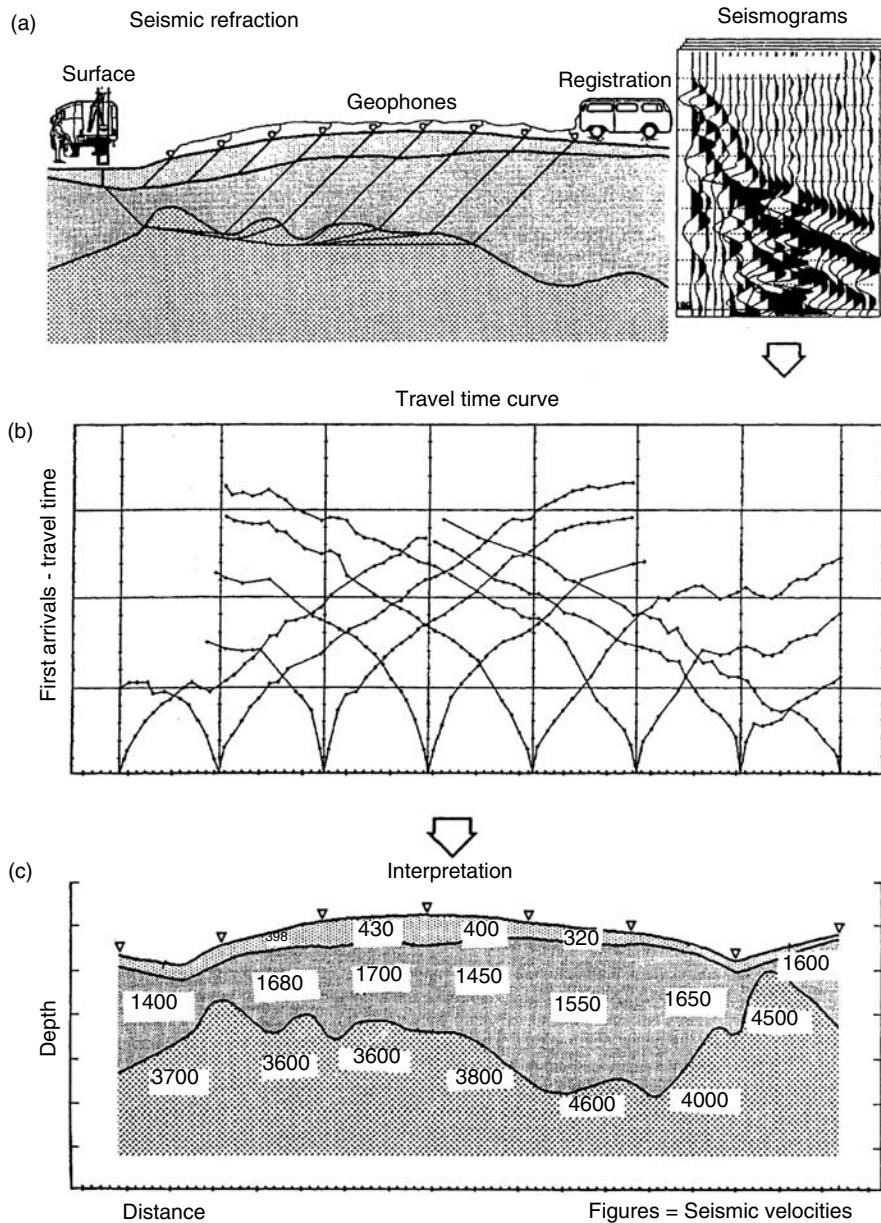


FIGURE 15.5 A diagram of a seismic refraction survey with the corresponding distance versus time plot for direct, reflected, and refracted waves. (From Vogelsang, D. 1995. *Environmental Geophysics: A Practical Guide*. Springer-Verlag, Berlin, 173 p. With permission.)

15.4 Ground-Penetrating Radar

Ground-penetrating radar (GPR) methods typically involve generating high-frequency electromagnetic waves (typically 10–1500 MHz) at a source antenna directed toward the ground and measuring the energy after it has moved through the subsurface at a receiver antenna. There is a trade-off between the high resolution obtained with high-frequency methods and the corresponding rapid signal *attenuation*. Preliminary surveys are often conducted with low-frequency radar antennas (50–100 MHz) that have

deeper signal penetration, with detailed higher frequency (200–900 MHz) surveys conducted in selected regions where higher resolution information about shallow properties is needed. The propagated direct wave (in cross-well GPR) or the reflected wave (in surface GPR) is measured at a receiver antenna. Surface GPR data are generally collected with a single spacing between antennas to provide nearly normal incidence reflections, although more information can be obtained using multiple offsets around each central point of interest, called common midpoint gathers, as described in Section 15.3.1. In both cases, multiple traces should be collected at each measurement point and stacked (or averaged) to improve the signal-to-noise ratio. Normal incidence reflection surveys have the benefit of rapid collection, since antennas can be towed across the surface at a constant rate while a signal is generated at a specified delay. This allows large regions to be rapidly surveyed and variable towing rates can be compensated using a timing mark at specified locations in a survey grid. A global positioning system (GPS) is a valuable addition to large radar surveys to continuously monitor the location of the radar system and map topography. Topographic data is necessary to correct the interpreted GPR depths to elevations above a reference, and lateral location information allows for rapid surveys of irregularly gridded regions.

Collecting traces in common midpoint gathers provides a suite of data that can be processed using surface seismic methods, yet the survey is more difficult to obtain because radar systems with multiple antennas are not readily available. The measured wavefields are processed for the first arrival traveltime and wave amplitudes. These wave attributes can then be processed using normal moveout corrections and mapped into depth versus amplitude plots which can be compared to lithological and moisture content variations in cores across the survey region. Tools that are commonly used to process seismic data such as frequency filtering, time-dependent gain, and migration can be used to improve interpretation of GPR data as described by Knoll et al. (1991) for the Cape Cod Aquifer. Rea and Knight (1998) estimated the near horizontal *variogram* (correlation vs. offset plot) of processed radar data from a gravel pit in southwestern British Columbia and compared this to the modeled variogram of a digitized image of a cliff face at the site. They found good agreement between both the direction of maximum correlation and the estimated correlation length determined from the two types of data. They then used this approach to infer the paleoflow direction of a buried channel and the correlation length in that direction from a three-dimensional GPR survey (Figure 15.6) over this site.

The subsurface propagation of electromagnetic energy depends on the electrical properties of the ground, commonly described using the relative dielectric permittivity ϵ_r in the electromagnetic wave equation

$$\nabla^2 E + i\mu\omega\sigma E + \omega^2 \epsilon^* \mu E = 0 \quad (15.9)$$

where E is the electrical field vector {Volts/m}; $i = \sqrt{-1}$; μ is the magnetic permeability {Henry/m}; ω is the angular frequency {radian/sec} = $2\pi\nu$ (ν = frequency in Hz); σ is the electrical conductivity {Siemen/m}; and ϵ^* is the complex dielectric permittivity of the media {Farad/m}.

The full electromagnetic wave equation is not generally used because of its complexity and uncertainty in all the parameters (i.e., μ , σ , ϵ^*). Instead, the imaginary portion of the dielectric permittivity is generally ignored, unless the goal is to describe electromagnetic wave attenuation, while the real portion of the relative dielectric permittivity is used to describe radar propagation velocities using

$$V = \frac{c}{\sqrt{\epsilon_r}} \quad (15.10)$$

where V is the electromagnetic wave velocity {m/nsec}; c is the free space light velocity (0.3 m/nsec); ϵ_r is the relative dielectric permittivity = (ϵ/ϵ_0) {dimensionless ratio}; ϵ is the real portion of the media dielectric permittivity; and ϵ_0 is the real portion of the free space dielectric permittivity (8.854×10^{-12} Farad/m). This approach has been successful in low electrical loss materials such as dry sand, which allow reasonable radar penetration depths.

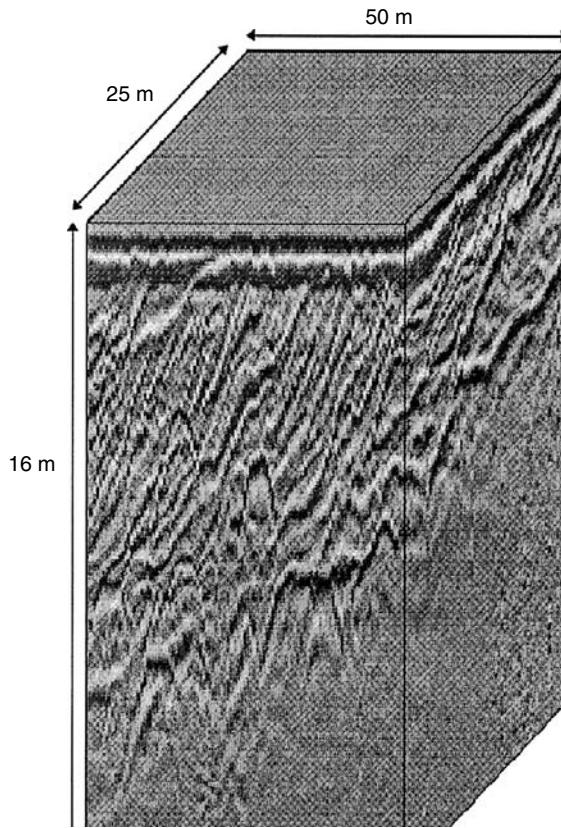


FIGURE 15.6 A 3D GPR survey conducted with 100 MHz antennas over dipping foreset beds in southwestern British Columbia. (From Rea, J. and Knight, R. 1998. *Water Resour. Res.* 34(3).)

The relative dielectric permittivity represents the ability of an electric field to polarize a material and is dependent on properties of both the solid and fluid. This property is inversely proportional to frequency, especially in saturated sediments (Telford et al., 1990). Some typical values are plotted in Figure 15.7 for high-frequency (> 100 MHz) measurements.

The large contrast between the dielectric permittivity of water and that of dry sediment can be used to infer the saturation of sediments from radar measurements. In this sense, it is possible to dynamically image wetting fronts, which would provide information about the hysteretic nature of unsaturated flow (e.g., Binley et al., 2001; Alumbaugh et al., 2002; Lunt et al., 2005). The capillary fringe often appears as a dominant reflector in radar records because of the large vertical gradient in saturation at this depth (Kim et al., 2000). The reflector depth should only be interpreted as the water table in coarse-grained sediments, as the capillary fringe will be significantly shallower than the water table in fine-grained sediments. A detailed GPR survey of the Cape Cod Aquifer imaged the capillary fringe as well as lithologic variability in the saturated zone (Figure 15.8) (Knoll et al., 1991). The GPR penetration depth is reduced in the saturated zone as a large portion of the incident energy is reflected from the capillary fringe due to the contrast in dielectric constant. At other sites, low-frequency radar (25 to 200 MHz) has provided deeper penetration and has been used to image lithologic contrasts in the saturated zone, although the trade-off for increased depth is decreased spatial resolution.

GPR has significant potential to image lithologic variations in the unsaturated zone due to the different water retention properties of sediments (Van Dam et al., 2003). Thus, lithologic heterogeneity may be detected by comparing the results of two radar surveys made under different moisture conditions. Although the dielectric permittivity of clay is different from sand with the same moisture content

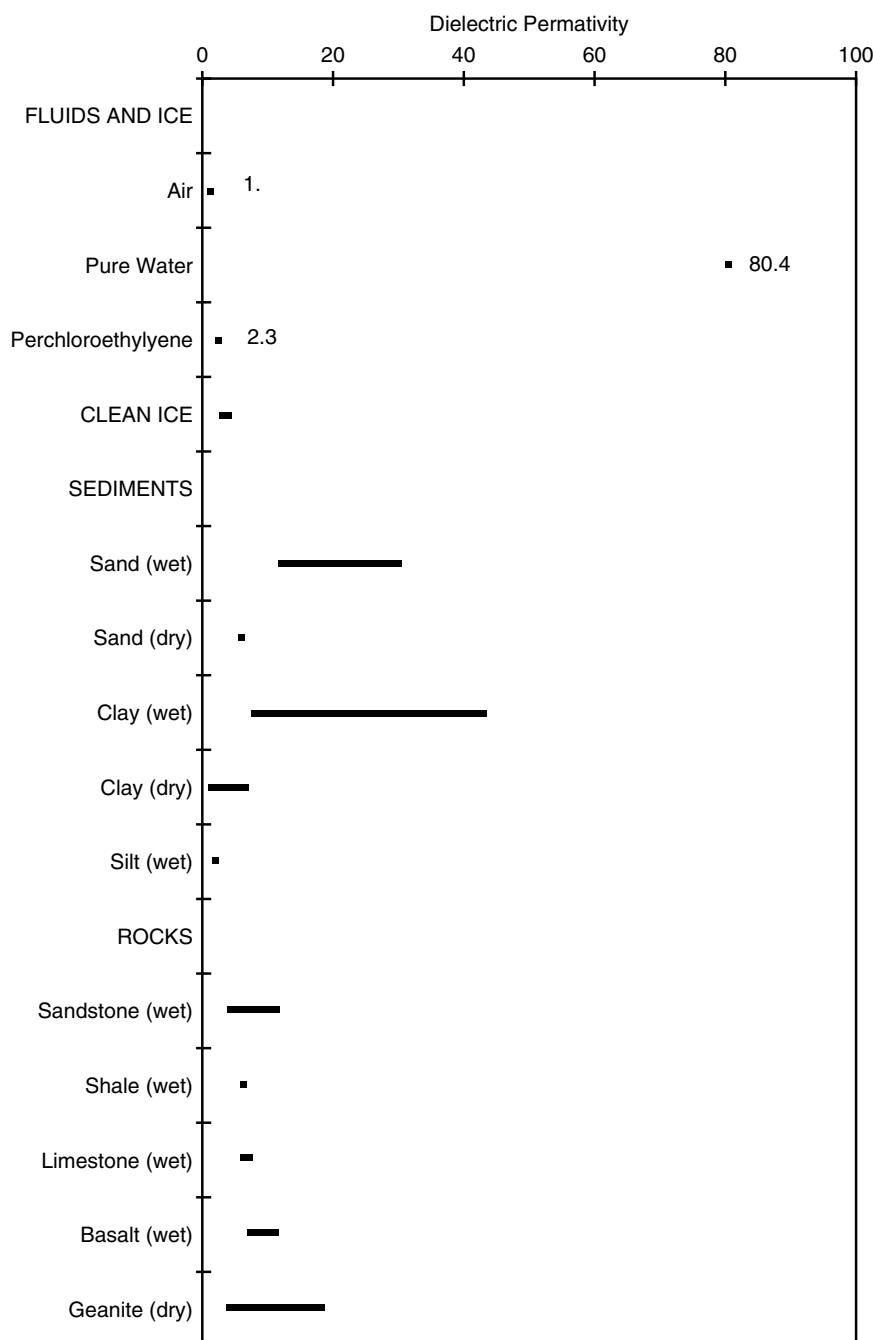


FIGURE 15.7 Plot of relative dielectric permittivity values of various materials. (Data from Placzek, G. and Haeni, F. P. 1995. U.S. Geological Survey Water Resources Investigations Report 95-4009, Denver, CO.; Beres, M. and Haeni, F. P. 1991. *Ground Water* 29(3):375–386; Arcane, S. A., Lawson, D. E., and Delaney, A. J. 1995. *J. Glaciology*. 41(137):68–86; Telford, W. M., Geldart, L. P., and Sheriff, R. E. 1990. *Applied Geophysics*. 2nd ed. Cambridge University Press, Cambridge.)

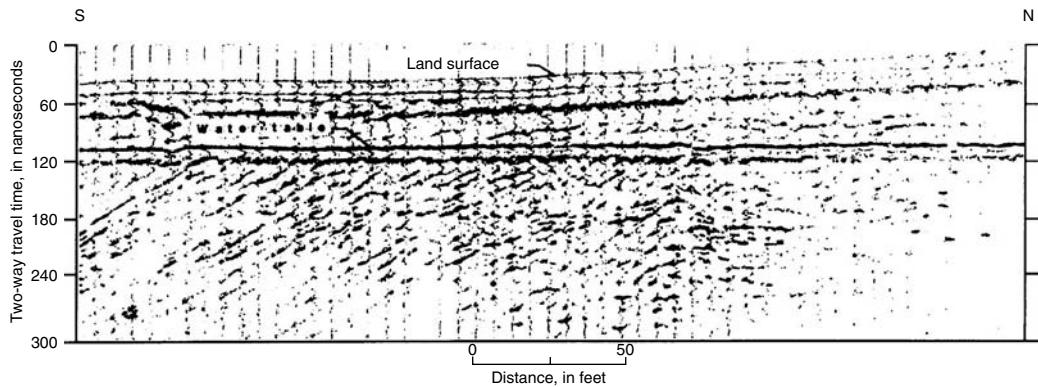


FIGURE 15.8 A single line of processed radar data from a 3D survey over the Cape Cod Aquifer. The land surface and capillary fringe (labeled as water table) are noted on the plot. (From Knoll, M. D., Haeni, F. P., and Knight, R. J. 1991. G. E. Mallard and D. E. Aronson, eds., U.S. Geological Survey toxic substances hydrology program — Proceedings of the technical meeting, March 11–15, 1991, Monterey, CA. With permission.)

(Figure 15.7), the change is more pronounced after a wetting front has passed through regions with different clay content as zones with higher clay content will have a higher moisture content in this case due to stronger capillary forces in the small pore spaces. Greaves et al. (1996) interpreted variations in moisture content from the high-quality multichannel radar data collected at the Chalk River site, Ontario, which were processed by Fisher et al. (1992) (Figure 15.9). The clarity of this processed data shows the ability of multichannel GPR surveys to infer heterogeneities in lithology and water content.

15.5 Cross-Well Tomography

Cross-well tomography is a high-resolution approach to image heterogeneities between pairs of wells. The philosophy of cross-well tomography is similar to that of the computed axial tomography (CAT) scan, except that seismic or radar waves are used to image the subsurface rather than using X-rays to image the human body. The resolution of cross-well tomography is much lower than what is obtained from medical imaging techniques because much lower frequencies are used for seismic imaging (1–20 KHz) and radar imaging (50–1500 MHz), and the coverage is limited to planes between well pairs, whereas three-dimensional coverage is available in a CAT scan.

Cross-well radar surveys provide the opportunity to image subsurface features with higher resolution and more control over the path of energy between source and receiver than surface radar surveys. In the cross-well radar configuration, the transmitted radar signal is measured in a receiver well, thus cross-well tomography can estimate both the attenuation and velocity of electromagnetic energy throughout the plane between wells (Olhoeft, 1988; Johnson et al., 2005). Radar signals are strongly attenuated by clays and saline fluids, thus limiting their use in these environments. This attenuation, however, can be used to identify such zones using short offset cross-well surveys in regions where clay lenses are embedded in a sand matrix.

Cross-well seismic tests involve generating sound waves at specified depth intervals in one well, while hydrophones or geophones measure the propagated wavefield at multiple depths in nearby wells. A variety of sources can be used including (1) a weight drop, (2) a hydraulic pulse generator, (3) an air pulse, or (4) a piezoelectric bender bar. Piezoelectric sources provide a high-quality, frequency-controlled signal by applying an electrical impulse to material that expands due to this charge. The received wavefield is analyzed for key events, such as the compressional- and shear-wave arrival times at each receiver (Telford et al., 1990) or reflections (Lazaratos et al., 1993).

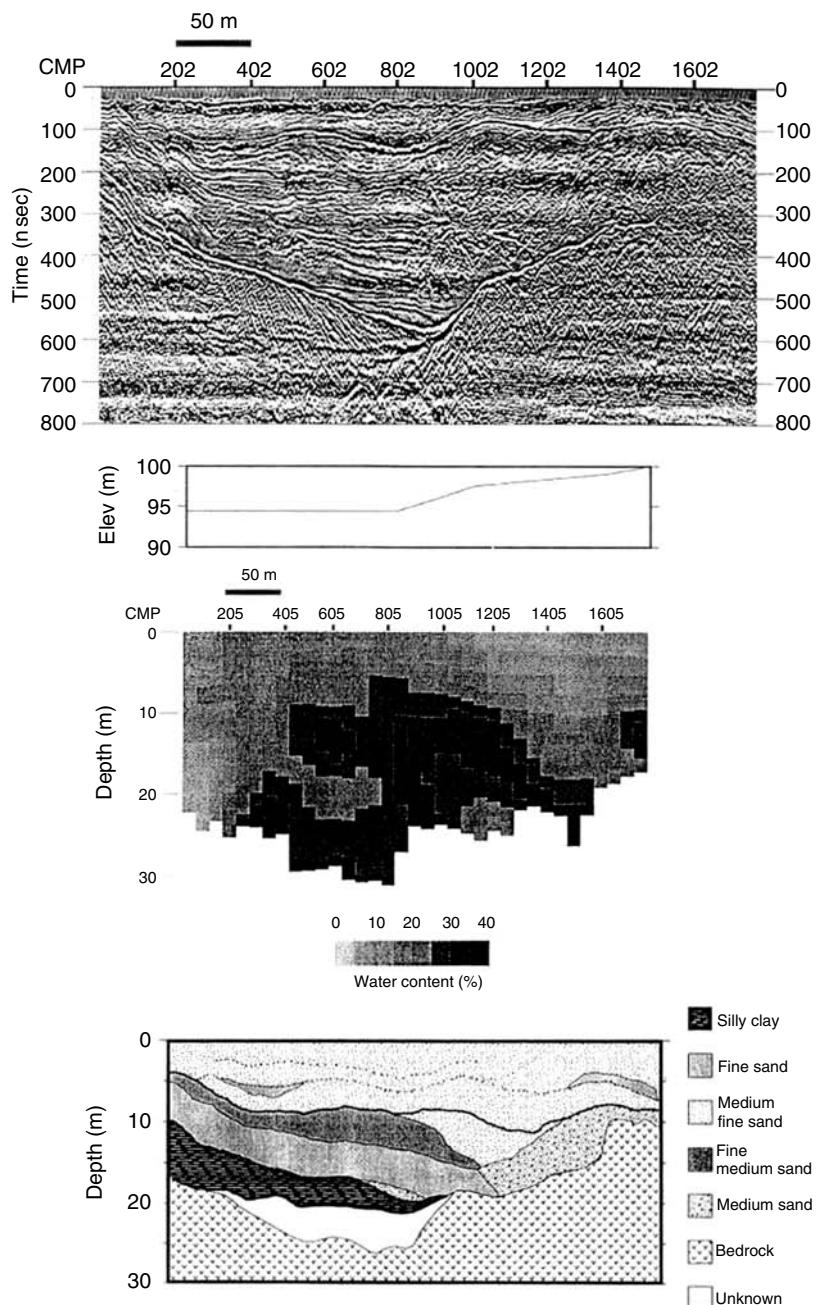


FIGURE 15.9 Plot of processed multiple offset GPR data collected at the Chalk River test site along with interpretations of water content and lithology for the surveyed region. (From Greaves, R. J., Lesmes, D. P., Le, J. M., and Toksoz, M. N. 1996. *Geophysics* 61(3):683–695. With permission. Original presentation in Fisher, E., McMechan, G., and Annan, A. P. 1992. *Geophysics* 57:495–504.)

Cross-well tomography fills a gap between sparse well logs and low-resolution surface geophysical estimates (Harris et al., 1996). These geophysical methods involve estimating seismic or radar properties that best reproduce aspects of the measured wavefield through wave propagation simulations. The best possible vertical resolution is a quarter wavelength (λ), which can be roughly calculated using the dominant frequency (f) of the source as $\lambda = V/f$, where V is the average seismic or radar velocity. This indicates that cross-well seismic tomography has the potential to image lithologic heterogeneities at scales from 10 cm to several meters, for typical sediment velocities (1000–2500 m/sec) (Telford et al., 1990; Kearey and Brooks, 1991) and source frequencies (1–10 KHz), while radar has resolution of less than 1 cm to more than 1 m for typical sediment velocities (0.05–0.15 m/nsec) and antenna frequencies (25 and 1500 MHz).

One seismic tomography approach involves simulating wave propagation through a homogeneous media with embedded scattering objects using the elastic wave equation

$$\rho \frac{\partial^2 \Delta}{\partial t^2} = (\lambda + 2\mu) \nabla^2 \Delta \quad (15.11)$$

where the compressional-wave velocity $\alpha^2 = (\lambda + 2\mu)/\rho$, Δ is the dilatation, or change in volume per unit volume due to strain, λ is Lame's constant, μ is the shear modulus, and ρ is the bulk density. Diffraction tomography provides one method of approximating wave propagation through a background model with embedded point objects, thus heterogeneous media can be split up into different levels of heterogeneities (Harris and Wang, 1996). Reiter and Rodi (1996) presented a nonlinear waveform tomography method that relaxes some of the assumptions necessary for diffraction tomography, however this approach is highly computationally intensive.

15.5.1 Traveletime Tomography

Rather than simulating the dynamics of wave propagation, this approach assumes that waves are of infinite frequency and thus can be represented using rays, which is fairly reasonable for high-frequency cross-well tomography. In this case the goal of the estimation problem is to minimize the absolute or squared difference between measured and simulated traveltimes for a particular type of seismic wave, such as a compressional wave. The wave equation can be simplified to the travelttime equation using the ray theory approximation

$$t_i = \int_{C_i} S(x, z) d\ell \quad i = 1, \dots, M \text{ rays} \quad (15.12)$$

where the travelttime (t_i) for source–receiver pair (i) is approximated as the integral of slowness (S) = (1/velocity) along the simulated ray path (C_i), and () is the increment of path length along the i th ray. Rays can be traced using a numerical approximation of Snell's law (which was originally derived for optics)

$$\frac{\sin \phi_1}{V_1} = \frac{\sin \phi_2}{V_2} = \frac{\sin \phi_i}{V_i} = \frac{\sin \phi_{i+1}}{V_{i+1}} \quad (15.13)$$

where ϕ_i is the incidence angle between the ray in zone i and the normal to an interface separating zone i from zone $i + 1$, and V_i is the seismic velocity in zone i . The top portion of Figure 15.5 illustrates an example from surface seismic characterization, in which refraction (or bending) of rays as they travel through progressively higher velocity regions with depth. The travelttime equation is nonlinear as both the ray path and the slowness field are unknown, thus we cannot directly invert this equation to estimate the slowness field. For each iteration, the travelttime equation can be linearized by iteratively tracing ray paths

through a known background slowness field (S_0) to calculate slowness perturbations (ΔS).

$$S = S_0 + \Delta S \quad (15.14)$$

The slowness estimate from the previous iteration is used for the known background field (S_0), except on the first iteration when an arbitrary starting model is used. A critical decision for tomographic inversions is the appropriate parameterization of the slowness model ($S(x, z)$). The most common parameterization for cross-well tomography has been rectangular constant-velocity pixels (S_j) as this simplifies the estimation problem. The solution to the traveltime equation reduces to a linear set of equations that relate traveltimes (Δt_i) to slowness perturbations in each of j pixels (ΔS_j).

$$\Delta t_i = \sum_{j=1}^N \Delta S_j \ell_{ij} \quad i = 1, \dots, M \text{ rays} \quad (15.15)$$

where the matrix (ℓ_{ij}) has elements of ray path length in each pixel when the slowness field is defined using rectangular pixels.

The slowness perturbations (ΔS_j) are calculated by inverting Equation 15.15 using an algorithm such as the simultaneous iterative reconstruction technique (McMechan et al., 1987). This algorithm simultaneously back-projects the traveltime residuals along the estimated ray paths and averages the result to determine the slowness perturbation to assign to each cell.

The slowness field can be parameterized in different ways depending on the goal of the inversion. The most common parameterization for tomography is a field of rectangular pixels, which are assigned slowness values using Equations 15.12 through 15.15. An alternative approach was presented by Michelena and Harris (1991) who parameterized the slowness inversion using “natural pixels” that account for the data collection geometry in cross-well tomography. These natural pixels are rectangular elements aligned in the direction of wave propagation. This approach can account for the sensitivity of the measured signal at a receiver to some average of the properties along a rectangular or ellipsoidal beam between wells.

For cases where the goal of the geophysical survey is identification of lithologic zones (e.g., sand, gravel, and clay), better results may be obtained by limiting the number of possible slowness values to the number of major lithologies identified in cores. For example, Hyndman and Harris (1996) parameterized a cross-well inversion using a small number of velocity values to estimate the geometry of the primary lithologies of the shallow Kesterson Aquifer in California’s San Joaquin Valley. This approach, called the multiple population inversion (MPI), converts traveltimes between all available well pairs for the spatial distribution of a small number of seismic slowness populations. This can provide more information about the dominant scale of subsurface heterogeneity than common parameterizations that provide smooth slowness fields. The lithologies and hydraulic parameters for the estimated populations can then be inferred from core data and hydraulic testing, or a co-inversion can be used to simultaneously estimate all parameters of interest as discussed in the next section of this chapter.

The MPI method iteratively assigns pixels to slowness populations based on the histogram of slowness perturbations, which are calculated by inverting Equation 15.15. By constraining the number of slowness values, this method is less susceptible to ray-based inversion artifacts and can resolve finer-scale sedimentary structures than methods that smooth the slowness field. Hyndman and Harris (1996) showed that the MPI converges to an equal or smaller average travelttime residual than obtained with an approach based on nonzonal parameterization for both a 2D synthetic aquifer, and for a 3D aquifer case from the Central Valley of California.

This approach reduced the artifacts commonly observed in ray-based inversions without smoothing the estimated field, which is the common method of reducing such artifacts. As tomography is an iterative estimation approach, the zonal slowness model can be used as the background model for an unconstrained velocity inversion. At the Kesterson site, this provided an updated seismic slowness estimate that preserved much of the structure of the zonal slowness estimates with finer-scale resolution (Figure 15.10). These

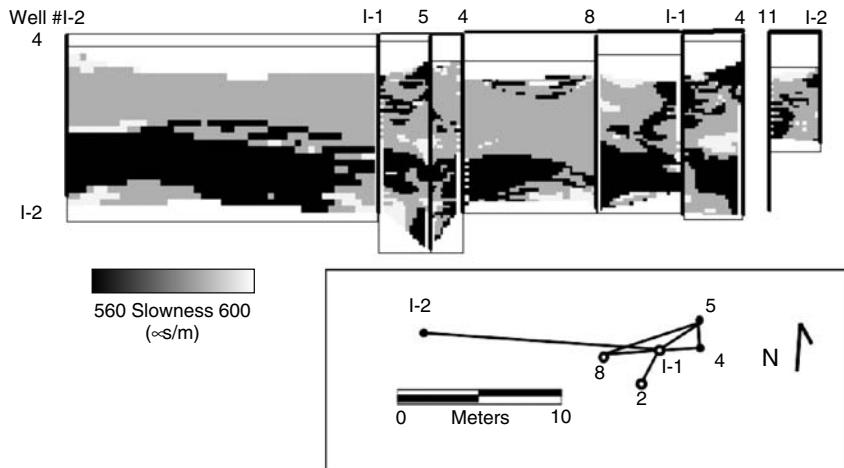


FIGURE 15.10 Location map for the Kesterson Aquifer site along with seismic velocity tomograms estimated by Hyndman and Harris. The low slowness zones were identified as low hydraulic conductivity regions. Note that well I-2 is the injection well, and I-6 is the extraction well, which is approximately the same distance from well 4 to well I-2, but to the right of well 4 along the approximate path from I-2 to 4. Well I-6 data are shown in Figure 15.11 (Modified from Hyndman, D. W. and Harris, J. M. 1996. *Geophysics* 61(6):1728–1737. With permission.)

Kesterson tomograms are used in later sections of this chapter for hydraulic and transport property estimation in three dimensions.

15.5.2 Attenuation Tomography

Traveltime tomography uses only one attribute of the measured wavefield, while attenuation tomography attempts to reproduce wave amplitudes due to wave propagation along the paths predicted by traveltome tomography. One philosophy of attenuation tomography is that the bulk of the seismic or radar energy travels near the ray paths predicted for traveltome tomography, thus the amplitude attenuation can also be described along such paths. Thus the attenuation coefficients can be iteratively updated to best reproduce some aspect of the amplitudes measured at the receivers. For example, Quan and Harris (1997) demonstrated an approach that used the observed frequency downshift in a window around the first arrival to estimate attenuation coefficients for an oil reservoir in Texas. If high-quality signals are measured, significant information can be gained about the lithologic heterogeneities through analysis of both attenuation and traveltome tomography results. Another approach to attenuation tomography is to minimize the difference between the simulated and measured amplitudes of the first arrival peak or energy in the first peak. Johnson et al. (2005) parameterized a radar tomography problem using fresnel zones, and calculated the attenuation in this region based on the difference in energy in the first half cycle of the received signal.

15.6 Combining Geophysical and Tracer Data

Geophysical and hydrogeologic data can be combined to provide aquifer property estimates over a range of spatial scales. Methods that estimate aquifer properties can benefit from different sampling scales for each type of data. Lithologic cores provide point values, while well-test data provide local information around a well. Cross-well seismic and radar tomography provides dense spatial information for a cross section between wells, while surface geophysical surveys can provide 3D information about variability in shallow lithologies and fluid saturation. Measured hydraulic heads and tracer concentrations are indirectly related to hydraulic conductivity through the groundwater flow and advection-dispersion equations. The scale measured by each type of data is different but overlapping, which provides a suite of information about spatial variability that could not be fully obtained using any single type of data alone. These independent data types complement each other because they contain information about a range of important aquifer scales.

15.6.1 Geostatistical Methods

Geostatistical methods provide one set of tools to combine geophysical and hydrogeologic data. These methods interpolate between available data points using weights that represent modeled spatial correlation structures and the uncertainty in the different measurement methods. Weights are assigned through a covariance matrix that incorporates the inferred spatial correlation of each variable and the interrelationships between variables. An excellent summary of basic geostatistical concepts is provided by Isaaks and Srivastava (1989) and a suite of powerful geostatistical computer codes is available in Deutsch and Journel (1997).

Geophysical surveys provide a good source of soft information for geostatistical estimation methods. Several geostatistical methods have been developed that use surface seismic and radar datasets to provide information about the heterogeneous structure of aquifers and reservoirs. The simplest approach is to *cokrig* the different datasets based on the inferred correlation structures of each variable and cross-correlation between datasets. The benefit of adding high-resolution seismic data to well-log data has been demonstrated in a number of studies. For example, Araktingi and Bashore (1992) cokriged 3D surface seismic velocity estimates with porosity measurements and found that even poor-quality seismic estimates provided information about reservoir properties. However, traditional cokriging cannot incorporate nonlinear and potentially non-unique relations between a parameter of interest and the accompanying soft geophysical estimates.

Sequential Gaussian co-simulation (Deutsch and Journel, 1997) provides an additional simple geostatistical method of developing hydraulic conductivity realizations based on hydraulic conductivity measurements as hard data and the slowness estimates as soft data. Sequential Gaussian simulation honors the values at all seismic tomogram locations, as well as the sample probability distribution and variograms of the tomographic slowness estimates. For the Kesterson Aquifer in California's San Joaquin Valley (Hyndman et al. 2000) found a correlation coefficient of +0.74 between slowness and natural log hydraulic measurements, which establishes the level of control the seismic data has in the geostatistical simulation (Figure 15.11). If this coefficient were +1.0 the seismic data would have a perfect positive linear correlation to log conductivity, with large slowness values indicating large $\ln(K)$ values. If this coefficient were 0.0, the seismic data would have no linear correlation to log conductivity and would thus have no influence on the $\ln(K)$ estimates. One limitation of this approach is the reliance on a linear relation between parameters.

Figure 15.11 demonstrates the value of adding seismic information to the conductivity estimation problem in cases where significant correlation is observed between slowness and log conductivity. The log conductivity field in Figure 15.11a is a randomly chosen realization generated using sequential Gaussian simulation of log hydraulic conductivity data alone, which results in the corresponding simulated concentration histories. Figure 15.11b shows a log conductivity realization generated using sequential Gaussian co-simulation for the same random path that was used to generate Figure 15.11a. For this realization, a slowness realization conditioned on the estimated seismic tomography planes (Figure 15.10) was used as soft data and hydraulic conductivity measurements were hard data. Thus to obtain the conductivity estimates first involved generating a slowness realization using sequential Gaussian simulation, then this field was updated using sequential Gaussian co-simulation to constrain the regions that are near conductivity measurements.

In the Kesterson case, Hyndman et al. (2000) illustrated that the slowness estimates improved the concentration histories at all wells, except for at well 8, relative to the simulations through the estimate based on conductivity data alone (Figure 15.11a). This improvement would likely be more pronounced if fewer conductivity measurements were available. In addition, the simulation with conductivity data alone required estimates of the correlation structure, which in this case were derived from the seismic estimates. Thus the simulations with conductivity alone would have been poorer if no seismic data were available to help infer the correlation structure of subsurface lithologies.

Bayesian geostatistics provides a more flexible alternative to traditional cokriging for incorporating "soft" geophysical data. Carty and Rubin (1995) developed a conditional log hydraulic conductivity

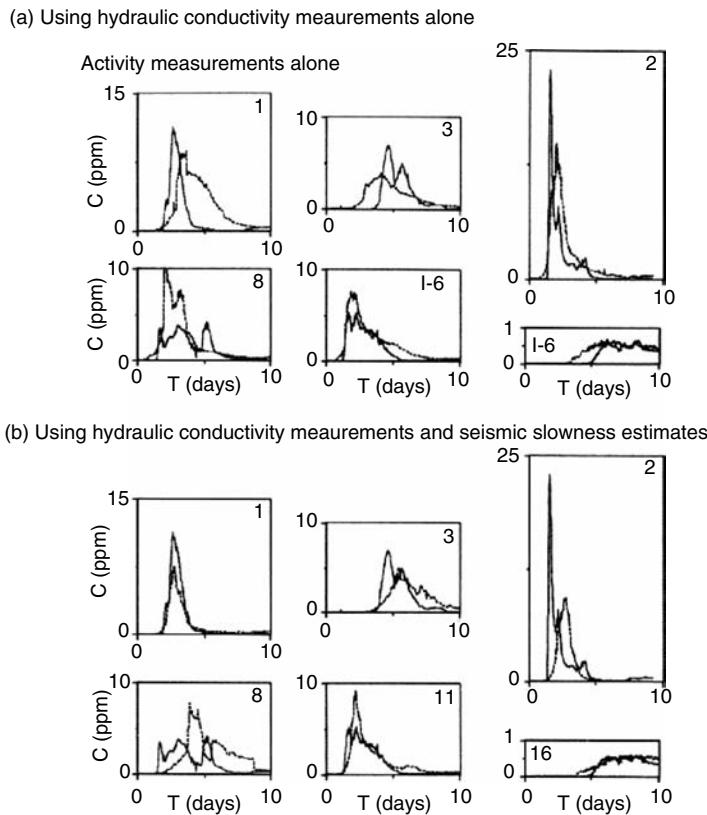


FIGURE 15.11 Comparison of observed (solid) and simulated (dashed) tracer concentration histories for the shallow Kesterson Aquifer, simulated through conductivity realizations (a) generated using conductivity measurements alone, and (b) generated using sequential co-simulation with hard conductivity data and a soft seismic slowness field based on estimates from seismic tomography (see Figure 15.10). (From Hyndman, D. W., Gorelick, S. M., and Harris, J. M. 2000. *Water Resour. Res.* 36(8):2121–2132. With permission.)

distribution based on seismic and lithologic data for a synthetic case study. They first estimated interval seismic velocities using *normal moveout corrections* (described in Section 15.3.1 of this chapter). Then they converted these interval velocities into a prior lithology indicator probability field, which represents the likelihood of each lithology across the survey region, using calibration curves that represented the relation between seismic velocity from well logs and lithology from core samples. This prior distribution was then conditioned on lithology logs from wells to provide an updated probability field. The log hydraulic conductivity pdf was then developed by weighting the conductivity probability distribution function (pdf) for each lithology by the indicator lithology pdf. The results of this study indicate that the best hydraulic conductivity estimates are obtained when both seismic and lithologic information are used. Hubbard et al. (1997) developed a similar *stochastic* technique that estimates binary permeability and saturation fields from GPR data based on petrophysical relationships from Knoll et al. (1995). In their numerical example, the GPR data dramatically improved the estimates relative to those based only on borehole data even with significant measurement error. In a field example at the Oyster, Virginia site, Hubbard et al. (2001) applied a Bayesian inversion approach to estimate hydraulic conductivity with prior estimates developed using borehole flowmeters, and posterior updates based on cross-well radar tomography. See 14-8 and Figure 14.18, along with the associated discussion in Chapter 14 of this volume for more information on this approach and example.

Other geostatistical techniques develop a covariance matrix between measurements and properties of interest through approximations of an appropriate governing equation. For example, hydraulic head can

be related to hydraulic conductivity using a first-order approximation of the steady-state groundwater flow equation (Hoeksema and Kitanidis, 1984), and solute tracer arrival times can be related to conductivity through a first-order approximation of the advection-dispersion equation (Harvey and Gorelick, 1995). The conductivity estimates obtained from such methods can be combined with high resolution seismic velocity estimates from inversion methods like those discussed in Section 15.3 of this chapter. For example, Rubin et al. (1992) and Carty et al. (1993) demonstrated the value of seismic velocity estimates given an assumed relation between seismic velocity and permeability (Marion et al., 1992). They showed that seismic information significantly enhanced estimates of hydraulic properties even in synthetic cases where random error was added to the seismic velocity values, when information about the relation between properties is available. For some sites, it may be possible to infer a relation between seismic and hydraulic properties using core data, which then allows the use of such stochastic methods to estimate the distribution of hydraulic properties across a region. Knowledge of the relationship between seismic or radar velocity and hydrogeologic properties (lithology or permeability), however, is generally unknown at the field scale.

Conditional indicator simulation can be used to combine a wide variety of datasets including geologic knowledge, seismic estimates, geophysical logs, and local lithologic and hydraulic conductivity measurements. This approach involves modeling the spatial correlation structure (i.e., variogram) for several sets of property ranges separated by indicator values, and developing realizations of a property that honor the correlation structure and are conditioned to all measured values. McKenna and Poeter (1995) classified the *hydrofacies* of a site using indicator seismic velocity thresholds. These thresholds were determined using discriminant analysis in a four-dimensional space representing borehole logs (sonic, neutron, electrical conductivity, and natural gamma). This approach allowed for a nonlinear and non-unique relation between hydraulic conductivity and seismic velocity. The concept behind this analysis is to group locations with similar lithologies as indicated by the combination of log responses at each location. Inverse flow modeling was then used to estimate the likely range of hydraulic conductivity for each hydrofacies and eliminate realizations that did not honor the hydraulic head data. This approach was used to delineate the hydrofacies of a fluvial deposit near Golden, Colorado, where the combined analysis of seismic, hydraulic, and geologic data reduced the uncertainty in the hydraulic conductivity estimates. Poeter and McKenna (1995) stressed the importance of incorporating hydraulic data to constrain the conductivity estimates. Using geologic and geophysical data (hydrofacies classification) from wells alone, they found that some of the geostatistical realizations were implausible.

15.6.2 Zonal Inversion Methods

Another approach to combining geophysical and hydrogeologic data involves inferring the large-scale zonation of lithologies and effective hydrogeologic properties that are consistent with seismic, tracer, and hydraulic data. The split inversion method (SIM) was introduced by Hyndman et al. (1994) to co-invert independently collected datasets (e.g., cross-well seismic traveltimes and tracer concentration histories) for the zonation of lithologies, effective hydraulic conductivity values for each zone, and an effective dispersivity for the region. The motivation for the approach is that the combined analysis of seismic energy and tracer concentrations, which have independently sampled portions of the same environment, should provide better estimates of the geometry of subsurface lithologies and effective zonal properties than obtained with either dataset alone. The SIM does not rely on knowledge of the relationship between seismic velocity and hydraulic conductivity, although it assumes some relationship exists for large-scale lithologic zones.

The ability to represent nonlinear and non-unique relations between these properties was illustrated by the accurate estimates obtained by Hyndman et al. (1994) for two synthetic aquifers with the same seismic profile (a heterogeneous high seismic velocity matrix with embedded low seismic velocity zones) but different hydraulic conductivity profiles. The low velocity zones in one aquifer represented low conductivity zones, while these zones in another aquifer represented both high and low conductivity zones. Thus seismic information alone did not provide enough information to discern the lithologic zonation or

hydrogeologic properties, yet the combined use of seismic and tracer data allowed for accurate estimates of both the zonation of lithologies and the hydrogeologic properties of the zones.

The SIM was also applied to the Kesterson Aquifer by Hyndman and Gorelick (1996). The first stage of this approach was to estimate the seismic slowness for the available cross sections between wells as discussed at the end of Section 15.5.1. The correlation structure of the estimated seismic tomograms was then modeled to provide horizontal and vertical slowness variograms for the region. These were used to develop multiple 3D conditional seismic slowness realizations (e.g., Figure 15.12) for a region surrounding the tomograms using sequential Gaussian simulation (Deutsch and Journel, 1992). The SIM was then used to split each slowness realization into three lithologic classes (e.g., Figure 15.13) and to estimate

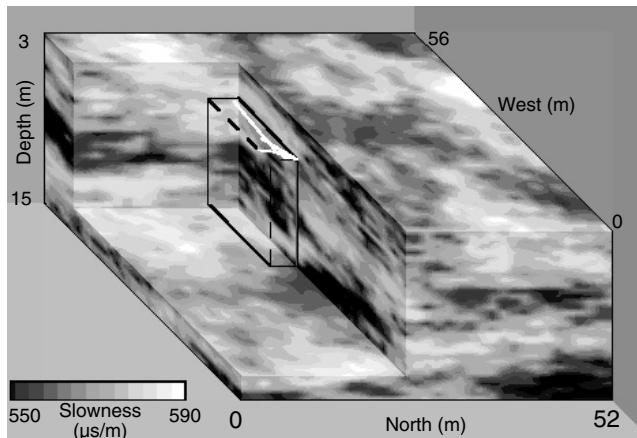


FIGURE 15.12 Three-dimensional seismic slowness realization for the shallow Kesterson Aquifer obtained using sequential simulation conditioned to the seismic tomograms in Figure 15.10. (From Hyndman, D. W. and Gorelick, S. M. 1996. *Water Resour. Res.* 32(9):2659–2670. With permission.)

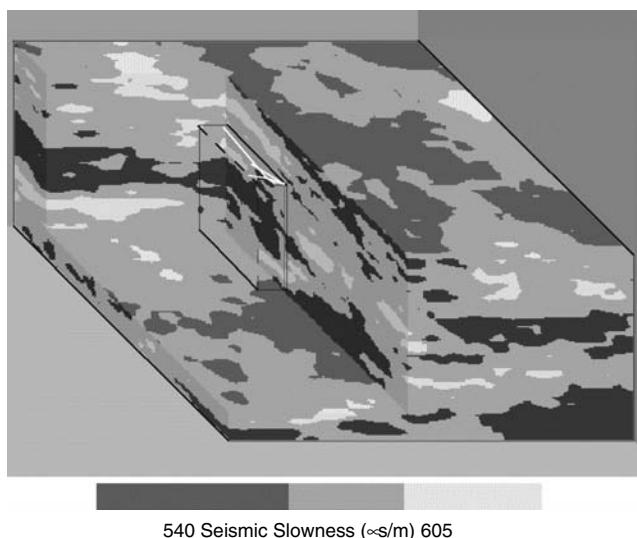


FIGURE 15.13 Zonation of hydraulic conductivity estimates obtained using the split inversion method (SIM). For this slowness realization (Figure 15.12), the estimated slowness splits are 569.4 and 587.8 s/m, and the hydraulic conductivity values are: 1.4×10^{-4} m/sec for the black regions; 3.6×10^{-4} m/sec for the gray regions; and 5.0×10^{-4} m/sec for the white regions. (From Hyndman, D. W. and Gorelick, S. M. 1996. *Water Resour. Res.* 32(9):2659–2670. With permission.)

both the effective hydraulic conductivity for each lithology and a single dispersivity value for the entire model domain. The objective function of this inversion was to minimize the squared residual between measured and simulated tracer arrival time quantiles, and drawdown during the tracer test. This study clearly illustrated the value of seismic tomography for estimating the hydraulic conductivity structure of a heterogeneous aquifer. The seismic slowness estimates provided information about the geometry of lithologic zones and the horizontal and vertical correlation structure of the aquifer lithologies. The drawdown measured during the tracer test provided information about the average regional hydraulic conductivity at the site, while the tracer concentrations provided information about the continuity of hydraulic conductivity across the region of the test.

15.6.3 Estimating the Relation between Geophysical and Hydrogeologic Properties

A major difficulty with incorporating geophysical data into hydrogeologic estimation methods is the uncertain and potentially non-unique relation between estimated geophysical properties (e.g., seismic velocity) and hydrogeologic properties (e.g., hydraulic conductivity). There is no reason to expect a fundamental relation between these properties as seismic velocities do not directly depend on the ability of sediment to transmit fluids. For particular frequencies of seismic energy, however, some researchers have found that fluid flow is important to the stiffness of the media, which is related to seismic velocity and attenuation (Bourbie et al., 1987). The effect of fluid moving through pores to accommodate the stress imposed by a sound wave was explored by Dvorkin et al. (1995) as an alternative to the Biot theory.

A primary goal in the field of rock physics (or petrophysics) is identifying relations between rock/sediment properties and measured geophysical signals to enable better predictions of properties in field settings. Petrophysics research has demonstrated empirical relations between seismic velocity and both porosity and clay content (Mavko et al., 2003). For example, Marion et al. (1992) developed a conceptual mixing model for the relationship between seismic velocity and porosity that involved sand-clay mixtures. The porosity had a minimum when the clay content was large enough to fill in the voids between sand grains, which resulted in a maximum seismic velocity. The porosity increased from this minimum toward both end member lithologies (sand and clay), with a corresponding reduction in the seismic velocity as the velocity of sound is lower through water than it is through sand or clay. Marion's model described one possible process leading to a relation between porosity and velocity. However, laboratory slowness estimates are often weakly correlated to the longer wavelength slowness estimates from the field because of factors such as frequency-dependent dispersion. It is thus difficult to extend the results of empirical laboratory studies to field scale with a wide variety of heterogeneous depositional environments.

Field-scale relationships between seismic slowness and hydraulic conductivity can be inferred by combining cross-well seismic tomography with hydraulic and tracer measurements in a joint estimation approach. The goal of this approach is to find a relation between seismic and hydraulic properties that minimizes the residual between simulated and measured hydraulic head and tracer concentration residuals. The concept is similar to that described above for the SIM, but for this approach the decision variables are parameters of the relation between seismic slowness and hydraulic conductivity. Hyndman et al. (2000) developed this approach to reduce computation time by reducing the number of decision variables adjusted in the estimation approach. The Kesterson field site was again used to test this approach, and the results were compared with the SIM results. For the Kesterson site a linear relation between slowness and the log of hydraulic conductivity appeared to be adequate based on: (1) laboratory measurements of hydraulic conductivity and seismic velocity, (2) the correlation between field cross-well slowness estimates and the hydraulic conductivity estimates in neighboring wells, which was used as an *a priori* relation for this iterative estimation approach, and (3) the zonal conductivity estimates from the split inversion method relative to the tomographic slowness values (see Figure 15.13).

A linear relation was not initially expected for the Kesterson Aquifer because of the nonlinear laboratory mixing relation of Marion et al. (1992). Even though the Marion model describes mixing of sand and clay, which are the predominant lithologies in the Kesterson region, the shallow aquifer examined in this study does not have the 100% clay end member, and the mixing process is likely quite different from

complete laboratory mixing. Estimating the relation between geophysical and hydrogeologic properties is a highly site- and scale-specific endeavor since no general relation is expected. The slowness (S in $\mu\text{s}/\text{m}$) at this site was positively correlated to log hydraulic conductivity (K in m/sec) for the shallow Kesterson Aquifer. The fit to the tracer and hydraulic data was significantly enhanced by the addition of the seismic data relative to those obtained using sequential simulation to interpolate between hydraulic conductivity measurements alone. This approach determined the likely conductivity structure of the shallow Kesterson Aquifer and explored the range of likely relations between slowness and conductivity at the field scale.

All the approaches described above indicated that geophysical information has the potential to increase the resolution of hydraulic property estimates. Changes in subsurface lithology are likely to cause a change in both geophysical and hydraulic properties at some scale, thus combining these different datasets provides better estimates than can be obtained with only one dataset. The primary difficulty in direct incorporation of these different datasets is the uncertain link between the measured properties. Different relations are expected for different depositional environments, and these relations may be nonlinear or non-unique if they are measurable at all. Moysey et al. (2005) presented an approach to develop field-scale petrophysical relationships using Monte Carlo simulation.

Other geophysical measurements could also be inverted to provide additional information about the hydraulic and lithologic properties of aquifers. For example, seismic attenuation is thought to be correlated to hydraulic conductivity. In all cases, high-quality data are necessary to perform detailed estimates of aquifer properties. This includes accurate surveys of well geometry for all cross-well surveys, since even small location errors may be larger than the variations in the property of interest between the wells. Hydraulic and tracer test data should be sampled with high spatial and temporal resolution to measure the fine scale variations in flow and transport that can be critical to solute transport predictions.

15.7 Dynamic Imaging of Tracer and Contaminant Plumes

A range of recent studies have used geophysical surveys to examine tracer or contaminant plumes. Static images of contaminant plumes provide valuable information which would be difficult or impossible to gather using well data alone. If a site has a conductive plume, as can occur from biodegradation of organic contaminants, the plume geometry can potentially be mapped using surface radar techniques (Sauck et al., 1998; Atekwana et al., 2000). Such plume maps could be used optimally design contaminant capture or remediation schemes.

Surface and cross-well radar have the potential to image the location of nonaqueous phase contaminant plumes (for nonpolar contaminants with permittivities very different from that of clean water), as was done in the Borden Aquifer (Greenhouse et al., 1993; Brewster and Annan, 1994). GPR reflections were observed as the dense nonaqueous phase liquid (DNAPL) pooled on successively deeper low-permeability lithologies. The late time GPR profiles illustrated the average increase in GPR velocity due to the introduction of the low dielectric constant DNAPL into the test region. The dynamic imaging of a tracer that influences the radar propagation velocity or attenuation can provide high resolution estimates of DNAPL saturation, which can be used to estimate the hydraulic conductivity structure of the aquifer and enhance our knowledge about transport processes in heterogeneous aquifer environments. Unfortunately, there has been limited success using geophysics to identify the location of NAPL plumes in the subsurface.

Dynamic geophysical imaging holds promise for estimating both changes in water content or concentrations of conductive or nonpolar solutes. Electrical resistivity tomography has been applied to image fractured rock (Slater et al., 1997), engineered barriers (Daily and Ramirez, 2000), and tracer transport through sediments (Mohrlock and Dietrich, 2001; Kemna et al., 2002; Singha and Gorelick, 2005). Such methods have defined the center of mass for tracer plumes, and provided information about solute plume geometry.

Cross-well radar tomography can also be used to dynamically image the transport of conductive tracer plumes. Attenuation difference tomography was used to successfully image the transport of a conductive

tracer plume through a highly heterogeneous test site in Boise Idaho (Barrash et al., 2003; Goldstein et al., 2003). This type of test provides the necessary data for fully coupled hydrogeophysical inversions, as the estimates of plume shape can be converted to estimates of solute velocity, which are closely related to the hydraulic properties of interest. Dynamic cross-well radar tomography was also used to image the properties of fractured rock at the U.S. Geological Survey Fractured Rock Hydrology Research Site in Grafton County, New Hampshire (Day-Lewis et al., 2002, 2003). Sequential imaging of the saline plume at this site allowed these researchers to identify preferential fracture flowpaths and estimate the peak tracer arrival times on planes between well pairs used for tomography. Such time-lapse geophysical imaging methods have the potential to provide insight into dynamic flow and transport processes through various coupled inversion methods.

Glossary

Aquifer Properties Characteristics of an aquifer that control the transport of fluid and solutes.

Attenuation Reduction in signal amplitude due to energy loss (because of reflection, dispersion, etc.).

Cokriging Interpolating between a series of measured values based on multiple types of data and the modeled spatial correlation structures of each data type.

Common Midpoint Gather A collection of geophysical data with source and receiver at different locations along a line, each equidistant from a central midpoint. If subsurface reflectors are flat and layered with homogeneous seismic properties in each layer, each trace sampled from the common midpoint gather will be reflected from the same point, which will be directly below the chosen common midpoint.

Converging Iteratively approaching a value or solution with successively smaller changes between iterations.

Global Minimum The smallest value of an objective function for all parameter combinations.

Hydrofacies Lithologies categorized by their flow and transport properties.

Inversion Iterative adjustment of model parameters to minimize the difference between simulated and observed values (e.g., hydraulic head or tracer concentrations).

Non-uniqueness For optimization problems, this condition indicates that multiple sets of parameters provide similar minimum objective values, thus no global minimum exists and the optimal parameter set cannot be determined using the provided data alone.

Normal Moveout Correction (NMO) Adjusting seismic or radar velocity estimates to flatten the parabolic appearance of reflectors due to offset between sources and receivers.

Objective Function A mathematical function that is minimized or maximized (i.e., the residual between measured and simulated head and concentration values for parameter estimation, or total cost for managing aquifer pumping).

Offset Horizontal distance between a source and a receiver.

Optimization Finding parameters that achieve the optimal objective value (minimum or maximum).

Pump Tests or Slug Tests Methods of estimating hydraulic conductivity near a well by pumping water into or out of the well and measuring the hydraulic response. A slug test involves rapid removal or addition of a small volume of water (or a solid slug that displaces water), while a pump test involves continuous pumping of water for a period of time.

Reflection Energy that bounces back from a surface due to a change in physical properties, such as seismic impedance in the case of sound waves.

Refraction Energy that bends according to changes in the velocity of sound waves in adjacent media. The angle can be described according to Snell's law, which is included in the text of this chapter.

Regression Iteratively adjusting the parameters of a model to minimize (or in some cases maximize) the objective value.

Simulation-Optimization Methods These methods estimate optimal parameter values by applying an optimization package, such as one based on nonlinear least-squares regression, to physically

based models of groundwater flow and solute transport to obtain the most desirable system response.

Slowness 1/seismic velocity. This terminology is used because the traveltime equation is solved using slowness.

Solute A dissolved substance (e.g., a contaminant or a tracer in groundwater).

Spatial Moments Statistical representations for the location, spread, and shape of a solute plume. In normalized forms, the zeroth moment defines the total mass, the first moment defines the location of the center of mass, and the second moment defines the spread of the plume.

Stacking Averaging multiple radar or seismic traces collected from the same location to improve the signal-to-noise ratio.

Travelttime Tomography Estimating the seismic velocity structure between wells based on the time for a type of wave (e.g., a compressional- or P-wave) to travel from the source to the receiver.

References

- Adams, E. E. and Gelhar, L. 1992. Field study of dispersion in a heterogeneous aquifer, 2. Spatial moments analysis. *Water Resour. Res.* 28(12): 3293–3307.
- Alumbaugh, D., Chang, P. Y., Paprocki, L., Brainard, J. R., Glass, R. J., Rautman, C. A., 2002. Estimating moisture contents in the vadose zone using cross-borehole ground penetrating radar: a study of accuracy and repeatability. *Water Resour. Res.* 38(12): 1309.
- Araktingi, U. G. and Bashore, W. M. 1992. Effects of properties in seismic data on reservoir characterization and consequent fluid-flow predictions when integrated with well logs, *Society of Petroleum Engineers* (SPE 24752, Presented at 67th Annual SPE conference in Washington, DC).
- Arcone, S. A., Lawson, D. E., and Delaney, A. J. 1995. Short-pulse radar wavelet recovery and resolution of dielectric contrasts within englacial and basal ice of Matanuska Glacier, Alaska, USA. *J. Glaciology* 41(137): 68–86.
- Aster, R., Borchers, B., and Thurber, C. 2005, *Parameter Estimation and Inverse Problems*, Elsevier Press, 390 p.
- Atekwana, E. A., Sauck W. A., and Werkema, D. 2000, Investigations of geoelectrical signatures at a hydrocarbon contaminated site; *J. Appl. Geophys.*, 44(2–3): 167–180.
- Barlebo, H. C., Hill, M. C., and Rosbjerg, D. 2004. Investigating the Macrodispersion Experiment (MADE) site in Columbus, Mississippi, using a three-dimensional inverse flow and transport model. *Water Resour. Res.* 40, W04211, doi:10.1029/2002WR001935.
- Barrash, W., Clemo, T., Fox, J. J., and Johnson, T. C. (2006). Field, laboratory, and modeling investigation of the skin effect at wells with slotted casing, Boise hydrogeophysical research site, *J. Hydrol.*
- Barrash, W., Knoll, M., Hyndman, D., Clemo, T., Reboulet, E. and Haurath, E. 2003. Tracer/time-lapse radar imaging test at the Boise Hydrogeophysical Research Site, in *Proc. SAGEEP 2003 (San Antonio, TX)*, pp. 163–174, Environmental and Engineering Geophysical Society, Denver, CO.
- Benson, S. M. 1988. Characterization of the hydrogeologic and transport properties of the shallow aquifer under Kesterson Reservoir. Ph.D. thesis, University of California Berkeley.
- Benson, S. M., White, A. F., Halfman, S., Flexser, S., and Alavi, M. 1991. Groundwater contamination at the Kesterson Reservoir, California: 1. Hydrogeologic setting and conservative tracer transport. *Water Resour. Res.* 27(6): 1071–1084.
- Beres, M., and Haeni, F. P. 1991. Application of ground penetrating radar methods in hydrologic studies. *Ground Water* 29(3): 375–386.
- Binley, A., Winship, P., West, J. L., Pokar, M., and Middleton, R. 2001. High-resolution characterization of vadose zone dynamics using cross-borehole radar. *Water Resour. Res.* 37: 2639–2652.
- Biteman, S. E., Hyndman, D. W., Phanikumar, M. S., and Weissmann, G. S. 2004. Integration of sedimentologic and hydrogeologic properties for improved transport simulations, In *Aquifer Characterization*, J. S. Bridge and D. W. Hyndman, eds. SEPM Special Publication 80.

- Boggs, M. J., Yound, S. C., Beard, L. M., Gelhar, L. W., Rehfeldt, K. R., and Adams, E. E. 1992. Field study of dispersion in a heterogeneous aquifer: 1. Overview and site description. *Water Resour. Res.* 28(12): 3281–3291.
- Bourbie, T., Coussy, O., and Zinszner, B. 1987. *Acoustics of Porous Media*. Gulf Publishing Company, Houston, TX.
- Brewster, M. L. and Annan, A. P. Ground-penetrating radar monitoring of a controlled DNAPL release: 200 MHz radar, *Geophysics* 59(8): 1211–1221.
- Claerbout, J. 1992. *Earth Sounding Analysis: Processing Versus Inversion*. Blackwell Scientific, Boston, MA, 304 p.
- Cpty, N. and Rubin, Y. 1995. A stochastic approach to the characterization of lithofacies from surface seismic and well data. *Water Resour. Res.* 31(7): 1673–1686.
- Cpty, N., Rubin, Y., and Mavko, G. 1993. Geophysical-hydrological identification of field permeabilities through Bayesian updating. *Water Resour. Res.* 29(8): 2813–2825.
- Daily, W. D. and Ramirez, A. L. 2000 Electrical imaging of engineered hydraulic barriers. *Geophysics* 65(1): 83–94.
- Day-Lewis, F. D., Harris, J. M., and Gorelick, S. M. 2002. Time-lapse inversion of crosswell radar data. *Geophysics* 67(6): 1740–1752.
- Day-Lewis, F. D., Lane, Jr. J. W., Harris, J. M. and S. M. Gorelick, 2003. Time-lapse imaging of saline-tracer transport in fractured rock using difference-attenuation radar tomography. *Water Resour. Res.*, 39(10): 1290,10.1029/2002WR001722.
- Deutsch, C. V. and Journel, A. G. 1997. *Geostatistical Software Library and User's Guide*. 2nd Edition, Oxford University Press, Oxford.
- Dvorkin J., Mavco, G., and Nur, A. 1995. Squirt flow in fully saturated rocks. *Geophysics* 60(1): 97–107.
- Dybas, M.J., Hyndman, D. W., Heine, R., Tiedje, J., Lanning, K., Wiggert, D., Voice, T., Zhao, X., Dybas, L., and Criddle, C. S. 2002. Development, operation, and long-term performance of a full-scale biocurtain utilizing bioaugmentation. *Environ. Sci. Technol.* 16(36): 3635–3644.
- Fisher, E., McMechan, G., and Annan, A. P. 1992. Acquisition and processing of wide-aperture ground penetrating radar data. *Geophysics* 57: 495–504.
- Freyberg, D. L. 1986. A natural gradient experiment on solute transport in a sand aquifer 2. Spatial moments and the advection and dispersion of nonreactive tracers. *Water Resour. Res.* 22(13): 2031–2046.
- Garabedian, S. P., LeBlanc, D. R., Gelhar, L. W., and Celia, M. A. 1991. Large-scale natural gradient tracer test in sand and gravel, Cape Cod, Massachusetts. 2. Analysis of spatial moments for a nonreactive tracer. *Water Resour. Res.* 27(5): 911–924.
- Gelhar, L. W. 1993. *Stochastic Subsurface Hydrology*. Prentice Hall, Englewood Cliffs, NJ, 390 p.
- Gelhar, L. W. and Axness, C. L. 1983. Three dimensional stochastic analysis on macrodispersion in aquifers. *Water Resour. Res.* 19(1): 161–180.
- Goldstein, S., Knoll, M., Barrash, W., and Clement, W. 2003. Borehole radar attenuation tomography during Tracer/Time-Lapse Test at the Boise Hydrogeophysical Research site, in Symposium on the Application of Geophysical Engineering and Environmental problems, San Antonie, TX., 2003.
- Greaves, R. J., Lesmes, D. P., Le, J. M., and Toksoz, M. N. 1996. Velocity variations and water content estimated from multi-offset, ground penetrating radar. *Geophysics* 61(3): 683–695.
- Greenhouse, J., Brewster, M., Schneider, G., Redman, G., Annan, P., Olhoeft, G., Lucius, J., Sander, K., and Mazzella, A. 1993. Geophysics and solvents: the Borden experiment. *The Leading Edge* 12(4): 261–267.
- Han, D., Nur, A., and Morgan, D. 1986. Effect of porosity and clay content on wave velocity in sandstones. *Geophysics* 51(11): 2093–2107.
- Harris, J. M., Nolen-Hoeksema, R. C., Langan, R. T., Schaack, M. V., Lazaratos, S. K., and Rector III, J. W. 1996. High resolution cross-well imaging of a west Texas carbonate reservoir: Part 1 — Project summary and interpretation. *Geophysics* 60(3): 667–681.

- Harris, J. M. and Wang, G. Y. 1996. Diffraction tomography for inhomogeneities in layered background medium. *Geophysics* 61(2): 570–583.
- Harvey, C. F. and Gorelick, S. M. 1995. Mapping hydraulic conductivity: sequential conditioning with measurements of solute arrival time, hydraulic head, and local conductivity. *Water Resour. Res.* 31(7): 1615–1626.
- Hess, K. M., Wolf, S. H., and Celia, M. A. 1992. Large-scale natural gradient tracer test in sand and gravel, Cape Cod, Massachusetts. 3. Hydraulic conductivity variability and calculated macrodispersivities. *Water Resour. Res.* 28(8): 2011–2027.
- Hoeksema, R. J. and Kitanidis, P. K. 1984. An application of the geostatistical approach to the inverse problem in two-dimensional groundwater modeling. *Water Resour. Res.* 20(7): 1003–1020.
- Hubbard, S. S., Chen, J., Peterson, J., Swift, D.J., Mailloux, B., and Rubin, Y. 2001. Hydrogeological Characterization of the South Oyster bacterial transport site using geophysical data. *Water Resour. Res.* 37(10): 2431–2456.
- Hubbard, S. S., Rubin, Y., and Majer, E. 1997. Ground-penetrating-radar-assisted saturation and permeability estimation in bimodal systems. *Water Resour. Res.* 33(5): 971–990.
- Hyndman, D. W., Dybas, M. J., Forney, L., Heine, R., Mayotte, T., Phanikumar, M.S., Tatara, G., Tiedje, J., Voice, T., Wallace, R., Wiggert, D., Zhao, X. and Criddle, C. S. 2000, Hydraulic characterization and design of a full-scale biocurtain. *Ground water* 38(3): 462–474.
- Hyndman, D. W. and Gorelick, S. M. 1996. Mapping lithologic and transport properties in three dimensions using seismic and tracer data: the Kesterson aquifer. *Water Resour. Res.* 32(9): 2659–2670.
- Hyndman, D. W. and Harris, J. M. 1996. Traveltime inversion for the geometry of aquifer lithologies. *Geophysics* 61(6): 1728–1737.
- Hyndman, D. W., Harris, J. M., and Gorelick, S. M. 1994. Coupled seismic and tracer test inversion for aquifer property characterization. *Water Resour. Res.* 30(7): 1965–1977.
- Hyndman, D. W., Gorelick, S. M., and Harris, J. M. 2000 Inferring the relationship between seismic slowness and hydraulic conductivity in heterogeneous aquifers. *Water Resour. Res.* 36(8): 2121–2132.
- Isaaks, E. H. and Srivastava, R. M. 1989. *An Introduction to Applied Geostatistics*. Oxford Press, New York, 561 p.
- Johnson, T.C., P.S. Routh and M.D. Knoll. 2005. Fresnel volume georadar attenuation-difference tomography, *Geophysical Jour. International* 162, 9–24.
- Kemna, A., Vanderborght, J., Kulessa, B., and Vereecken, H. 2002. Imaging and characterization of subsurface solute transport using electrical resistivity tomography (ERT) and equivalent transport models. *J. Hydrology* 267(3–4): 125–146.
- Kearey, P. and Brooks, M. 1991. *An Introduction to Geophysical Exploration*, 2nd Edition, Blackwell Scientific Publications, Oxford.
- Kim, C., Daniels, J. J., Guy, E. D., Radzevicius, S. J., Holt, J. 2000. Residual hydrocarbons in a water-saturated medium: a detection strategy using ground penetrating radar. *Environ. Geosci.* 7(4): 169.
- Knoll, M. D., Haeni, F. P., and Knight, R. J. 1991. Characterization of a sand and gravel aquifer using ground-penetrating radar, Cape Cod, Massachusetts. G. E. Mallard and D. E. Aronson, eds. U. S. Geological Survey toxic substances hydrology program — Proceedings of the technical meeting, March 11–15, 1991, Monterey, CA.
- Knoll, M., Knight, R., and Brown, E. 1995. Can accurate estimates of permeability be obtained from measurements of dielectric properties? *SAGEEP Annual Meeting Extended Abstracts*, Orlando, FL.
- Lazaratos, S. K., Rector III, J. W., Harris, J. M., and Van Schaack, M., 1993 High-resolution, cross-well reflection imaging: potential and technical difficulties. *Geophysics* 58(9): 1270–1280.
- LeBlanc, D. R., Garabedian, S. P., Hess, K. M., Gelhar, L. W., Quadri, R. D., Stollenwerk, K. G., and Wood, W. W. 1991. Large-scale natural gradient tracer test in sand and gravel, Cape Cod, Massachusetts. 1. Experimental design and observed tracer movement. *Water Resour. Res.* 27(5): 895–910.

- Lunt, I., S. Hubbard and Y. Rubin, 2005. Soil moisture content estimation using ground-penetrating radar reflection data. *Journal of Hydrology* 307, 1–4, 254–269.
- Marion, D., Nur, A., Yin, H., and Han, D. 1992. Compressional velocity and porosity in sand–clay mixtures. *Geophysics* 57(4): 554–563.
- Mavko, G., Mukerji, T., and Dvorkin, J. 2003. *The Rock Physics Handbook: Tools for Seismic Analysis in Porous Media*. 2nd ed. Cambridge University Press, New York.
- McKenna, S. A. and Poeter, E. P. 1995. Field example of data fusion in site characterization. *Water Resour. Res.* 31(12): 3229–3240.
- McLaughlin, D. and Townley, L. R. 1996. A reassessment of the groundwater inverse problem. *Water Resour. Res.* 32(5): 1131–1161.
- McMechan, G. A., Harris, J. M., and Anderson, L. M., 1987. Cross-hole tomography for strongly variable media with applications to scale model data. *Bul. Seis. Soc. Am.* 77(6): 1945–1960.
- Michelena, R. and Harris, J. 1991. Tomographic traveltimes inversion using natural pixels. *Geophysics* 56(5): 635–644.
- Mohrlock, U. and Dietrich. 2001. Exploration of preferential transport paths using geoelectrical salt tracer tests, Field Screening Europe, 4 pp.
- Molz, F. 1994. Borehole flowmeters: field application and data analysis. *J. Hydrology* 163(3–4): 347–371.
- Moysey, S., Singha, K., and Knight, R. 2005. A framework for inferring field-scale rock physics relationships through numerical simulation. *Geophys. Res. Lett.* 32, L08304, doi:10.1029/2004GL022152.
- Olhoeft, G. R. 1988. Interpretation of hole-to-hole radar measurements. *Proceedings of the Third Technical Symposium on Tunnel Detection — Golden, Colorado*. pp. 616–626. (also in USGS Open File Report 900414).
- Phanikumar, M. S., Hyndman, D. W., Zhao, X., Dybas, M. J. 2005. A three-dimensional model of microbial transport and biodegradation at the Schoolcraft, Michigan, site. *Water Resour. Res.* 41, W05011, doi:10.1029/2004WR003376.
- Placzek, G. and Haeni, F. P. 1995. Surface-geophysical techniques used to detect existing and infilled scour holes near bridge piers. U.S. Geological Survey Water Resources Investigations Report 954009, Denver, CO.
- Poeter, E. P. and McKenna, S. A. 1995. Reducing uncertainty associated with ground-water flow and transport predictions. *Ground Water* 33(6): 899–904.
- Quan, Y. and Harris, J. 1997. Seismic attenuation tomography using the frequency shift method. *Geophysics* 62(3): 895–905.
- Rea, J. and Knight, R. 1998. Geostatistical analysis of ground penetrating radar data: a means of describing spatial variations in the subsurface. *Water Resour. Res.* 34(3): 329–340
- Rehfeldt, K. R., Boggs, M. J., and Gelhar, L. 1992. Field study of dispersion in a heterogeneous aquifer, 3. Geostatistical analysis of hydraulic conductivity. *Water Resour. Res.* 28(12): 3281–3291.
- Reiter, D. T. and Rodi, W. 1996. Nonlinear waveform tomography applied to crosshole seismic data. *Geophysics* 61(3): 902–913.
- Rubin, Y., Mavko, G., and Harris, J. 1992. Mapping permeability in heterogeneous aquifers using hydrologic and seismic data. *Water Resour. Res.* 28(7): 1809–1816.
- Sanchez-Villa, X., Carrera, J., and Colominas, I. 1992. Directional effects on convergent flow tracer tests, In *Tracer Hydrology*, H. Hotzl and A. Werner, eds. A. A. Balkema, Rotterdam, pp. 407–414.
- Sauck, W. A., Atekwana, E. A., and Nash, M. S. 1998. High conductivities associated with an LNAPL plume imaged by integrated geophysical techniques. *J. Environ. Eng. Geophys.* 2: 203–212.
- Singha, K., and Gorelick, S. M. 2005. Saline tracer visualized with three-dimensional electrical resistivity tomography: field-scale spatial moment analysis. *Water Resour. Res.* 41(5): W05023, doi 10.1029/2004WR003460, 17 pp.
- Slater, L., Binley, A., and Brown, D. 1997. Electrical imaging of fractures using ground-water salinity change. *Ground Water* 35, 436–442.

- Steebles, D. W. and Miller, R. D. 1990. Seismic reflection methods applied to engineering, environmental, and groundwater problems, In *Geotechnical and Environmental Geophysics*, Ward, S. ed. Vol. I, Investigations in Geophysics, SEG Series. SEG, Tulsa, OK.
- Sudicky, E. A. 1986. A natural gradient experiment on solute transport in a sand aquifer: spatial variability of hydraulic conductivity and its role in the dispersion process. *Water Resour. Res.* 22(13): 2069–2082.
- Telford, W. M., Geldart, L. P., and Sheriff, R. E. 1990. *Applied Geophysics*. 2nd Edition. Cambridge University Press, Cambridge.
- Tiedeman, C. R. and Hsieh, P. A. 2004. Evaluation of longitudinal dispersivity estimates from simulated forced- and natural-gradient tracer tests in heterogeneous aquifers. *Water Resour. Res.* 40, W01512, doi:10.1029/2003WR002401.
- Van Dam, R. L., Van Den Berg, E. H., Schaap, M. G., Broekema, L. H., and Schlager, W. 2003. Radar reflections from sedimentary structures in the vadose zone, In *Ground Penetrating Radar in Sediments: Applications and Interpretation*, Bristow, C. S. and Jol, H. M. eds., Geological Society Special Publication No. 211, pp. 257–273.
- Vogelsang, D. 1995. *Environmental Geophysics: A Practical Guide*. Springer-Verlag, Berlin, 173 p.
- Ward, S. (ed). 1990. *Geotechnical and Environmental Geophysics*, Vols. I, II, and III. Investigations in Geophysics, SEG Series. SEG, Tulsa, OK.
- Zhao, X., Wallace, R. B., Hyndman, D. W., Dybas, M. J. and Voice, T. C. 2005. Heterogeneity of chlorinated hydrocarbon sorption properties in a sandy aquifer. *J. Contam. Hydro.* 78(4): 327–342.

Further Information

Aster et al. (2005) provide a detailed overview of parameter estimation methods. Telford et al. (1990) provide an excellent overview of introductory geophysical methods, but have a limited discussion of tomography or GPR. Ward (1990) compiled a large set of papers relating to environmental geophysics, including case studies and theoretical investigations. McLaughlin and Townley (1996) provide an excellent review of inverse methods for estimating groundwater flow and solute transport properties.

16

Geostatistics: Interpolation and Inverse Problems

16.1	Introduction.....	16-1
16.2	Preliminaries	16-2
	Model • Stochastic Approach • Least Squares	
16.3	Ordinary Kriging	16-3
	Interpolation • Function Estimate • Conditional Realizations	
16.4	Variogram	16-6
	Variogram Model • Variogram Selection • Experimental Variogram • Residuals • Variogram Testing • Variogram Calibration • Nonnegativity	
16.5	Kriging Variants.....	16-11
16.6	Cokriging.....	16-13
16.7	Generalized Linear Estimation Equations	16-13
	Model • Blue • Parameter Estimation	
16.8	Inverse Problems.....	16-14
	Definition • Methodology • Derivative Computation	
	Glossary.....	16-17
	References	16-17
	Further Information	16-19

Peter K. Kitanidis
Stanford University

16.1 Introduction

Hydrogeologic applications often require the estimation of functions from data. For example:

Interpolation From observations of water elevation at a number of wells, draw the contour map of the water table in an aquifer.

Inverse Problems Estimate transmissivity from head observations and other information.

Most applications involve quantities that vary in space but the methods may also apply to quantities that vary in time. A variety of methods have been developed to address such problems in hydrogeology but also elsewhere, interpolation and inverse problems being ubiquitous. Methods include splines (de Boor, 1978), objective analysis (Thiebaux and Pedder, 1987), and the like. There are considerable formal similarities among available methods (Cressie, 1993), but even formally equivalent methods may differ in

emphasis or in how they are applied and understood. The focus here is on *geostatistics* (Matheron, 1971), a practical methodology popular in hydrology and other geophysical sciences.

We will put in a nutshell results useful in solving interpolation and inverse problems. The presentation assumes familiarity with linear algebra and probability theory.

16.2 Preliminaries

16.2.1 Model

A model of variability is essential for the solution of interpolation problems. The unknown function, denoted as z , is a function of variable \mathbf{x} (space and time coordinates). A practical and versatile model is:

$$z(\mathbf{x}) = c(\mathbf{x}) + \underbrace{\sum_{k=1}^p f_k(\mathbf{x})\beta_k}_{\text{deterministic}} + \underbrace{\varepsilon(\mathbf{x})}_{\text{random}} \quad (16.1)$$

where $c(\mathbf{x})$ and $f_1(\mathbf{x}), \dots, f_p(\mathbf{x})$ are known functions. The bold letter \mathbf{x} represents an array of point coordinates. The coefficients β_1, \dots, β_p are unknown. The term $\varepsilon(\mathbf{x})$ is a *random function* with zero expected value.

The *deterministic part*, known also as the *mean*, or *drift*, or the *trend*, represents information about large-scale trends, zones, and other features that can be represented through deterministic functions. This part may include parameters β to be “fitted” from the data. However, a crucial requirement is that the number of β coefficients must be much smaller than the number of observations. The *random part* represents erratic, complex, and usually small-scale variability about which we know too little to model through deterministic functions but enough to describe through statistics (or probabilistic averages). In addition to stating that the expected value of ε is zero, we specify a covariance function, variogram, or generalized covariance function. These statistics furnish information about the variance and the degree of continuity of ε and are needed to obtain estimates with error bars.

16.2.2 Stochastic Approach

The objective is to find a function z , a hydrogeologic variable, given limited information. If nothing were known about z , the search should be over an ensemble that includes all functions. As more information becomes available, many of these functions are eliminated or become less likely. The stochastic approach involves assigning probabilities to these functions that form the ensemble.

Geostatistical estimation is for convenience broken into two phases:

Structural analysis. The search is narrowed down to functions that share certain characteristics, collectively known as *structure*, such as continuity and smoothness, scale of variability, etc. Information about the structure is represented here through equations for the deterministic part and the covariance function of the stochastic part (Equation 16.1).

Conditioning. The search is further narrowed by eliminating functions that are not consistent with observations.

16.2.3 Least Squares

Consider the interpolation problem: estimate the value of z at a location \mathbf{x}_0 from observations at locations $\mathbf{x}_1, \dots, \mathbf{x}_n$. One approach is to fit parameters β_1, \dots, β_p , and $z(\mathbf{x}_0)$ from data $z(\mathbf{x}_1), \dots, z(\mathbf{x}_n)$ using Equation 16.1 and a weighted least squares criterion. Using this intuitive and time-honored approach, one obtains an estimate and its mean square error of estimation. It turns out that formally this is equivalent to the methods of linear geostatistics. However, in implementation, there are considerable

differences between some trend analysis approaches (traditionally associated with least squares fitting) and geostatistics, including:

- Trend analysis attempts to capture the structure of the function through an elaborate deterministic part, whereas the random part is represented as lacking spatial correlation. However, the deterministic part is not well suited to represent information about small-scale and erratic variability. In geostatistics, such variability may be represented through the structure of the random part that is appropriately parameterized through a covariance function.
- In geostatistics, a central issue is the selection of the covariance function or the variogram, which is accomplished in the phase that is known as structural analysis.
- In geostatistics, because one is not really interested in estimating the β coefficients, one may use simplified versions of the covariance function, which is what variograms and generalized covariance functions are. One can also use algorithms that are well suited for interpolation.

16.3 Ordinary Kriging

16.3.1 Interpolation

In many applications, it is appropriate to adopt the simplest and most conservative representation of the mean, which is as an unknown constant, that is, referring to Equation 16.1,

$$z(\mathbf{x}) = \beta + \varepsilon(\mathbf{x}) \quad (16.2)$$

The information about the function structure is described only through the *variogram* (used instead of the traditional term *semivariogram*) defined as:

$$\gamma(\mathbf{x} - \mathbf{x}') = \frac{1}{2} E[(z(\mathbf{x}) - z(\mathbf{x}'))^2] \quad (16.3)$$

where E denotes a probabilistic average or expected value. *Ordinary kriging*, the most common geostatistical method, gives an estimate and its mean square error of estimation of z at any location \mathbf{x}_0 given the variogram γ and n observations $z(\mathbf{x}_1), z(\mathbf{x}_2), \dots, z(\mathbf{x}_n)$. To infer the unmeasured value of z at a location \mathbf{x}_0 , a *linear estimator* is adopted, that is, formula:

$$\hat{z}_0 = \sum_{i=1}^n \lambda_i z(\mathbf{x}_i) \quad (16.4)$$

The weights $\lambda_1, \dots, \lambda_n$ are chosen based on the requirements of (1) *unbiasedness*, that is, the expected value of the estimation error $\hat{z}_0 - z(\mathbf{x}_0)$ must be zero and (2) *best* or minimum variance estimation, that is, the expected value of the square error must be as small as possible. These requirements result in a system of $n + 1$ linear equations with $n + 1$ unknowns, known as the kriging system:

$$\begin{aligned} - \sum_{j=1}^n \lambda_j \gamma(\mathbf{x}_i - \mathbf{x}_j) + v &= -\gamma(\mathbf{x}_i - \mathbf{x}_0), \quad i = 1, 2, \dots, n \\ \sum_{j=1}^n \lambda_j &= 1 \end{aligned} \quad (16.5)$$

where ν is a Lagrange multiplier. The mean square error of estimation is given by the formula:

$$\sigma_0^2 = E[(\hat{z}_0 - z(\mathbf{x}_0))^2] = -\nu + \sum_{j=1}^n \lambda_j \gamma(\mathbf{x}_j - \mathbf{x}_0) \quad (16.6)$$

Often, the reliability of the estimate is presented in the form of a 95% *confidence interval*. This is the interval that contains the actual value of $z(\mathbf{x}_0)$ with probability 0.95. Under normality, that is, Gaussian distribution of the estimation errors (a common implicit assumption in linear estimation), the confidence interval is approximately $[\hat{z}_0 - 2\sigma_0, \hat{z}_0 + 2\sigma_0]$.

In the *global neighborhood* approach, all observations are used. In the *local neighborhood* approach, one uses a subset of all observations; one may use a fixed number of observations that are closest to the point \mathbf{x}_0 , all observations within a certain radius, or a combination of these strategies. The global neighborhood approach is recommended except in special cases such as when there is a large number of observations. If the local neighborhood approach is used for interpolation, care should be taken to avoid introducing inadvertent discontinuities in the estimate.

16.3.2 Function Estimate

An expression for the estimate at any point \mathbf{x} that is useful in interpolation and graphing a function is the following:

$$\hat{z}(\mathbf{x}) = - \sum_{i=1}^n \gamma(\mathbf{x} - \mathbf{x}_i) \xi_i + \hat{\beta} \quad (16.7)$$

where the ξ coefficients and $\hat{\beta}$ are found from the solution of a single system of $n+1$ equations with $n+1$ unknowns:

$$\begin{aligned} - \sum_{j=1}^n \gamma(\mathbf{x}_i - \mathbf{x}_j) \xi_j + \hat{\beta} &= z(\mathbf{x}_i), \quad \text{for } i = 1, \dots, n \\ \sum_{j=1}^n \xi_j &= 0 \end{aligned} \quad (16.8)$$

This approach is useful when performing kriging on a fine grid, and expression 16.7 reveals how the form of the estimate depends on the form of the variogram.

The conditional covariance function R_c (i.e., the covariance function given the data) is not stationary and may be obtained explicitly as follows:

1. Form matrix of coefficients of system 16.8

$$\begin{bmatrix} Q & X \\ X^T & 0 \end{bmatrix} \quad (16.9)$$

where Q is matrix n by n with $Q_{ij} = -\gamma(\mathbf{x}_i - \mathbf{x}_j)$, X is n by 1 with $X_i = 1$, and exponent T denotes matrix transpose.

2. Compute matrix inverse

$$\begin{bmatrix} Q & X \\ X^T & 0 \end{bmatrix}^{-1} = \begin{bmatrix} P & a \\ a^T & B \end{bmatrix}$$

where P is n by n , a is n by 1, and B is a scalar.

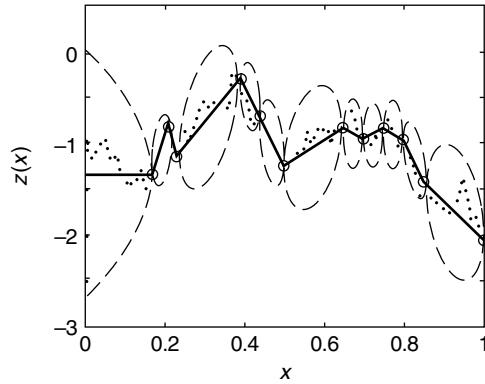


FIGURE 16.1 Best estimate (solid line), 95% confidence interval (dashed lines), vs. actual (fine dotted line) z function.

3. The conditional covariance function is given from:

$$R_c(\mathbf{x}, \mathbf{x}') = -\gamma(\mathbf{x} - \mathbf{x}') - \sum_{i=1}^n \sum_{j=1}^n \gamma(\mathbf{x} - \mathbf{x}_i) P_{ij} \gamma(\mathbf{x}_j - \mathbf{x}') - B + \sum_{j=1}^n a_j \gamma(\mathbf{x}_j - \mathbf{x}') + \sum_{i=1}^n \gamma(\mathbf{x} - \mathbf{x}_i) a_i \quad (16.10)$$

The mean square error is given from this formula by setting $\mathbf{x}' = \mathbf{x}$.

For illustration, the true function $z(\mathbf{x})$, the best estimate $\hat{z}(\mathbf{x})$, and the 95% confidence interval are shown in Figure 16.1, for a hypothetical case of a function of one variable. The data are indicated by small open circles and the variogram $\gamma(h) = 1.4$ h was used for estimation. The best estimate reproduces the observations and is a simplified version of the actual function. The 95% confidence interval is indicative of the potential error (difference between estimate and actual function).

16.3.3 Conditional Realizations

It is often useful to generate a large number of equally likely (given the available information) possible solutions perform probabilistic risk analysis. These solutions are called *conditional realizations* or conditional sample functions or conditional simulations. The average of the ensemble of all conditional realizations is equal to the best estimate and the covariance is equal to the conditional covariance function. We focus on the prevalent case of Gaussian realizations.

16.3.3.1 Point Simulation

Step 1. Generate an unconditional realization $z(\mathbf{x}; k)$, with zero mean, consistent with the assumed variogram (where the index k is used to remind that there are many possible solutions). This is a technical issue (Mantoglou and Wilson, 1982; Tompson et al., 1989; Dykaar and Kitanidis, 1992; Gutjahr et al., 1994; Robin et al., 1993) that goes beyond the scope of this chapter.

Step 2. Generate a conditional realization $z_c(\mathbf{x}_0; k)$ at a given point \mathbf{x}_0 ,

$$z_c(\mathbf{x}_0; k) = z(\mathbf{x}_0; k) + \sum_{j=1}^n \lambda_j (z(\mathbf{x}_j) - z(\mathbf{x}_j; k)) \quad (16.11)$$

where $z(\mathbf{x}_i)$ is actual observation and the λ coefficients are the same as that were used to obtain the best estimate.

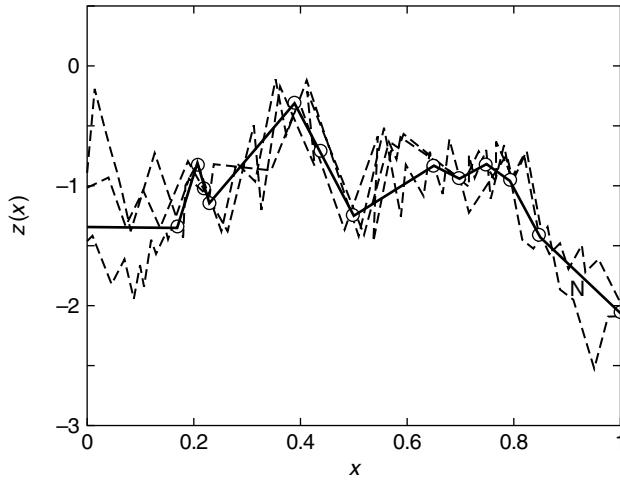


FIGURE 16.2 Three conditional realizations (discontinuous lines) and the best estimate (solid line).

16.3.3.2 Function Simulation

The functional form of $z_c(\mathbf{x}; k)$ is developed similarly as the best estimate in Section 16.3.2. First, generate an unconditional realization $z(\mathbf{x}; k)$, with zero mean. Then

$$z_c(\mathbf{x}; k) = - \sum_{j=1}^n \gamma(\mathbf{x} - \mathbf{x}_j) \xi_j + \hat{\beta} + z(\mathbf{x}; k) \quad (16.12)$$

where the ξ and $\hat{\beta}$ coefficients are found from system:

$$\begin{aligned} - \sum_{j=1}^n \gamma(\mathbf{x}_i - \mathbf{x}_j) \xi_j + \hat{\beta} &= z(\mathbf{x}_i) - z(\mathbf{x}_i; k), \quad \text{for } i = 1, \dots, n \\ \sum_{j=1}^n \xi_j &= 0 \end{aligned} \quad (16.13)$$

For illustration, three conditional realizations (out of many possible) of function $z(\mathbf{x})$ are shown together with the best estimate $\hat{z}(\mathbf{x})$ in Figure 16.2 for the same case as Figure 16.1. The conditional realizations have the structure described by the variogram and also reproduce the observations. Each conditional realization is equally likely to be the sought-after function. The average of all realizations is the estimate, which is a smooth function as shown in Figure 16.2. At any point \mathbf{x} , 95% of the conditional realizations should fall within the confidence interval (see Cressie, 1993; Kitanidis, 1996).

16.4 Variogram

16.4.1 Variogram Model

To perform ordinary kriging, an expression for the variogram is needed. In the isotropic case, the variogram γ is a function of separation distance only, $h = \|\mathbf{x} - \mathbf{x}'\|$, not the direction of the separation vector $\mathbf{x} - \mathbf{x}'$

TABLE 16.1 Common Variogram Models

Model	Expression, $\gamma(h)$	Parameters
Gaussian	$\sigma^2 \left(1 - \exp \left(-\frac{h^2}{L^2} \right) \right)$	$\sigma^2 > 0, L > 0$
Exponential	$\sigma^2 \left(1 - \exp \left(-\frac{h}{\ell} \right) \right)$	$\sigma^2 > 0, \ell > 0$
Spherical	$\begin{cases} \left(\frac{3}{2} \frac{h}{\alpha} - \frac{1}{2} \frac{h^3}{\alpha^3} \right) \sigma^2, & \text{for } 0 \leq h \leq \alpha \\ \sigma^2, & \text{for } h > \alpha \end{cases}$	$\sigma^2 > 0, \alpha > 0$
Nugget	$\begin{cases} C_0, & h > 0 \\ 0, & h = 0 \end{cases}$	$C_0 > 0$
Linear	$\theta \cdot h$	$\theta > 0$
Power	$\theta \cdot h^s$	$\theta > 0, 2 > s > 0$

between locations \mathbf{x} and \mathbf{x}' . For example, in three dimensions,

$$\mathbf{x} = \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \mathbf{x}_3 \end{bmatrix}, \quad \mathbf{x}' = \begin{bmatrix} \mathbf{x}'_1 \\ \mathbf{x}'_2 \\ \mathbf{x}'_3 \end{bmatrix}, \quad h = \sqrt{(\mathbf{x}_1 - \mathbf{x}'_1)^2 + (\mathbf{x}_2 - \mathbf{x}'_2)^2 + (\mathbf{x}_3 - \mathbf{x}'_3)^2} \quad (16.14)$$

We will consider the isotropic case first.

The variogram model must satisfy some mathematical requirements. In practice, the expression for the variogram is formed by combining a small number of simple mathematically acceptable expressions. Table 16.1 contains the most used isotropic variogram models.

By adding variograms from Table 16.1, one can obtain other mathematically acceptable variograms, for example, combining the linear and nugget effect variograms:

$$\gamma(h) = \begin{cases} C_0 + \theta h & h > 0 \\ 0, & h = 0 \end{cases} \quad (16.15)$$

with two parameters, $C_0 \geq 0$ and $\theta \geq 0$. This way, one can construct a model that adequately and parsimoniously (i.e., without unnecessary terms and parameters) represents information about the structure of z .

In many applications, due to geologic stratification, hydrodynamic mechanisms, and the like, the spatial structure may depend on direction as well as on distance. The simplest approach (Cressie, 1993; Kitanidis, 1997, describe more general cases) is to align the axes with the principal axes of anisotropy (e.g., parallel and perpendicular to layering) and to use as h :

$$h = \sqrt{\left(\frac{\mathbf{x}_1 - \mathbf{x}'_1}{l_1} \right)^2 + \left(\frac{\mathbf{x}_2 - \mathbf{x}'_2}{l_2} \right)^2 + \left(\frac{\mathbf{x}_3 - \mathbf{x}'_3}{l_3} \right)^2}$$

where l_1, l_2, l_3 are length parameters.

16.4.2 Variogram Selection

The variogram represents information about the structure of the function z and consequently should be consistent with (1) prior information and (2) the observations.

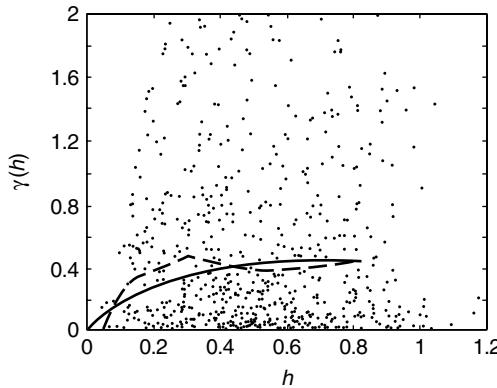


FIGURE 16.3 Raw (dots), experimental (dashed line), and model (solid line) variograms.

Prior information, sometimes referred to as “soft data,” is knowledge about the unknown that is additional to the observations. Prior information includes the experience and judgment of the analyst. In any specific application, there is always some prior information to the extent that even without observations one should be able to judge whether a variogram makes sense. Prior information is also important because certain features of the variogram cannot be established from the observations, no matter how dense. For example, if prior information suggests that function z should be smooth, then the Gaussian variogram (see Table 16.1) may be appropriate. If one expects that due to function “nonstationarity” (as when the likelihood of change in the hydrogeologic environment or geochemical regime increases rapidly with the distance) the estimate should not depend on measurements far away, one should choose a linear variogram. Certain decisions are easy to justify in the context of specific applications but, because here we deal with general methodologies, we will focus on the estimation of the variogram from data.

The process of developing a variogram from the data involves a number of steps and iterations. A variogram is selected tentatively based on an examination of the *experimental variogram*, as well as other information. The variogram is then calibrated (if it includes adjustable parameters) and tested on the basis of observations (or “hard data” or just “data”) using *residuals*.

16.4.3 Experimental Variogram

For all $n(n - 1)/2$ measurement pairs, plot the square difference $\frac{1}{2}[z(\mathbf{x}_i) - z(\mathbf{x}'_i)]^2$ against the separation distance $h = \|\mathbf{x}_i - \mathbf{x}'_i\|$. The plotting produces a cloud of points known as the *raw variogram*; the *experimental variogram* is a relatively regular line through this scatter plot (see Figure 16.3).

In the prevalent approach, the axis of separation distance is divided into consecutive intervals: the k th interval is denoted $[h_k^\ell, h_k^u]$ and contains N_k pairs of measurements $[z(\mathbf{x}_i), z(\mathbf{x}'_i)]$.

Compute:

$$\hat{\gamma}(h_k) = \frac{1}{2N_k} \sum_{i=1}^{N_k} [z(\mathbf{x}_i) - z(\mathbf{x}'_i)]^2 \quad (16.16)$$

where index i refers to a pair of measurements $z(\mathbf{x}_i)$ and $z(\mathbf{x}'_i)$ which meet the requirement

$$h_k^\ell < \|\mathbf{x}_i - \mathbf{x}'_i\| \leq h_k^u \quad (16.17)$$

The k th interval is represented by a single point h_k . Take h_k equal to the average value,

$$h_k = \frac{1}{N_k} \sum_{i=1}^{N_k} \|\mathbf{x}_i - \mathbf{x}'_i\| \quad (16.18)$$

Next, these points $[h_k, \hat{\gamma}(h_k)]$ are connected to form the experimental variogram (see Figure 16.3). Numerous modifications to this basic approach have been proposed to improve its robustness or to account for other effects (Cressie, 1993).

The appearance of the experimental variogram is affected by the choice of intervals. By choosing longer intervals, one averages over more points in each interval thus reducing jaggedness in the experimental variogram; however, one may sacrifice resolution, that is, “straighten” the variogram. One must use judgment in selecting the intervals to achieve reasonable trade-offs between the objectives of reducing jaggedness and preserving resolution. It usually suffices to use three to six intervals, and it makes sense to use shorter intervals where the points of the raw variogram are less spread out (in the vertical direction).

An experimental variogram that depends on orientation and is thus useful in identifying anisotropy is the so-called directional variogram (Journel and Huijbregts, 1978).

The experimental variogram, though it may provide a reasonable preliminary estimate of the variogram as shown in Figure 16.3, has its limitations (Armstrong, 1984) and should not monopolize the analysis. The selection of the variogram to be used in kriging should rely on the examination of the residuals.

16.4.4 Residuals

Arrange the n measurements in a given sequence. Calculate the kriging estimate of z at the second point, \mathbf{x}_2 , given only the first measurement, \mathbf{x}_1 : $\hat{z}_2 = z(\mathbf{x}_1)$ and $\sigma_2^2 = 2\gamma(\mathbf{x}_1 - \mathbf{x}_2)$. Calculate the actual error.

$$\delta_2 = z(\mathbf{x}_2) - \hat{z}_2 \quad (16.19)$$

and normalize by the standard error

$$\varepsilon_2 = \frac{\delta_2}{\sigma_2} \quad (16.20)$$

Use the same procedure to construct the other residuals. For the k th measurement location, estimate through kriging the value of z using only the first $k - 1$ measurements and normalize by the standard error. Thus,

$$\delta_k = z(\mathbf{x}_k) - \hat{z}_k, \quad \text{for } k = 2, \dots, n \quad (16.21)$$

$$\varepsilon_k = \frac{\delta_k}{\sigma_k}, \quad \text{for } k = 2, \dots, n \quad (16.22)$$

Using the actual data, the variogram model, and kriging, one thus computes the actual (or *experimental*) residuals, which are $n - 1$ δ or ε values.

According to the probabilistic model, the ε residuals satisfy:

$$E[\varepsilon_k] = 0, \quad k = 2, \dots, n \quad (16.23)$$

$$E[\varepsilon_k \varepsilon_\ell] = \begin{cases} 1, & \text{if } k = \ell, \\ 0, & \text{if } k \neq \ell, \end{cases} \quad k, \ell = 2, \dots, n \quad (16.24)$$

and are called *orthonormal* (Kitanidis, 1991, 1997).

16.4.5 Variogram Testing

Model validation means testing the agreement of the model with the data. In simple terms, a validation test is like an experiment with the following steps:

1. Predict the outcome to the experiment using the theory
2. Observe the actual outcome of the experiment
3. Compare the predicted and observed outcomes

The following tests are suggested to evaluate the model.

16.4.5.1 Q_1 Statistic

Experiment: Compute the average of the orthonormal residuals:

$$Q_1 = \frac{1}{n-1} \sum_{k=2}^n \varepsilon_k \quad (16.25)$$

Model Prediction: The model predicts that Q_1 is normally distributed with mean 0 and variance $1/(n-1)$.

Experimental Observation: Measure the experimental value of Q_1 (from the actual data).

Compare Model with Experiment: Agreement between the model and the data means that the experimental value of Q_1 is close to zero.

A reasonable rule is to reject the model if:

$$|Q_1| > \frac{2}{\sqrt{n-1}} \quad (16.26)$$

This rule involves a 5% probability that the correct model may be rejected.

16.4.5.2 Q_2 Statistic

The same procedure is followed with

$$Q_2 = \frac{1}{n-1} \sum_{k=2}^n \varepsilon_k^2 \quad (16.27)$$

Agreement between model and data means that the experimental value of Q_2 is near 1.

For $n > 40$, one may use (see Kitanidis, 1997, for procedure for smaller n): reject the model if

$$|Q_2 - 1| > 2.8/\sqrt{n-1} \quad (16.28)$$

16.4.5.3 Normality

The ε residuals should follow an approximately Gaussian distribution, because BLUE method makes most sense when estimation errors are approximately normal. Plotting the experimental $\varepsilon_2, \dots, \varepsilon_n$ on normal probability paper one may visually detect departures from normality, or one may perform goodness-of-fit tests (e.g., Shapiro and Wilk, 1965).

16.4.5.4 No Correlation

Finally, the experimental ε residuals should be uncorrelated, that is, they should have the pure nugget effect variogram,

$$\gamma(\mathbf{x}_i - \mathbf{x}_j) = \begin{cases} 1, & \text{for } \mathbf{x}_i \neq \mathbf{x}_j \\ 0, & \text{for } \mathbf{x}_i = \mathbf{x}_j \end{cases} \quad (16.29)$$

One may thus perform a variogram analysis and test the hypothesis that the experimental variogram of $\varepsilon_2, \dots, \varepsilon_n$ is indeed this γ . One can also plot the signs of the residuals seeking regular patterns that would contradict the model, which predicts that the residuals are uncorrelated.

16.4.6 Variogram Calibration

A measure of how close the model reproduces or fits the data is given by:

$$cR = Q_2 \exp \left(\frac{1}{n-1} \sum_{i=2}^n \ell n(\sigma_i^2) \right) \quad (16.30)$$

where σ_i^2 was defined in Section 16.4.4.

This quantity and Q_2 can be used to calibrate a variogram, that is, to estimate its parameters (Kitanidis, 1991). Consider, for example, that the model of Equation 16.15 has been selected but its parameters C_0 and θ need to be estimated from the data. Select the values $C_0 \geq 0$ and $\theta \geq 0$ that minimize

$$cR = \exp \left(\frac{1}{n-1} \sum_{i=2}^n \ell n(\sigma_i^2) \right) \quad (16.31)$$

subject to the constraint

$$\frac{1}{n-1} \sum_{i=2}^n \varepsilon_i^2 = 1 \quad (16.32)$$

This method is related to more general methods of statistics (Rao, 1973; Corbeil and Searle, 1976; Kitanidis, 1983). Also, such a parameter estimation method is sometimes referred to as a “cross-validation” method.

16.4.7 Nonnegativity

Quite often z , the function of interest, is known to be nonnegative. For example, z may represent transmissivity or solute concentration. However, kriging does not account for this nonnegativity requirement. If the 95% confidence interval turns out unrealistic because it includes negative values, then the nonnegativity constraint needs to be enforced.

The simplest approach is to make a one-to-one variable transformation. The prevalent transformation is the logarithmic:

$$y(\mathbf{x}) = \ln(z(\mathbf{x})) \quad (16.33)$$

Then, we proceed to perform kriging on the transformed variable. After the result is obtained, we back-transform the best estimate and the 95% confidence intervals or the conditional realizations. The results are guaranteed to be nonnegative. Issues of interpretation are discussed in Kitanidis (1997).

16.5 Kriging Variants

Consider estimation when z conforms to the general model of Equation 16.1. This formulation allows us to account for structural information by appropriately selecting the deterministic part in Equation 16.1,

as the following two examples illustrate:

Example 16.1

If $z(\mathbf{x})$ is the (depth averaged) piezometric head in an aquifer where a linear trend surface is known to exist,

$$f_1 = 1, \quad f_2 = x_1, \quad f_3 = x_2 \quad (16.34)$$

where x_1 and x_2 are the two Cartesian coordinates of location \mathbf{x} .

Example 16.2

If z represents log conductivity and the domain includes two zones with different hydro-geologic properties, interzonal differences in the mean are accounted for through

$$f_1 = \begin{cases} 1, & \text{in zone 1,} \\ 0, & \text{in zone 2,} \end{cases} \quad f_2 = \begin{cases} 0, & \text{in zone 1} \\ 1, & \text{in zone 2} \end{cases} \quad (16.35)$$

The covariance function $R(\mathbf{x} - \mathbf{x}')$ of the random part is defined as

$$R(\mathbf{x} - \mathbf{x}') = E[\varepsilon(\mathbf{x})\varepsilon(\mathbf{x}')] \quad (16.36)$$

The estimate of z at location \mathbf{x}_0 is

$$\hat{z}_0 = c(\mathbf{x}_0) + \sum_{i=1}^n \lambda_i(z(\mathbf{x}_i) - c(\mathbf{x}_i)) \quad (16.37)$$

where the coefficients are found (using the previously mentioned unbiasedness and best requirements) from the kriging system of $n + p$ linear equations with $n + p$ unknowns:

$$\begin{aligned} \sum_{j=1}^n R(\mathbf{x}_i - \mathbf{x}_j) \lambda_j + \sum_{k=1}^p f_k(\mathbf{x}_i) v_k &= R(\mathbf{x}_i - \mathbf{x}_0), \quad i = 1, \dots, n \\ \sum_{j=1}^n f_k(\mathbf{x}_j) \lambda_j &= f_k(\mathbf{x}_0), \quad k = 1, \dots, p \end{aligned} \quad (16.38)$$

where v_1, \dots, v_p are Lagrange multipliers. The mean square error is

$$E[(\hat{z}_0 - z(\mathbf{x}_0))^2] = - \sum_{k=1}^p f_k(\mathbf{x}_0) v_k - \sum_{i=1}^n \lambda_i R(\mathbf{x}_i - \mathbf{x}_0) + R(0) \quad (16.39)$$

From the general case, the following special cases are obtained:

Ordinary Kriging: When $c(\mathbf{x}) = 0, p = 1, f_1 = 1$. The ordinary kriging equations are obtained if we set $-\gamma$ in place of R in Equation 16.38 and Equation 16.39.

Kriging with External Drift and Simple Kriging: When $c(\mathbf{x})$ is given and there are no terms with unknown β coefficients. The kriging system is n equations with n unknowns and the v coefficients vanish. The even more special case of $c(\mathbf{x}) = \text{constant}$, is known as simple kriging.

Universal Kriging: When $c(\mathbf{x})$ vanishes and there are p unknown drift coefficients. This is also equivalent to kriging with generalized covariances because one needs a simplified version of the actual covariance function (in ordinary kriging, one can use $-\gamma$ instead of the covariance function). Issues of covariance estimation are discussed in Kitanidis (1993), and also the method of Section 16.7.3 can be applied.

16.6 Cokriging

Consider the case of N correlated functions, $z_1(\mathbf{x}), \dots, z_N(\mathbf{x})$, such as log-transmissivity, head, and accretion rate. One may want to estimate one of these functions, say $z_1(\mathbf{x})$, from observations of $z_1(\mathbf{x}), \dots, z_N(\mathbf{x})$ functions. The BLUE approach, known as *cokriging*, is similar to the one in kriging: after describing the joint structure of the N functions, the best estimate is a linear function of the data with coefficients determined from the requirements of unbiasedness and minimum variance.

The joint structure is described by:

Mean functions:

$$E[z_k(\mathbf{x})] = m_k(\mathbf{x}), \quad k = 1, \dots, N \quad (16.40)$$

Each of the mean functions may be represented as the deterministic part in Equation 16.1.

Covariance (or autocovariance) functions:

$$E[(z_k(\mathbf{x}) - m_k(\mathbf{x}))(z_k(\mathbf{x}') - m_k(\mathbf{x}'))] = R_{kk}(\mathbf{x}, \mathbf{x}'), \quad k = 1, \dots, N \quad (16.41)$$

Cross-covariance functions:

$$E[(z_k(\mathbf{x}) - m_k(\mathbf{x}))(z_\ell(\mathbf{x}') - m_\ell(\mathbf{x}'))] = R_{k\ell}(\mathbf{x}, \mathbf{x}'), \quad \text{for } k, \ell = 1, \dots, N \text{ and } k \neq \ell \quad (16.42)$$

Mean and autocovariance functions describe the structure of individual functions. Cross-covariance functions describe the interdependence of pairs of functions.

The cokriging equations are incorporated into the general equations of the next section. Determination of the model (structural analysis) for a set of functions is a more advanced topic that cannot be covered here in detail. Nevertheless, the next section outlines a method for estimation of parameters of autocovariance and cross-covariance functions and examples of application of cokriging are given in Dagan (1985) and Hoeksema et al. (1989).

16.7 Generalized Linear Estimation Equations

All linear kriging and cokriging equations can be written in a unified way, using matrix notation, as follows.

16.7.1 Model

$$\mathbf{y} = \mathbf{c} + \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon} \quad (16.43)$$

where \mathbf{y} is the $n \times 1$ measurement vector, \mathbf{c} is known $n \times 1$ vector, \mathbf{X} is a known $n \times p$ matrix, $\boldsymbol{\beta}$ is a $p \times 1$ vector of parameters (“drift coefficients” in kriging), and $\boldsymbol{\epsilon}$ is a random vector with zero mean and covariance matrix $\mathbf{Q}_{yy}(\boldsymbol{\theta})$ (a function of some covariance parameters). Also,

$$\mathbf{s} = \mathbf{c}_s + \mathbf{X}_s\boldsymbol{\beta} + \boldsymbol{\epsilon}_s \quad (16.44)$$

where \mathbf{s} is the $m \times 1$ vector of unknowns, \mathbf{c}_s is known $m \times 1$ vector, \mathbf{X}_s is a known $m \times p$ matrix, and $\boldsymbol{\epsilon}_s$ is a random vector with zero mean, covariance matrix $\mathbf{Q}_{ss}(\boldsymbol{\theta})$, and cross-covariance to $\boldsymbol{\epsilon}$ that is $\mathbf{Q}_{sy}(\boldsymbol{\theta})$.

16.7.2 Blue

$$\hat{s} = c_s + \Lambda(y - c) \quad (16.45)$$

where:

$$\begin{bmatrix} Q_{yy} & X \\ X^T & 0 \end{bmatrix} \begin{bmatrix} \Lambda^T \\ M \end{bmatrix} \begin{bmatrix} Q_{ys} \\ X_s^T \end{bmatrix} \quad (16.46)$$

The estimation error covariance matrix is:

$$E[(\hat{s} - s)(\hat{s} - s)^T] = -X_s M + Q_{ss} - Q_{sy} \Lambda^T \quad (16.47)$$

16.7.3 Parameter Estimation

Select covariance parameters θ that minimize:

$$\left(|Q_{yy}| |X^T Q_{yy}^{-1} X| |X^T X|^{-1} \right)^{1/(n-p)} \quad (16.48)$$

while at the same time satisfying:

$$Q_2 = \frac{1}{n-p} y^T \left(Q_{yy}^{-1} - Q_{yy}^{-1} X (X^T Q_{yy}^{-1} X)^{-1} X^T Q_{yy}^{-1} \right) y = 1 \quad (16.49)$$

where $| |$ applied on a square matrix indicates matrix determinant. Note that this method, which is a generalization of the method of minimizing cR in ordinary kriging, is appropriate for finding the parameters of generalized covariance functions and variograms (which are needed in kriging or cokriging) and not necessarily of ordinary covariance functions (see Kitanidis, 1993).

16.8 Inverse Problems

16.8.1 Definition

To be successful in modeling groundwater flow and transport, one needs to obtain representative values of the required parameters, such as transmissivity, storage coefficient, conductivity, dispersivity, retardation factor, reaction rate, and so on. The task is particularly difficult when a parameter is known to vary in space but is not measured directly. For example, conductivity (or transmissivity) varies over the flow domain, but there are not enough direct observations to resolve its spatial variability.

The inverse problem of groundwater modeling is defined broadly as the estimation of parameters from observations of the system response. The quintessential problem is the determination of conductivity from observations of the hydraulic head. Usually, this goal is achieved through systematic adjustment of parameter values to match the observations (as illustrated in Carrera and Neuman, 1986; Cooley et al., 1986), hence the term *history matching* is also used. Many approaches have been proposed, and the literature on the subject is vast. There have been several literature reviews (Yeh, 1986; Carrera, 1987; Ginn and Cushman, 1990) and a book devoted to the subject (Sun, 1994).

In this chapter, the inverse problem is defined as:

Estimation of at least one spatial function $z(\mathbf{x})$ from observations of other quantities related to z through a mathematical model as well as direct measurements of z .

The method to be presented is an extension of the geostatistical methods for the solution of the interpolation problem. We will examine exclusively static estimation problems (where all observations are given at the same time, as opposed to methods where data are taken into account as they become available). The justification for the approach that follows is given in Kitanidis (1995).

16.8.2 Methodology

Consider that we want to estimate parameter $z(\mathbf{x})$, a spatial function such as the log conductivity. The function is represented through the model of Equation 16.1. After discretization, $z(\mathbf{x})$ is represented through an n by 1 vector \mathbf{s} . The mean of \mathbf{s} is

$$E[\mathbf{s}] = \mathbf{X}\boldsymbol{\beta} \quad (16.50)$$

where \mathbf{X} is a known $n \times p$ matrix, $\boldsymbol{\beta}$ are p unknown drift coefficients. Furthermore, \mathbf{s} has a covariance matrix

$$E[(\mathbf{s} - \mathbf{X}\boldsymbol{\beta})(\mathbf{s} - \mathbf{X}\boldsymbol{\beta})^T] = \mathbf{Q}(\theta) \quad (16.51)$$

that is considered a known function of parameters θ . If there are two or more spatially variable parameters that need to be estimated (such as log transmissivity and storage coefficient), this case is included in this formulation by considering \mathbf{s} as the aggregation of all spatially variable parameters. The objective is to find the best estimate of \mathbf{s} and to evaluate the error of estimation.

There may also be other unknown parameters, such as boundary conditions or the intensity of sources or sinks, but these will not be included here because we focus on the estimation of the spatial process and its structural parameters. The $\boldsymbol{\beta}$ and θ parameters are treated as unknown constants and are supposed to be few in number, certainly far fewer than the observations. The observations are related to the unknown spatial process and this relation is denoted through:

$$\mathbf{y} = \mathbf{h}(\mathbf{s}) + \mathbf{v} \quad (16.52)$$

where \mathbf{y} is the $m \times 1$ vector of observations. The observation error \mathbf{v} is random with normal distribution, zero mean, and covariance matrix \mathbf{R} that is fixed or a known function of an expanded set of parameters θ . The standard deviations of the measurement errors, which are the square root of the diagonal elements of \mathbf{R} , define how closely the observations should be reproduced.

An intuitive approach is to minimize the weighted least squares criterion:

$$(z - \mathbf{h}(\mathbf{s}))^T \mathbf{R}^{-1} (z - \mathbf{h}(\mathbf{s})) + (\mathbf{s} - \mathbf{X}\boldsymbol{\beta})^T \mathbf{Q}^{-1} (\mathbf{s} - \mathbf{X}\boldsymbol{\beta}) \quad (16.53)$$

where the first term represents the objective of reproducing the observations and the second term the objective of being consistent with the structure. The minimization should be with respect to \mathbf{s} and $\boldsymbol{\beta}$ vectors. The covariance parameters are to be imposed by the analyst based on prior information or inferred from the data through a process of cross-validation.

This approach is systematized into the two phases of a geostatistical approach:

1. Structural analysis, where the form of the mean and the covariance function are selected and the structural parameters are estimated.
2. Conditioning on the data, where vector \mathbf{s} is estimated.

The implementation of the second step is an iterative cokriging approach. We start with an estimate of \tilde{s} and we improve at each iteration. Find the derivative of \mathbf{h} about s at \tilde{s} :

$$\mathbf{H} = \frac{\partial \mathbf{h}}{\partial s} \Big|_{s=\tilde{s}} \quad (16.54)$$

Then, assuming that the actual s is close to \tilde{s} , approximate

$$\mathbf{y} = \mathbf{h}(\tilde{s}) + \mathbf{H}(s - \tilde{s}) + \mathbf{v} \quad (16.55)$$

Define

$$\Sigma = \mathbf{H} \mathbf{Q} \mathbf{H}^T + \mathbf{R}$$

Solve the cokriging system of equations:

$$\begin{bmatrix} \Sigma & \mathbf{H} \mathbf{X} \\ (\mathbf{H} \mathbf{X})^T & 0 \end{bmatrix} \begin{bmatrix} \Lambda^T \\ \mathbf{M} \end{bmatrix} = \begin{bmatrix} \mathbf{H} \mathbf{Q} \\ \mathbf{X}^T \end{bmatrix} \quad (16.56)$$

where Λ is an $m \times n$ matrix of coefficients and \mathbf{M} is a $p \times n$ matrix of multipliers. Then, the cokriging estimate is:

$$\hat{s} = \Lambda(\mathbf{y} - \mathbf{h}(\tilde{s}) + \mathbf{H}\tilde{s}) \quad (16.57)$$

If \hat{s} is practically equal to \tilde{s} , the algorithm has converged and the covariance matrix of estimation is

$$\mathbf{V} = -\mathbf{X} \mathbf{M} + \mathbf{Q} - \mathbf{Q} \mathbf{H}^T \Lambda^T \quad (16.58)$$

Otherwise \tilde{s} is set equal to \hat{s} and the procedure is repeated. Examples are found in Kitanidis (1995) and Yeh et al. (1996).

The cokriging equations assume that the covariance parameters, θ , are known and in some approaches are not estimated in any formal way. They can be estimated (Kitanidis, 1995) by minimizing the criterion:

$$\begin{aligned} L = & \ln |\Sigma| + \ln |\mathbf{X}^T \mathbf{H}^T \Sigma^{-1} \mathbf{H} \mathbf{X}| + (\mathbf{y} - \mathbf{h}(\hat{s}) + \mathbf{H}\hat{s})^T \\ & \times (\Sigma^{-1} - \Sigma^{-1} \mathbf{H} \mathbf{X} (\mathbf{X}^T \mathbf{H}^T \Sigma^{-1} \mathbf{H} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{H}^T \Sigma^{-1}) (\mathbf{y} - \mathbf{h}(\hat{s}) + \mathbf{H}\hat{s}) \end{aligned} \quad (16.59)$$

This approach is essentially a restricted maximum likelihood or cross-validation methodology. A variation to this approach is to minimize the sum of the first two terms subject to the constraint that the third term is equal to 1, which is the method of Section 12.7.3. The two approaches yield similar or identical results.

After the covariance parameters are updated, one may need to repeat the iterative procedure of estimating s .

The error computed in the iterative cokriging or “quasilinear” approach is valid provided that the final estimation error is small. In small-variance cases (Kitanidis and Vomvoris, 1983; Hoeksema and Kitanidis, 1984, 1985, 1989; Dagan, 1985; Rubin and Dagan, 1987a; Wagner and Gorelick, 1989), the observation Equation 16.52 is linearized only once, about the mean (deterministic part). Sometimes, Monte Carlo methods can be used instead of the usual linearization (Hoeksema and Clapp, 1990). The approach is effectively the same with the method of Carrera and Neuman (1986) in the conditioning step but differs in the step of estimating θ . Conceptually and in simple terms, the relation of this approach to the least squares approach (Cooley and Naff, 1990, for example) is that in least squares the spatial variability is accounted for by including sufficient terms in the deterministic part; the covariance matrix \mathbf{Q} is proportional to the identity matrix; and the problem is formulated to estimate the β coefficients.

16.8.3 Derivative Computation

Computationally, the greatest challenge in the application of this method with numerical models is the computation of the derivative matrix \mathbf{H} in Equation 16.54. The direct approach of varying one element of \mathbf{s} at a time is inefficient because usually there are many components in \mathbf{s} and few components in \mathbf{y} . More efficient “adjoint state” methods have been developed. This topic, although important, is beyond the scope of this chapter, and the reader is referred to Chavent et al. (1975), Neuman (1980), Townley and Wilson (1985), Carrera et al. (1990), and Sun (1994).

Glossary

BLUE: Acronym for Best Linear Unbiased Estimation. Any estimation procedure where the estimate depends linearly on the data, with weights selected so that the estimation error has zero mean and minimum variance.

Cokriging: A geostatistical estimation method where one function is estimated using measurements from another related function.

Conditional Realizations: Functions that are possible solutions to the problem of finding a function from observations. They are consistent with the structure and with the data.

Deterministic Part: Also known as mean, drift, or trend. The part of the function that is described through a deterministic expression, such as a polynomial or zones with uniform values.

Experimental: Indicates computed from data.

Experimental Variogram: An estimate of the variogram obtained from the data using a graphical procedure. It is usually shown as piecewise linear.

Generalized Covariance: The part of the covariance that matters in universal kriging.

Geostatistics: A method for the estimation of spatial variables. It is based on the theory of random functions.

Kriging: Denotes any “geostatistical” interpolation or averaging method.

Ordinary Kriging: BLUE for the case that the mean is an unknown constant and the variogram is specified.

Orthonormal: Normalized random variables, which are uncorrelated and have zero mean and unit variance.

Random Function: Defined through probabilistic averages. Also known as *stochastic process*.

Random Part: The part of the unknown function that is described as a random function with zero mean. Usually, the less “regular” part of the function.

Residuals: Differences between model predictions and data. Useful in model testing and in parameter estimation.

Structure: Common characteristics of an ensemble of functions. In this chapter, structure is quantified by expressions for the deterministic part and the covariance function of the random part.

Variogram: Used synonymously with semivariogram. A measure of spatial structure. The expected value of half the square difference expressed as a function of the separation.

References

- Armstrong, M. 1984. Common problems seen in variograms. *Math. Geol.* 16(3), 305–313.
- ASCE Task Committee on Geohydrology. 1990. Review of geostatistics in geohydrology, Parts I and II. *ASCE J. Hydraul. Eng.* 116(5), 612–658.
- Carrera, J. and Neuman, S. P. 1986. Estimation of aquifer parameters under transient and steady state conditions, 1. Maximum likelihood method incorporating prior information. *Water Resour. Res.* 22(2), 199–210.
- Carrera, J. 1987. State of the art of the inverse problem applied to the flow and solute transport equations, in NATO ASI Ser., Eds. E. Curtodio, A. Gurgui, and J. P. Lobo-Ferreira. D. Reidel, Hingham, MA.

- Carrera, J., Navarrina, F., Vives, L., Heredia, J., and Medina, A. 1990. Computational aspects of the inverse problem, in *Computational Methods in Water Resources*. Computational Mechanics Publications, Billerica, MA, pp. 513–523.
- Chavent, G., Dupuy, M., and Lemonnier, P. 1975. History matching by use of optimal theory. *Soc. Pet. Eng. J.* 15(1), 74–86.
- Cooley, R. L., Konikow, L. F., and Naff, R. L. 1986. Nonlinear-regression groundwater flow modeling of a deep regional aquifer system. *Water Resour. Res.* 22(13), 1759–1778.
- Cooley, R. L. and Naff, R. L. 1990. Regression modeling of ground-water flow, USGS, Techniques of Water Resources Investigations, 03-B4. Washington, DC.
- Cressie, N. A. C. 1993. *Statistics for Spatial Data*. John Wiley & Sons, New York.
- Corbeil, R. R. and Searle, S. R. 1976. Restricted maximum likelihood estimation of variance components in the mixed model. *Technometrics* 18(1), 31–38.
- Dagan, G. 1985. Stochastic modeling of groundwater flow by unconditional and conditional probabilities: The inverse problem. *Water Resour. Res.* 21(1), 65–72.
- Dagan, G. 1989. *Flow and Transport in Porous Media*. Springer-Verlag, Berlin.
- de Boor, C. 1978. *A Practical Guide to Splines*. Springer-Verlag, New York.
- de Marsily, G. 1986. *Quantitative Hydrogeology*. Academic Press, New York.
- Deutsch, C. V. and Journel, A. G. 1992. *GSLIB: Geostatistical Software Library and User's Guide*. Oxford University Press, New York.
- Dykaar, B. B. and Kitanidis, P. K. 1992. Determination of the effective hydraulic conductivity for heterogeneous porous media using a numerical spectral approach 1. Method. *Water Resour. Res.* 28(4), 1155–1166.
- Gelhar, L. W. 1993. *Stochastic Subsurface Hydrology*. Prentice Hall, Englewood Cliffs, NJ.
- Ginn, T. R. and Cushman, J. H. 1990. Inverse methods for subsurface flow: A critical review of stochastic techniques. *Stoch. Hydrol. Hydraul.* 4, 1–26.
- Gutjahr, A., Bullard, B., Hatch, S., and Hughson, L. 1994. Joint conditional simulations and the spectral approach for flow modeling. *Stoch. Hydrol. Hydraul.* 8(1), 79–108.
- Hoeksema, R. J. and Kitanidis, P. K. 1984. An application of the geostatistical approach to the inverse problem in two-dimensional groundwater modeling. *Water Resour. Res.* 20(7), 1003–1020.
- Hoeksema, R. J. and Kitanidis, P. K. 1985. Comparison of Gaussian conditional mean and kriging estimation in the geostatistical solution of the inverse problem. *Water Resour. Res.* 21(6), 825–836.
- Hoeksema, R. J. and Kitanidis, P. K. 1989. Prediction of transmissivities, heads, and seepage velocities using mathematical models and geostatistics. *Adv. Water Resour.* 12(2), 90–102.
- Hoeksema, R. J., Clapp, R. B., Thomas, A. L., Hunley, A. E., Farrow, N. D., and Dearstone, K. C. 1989. Cokriging model for estimation of water table elevation. *Water Resour. Res.* 25(3), 429–438.
- Hoeksema, R. J. and Clapp, R. B. 1990. Calibration of groundwater flow models using Monte Carlo simulations and geostatistics, in *ModelCARE 90: Calibration and Reliability in Groundwater Modelling*. IAHS Publ. No 195, pp. 33–42.
- Isaaks, E. H. and Srivastava, R. M. 1989. *Applied Geostatistics*. Oxford University Press, New York.
- Journel, A. G. and Huigbrechts, Ch. J. 1978. *Mining Geostatistics*. Academic Press, New York.
- Kitanidis, P. K. 1983. Statistical estimation of polynomial generalized covariance functions and hydrologic applications. *Water Resour. Res.* 19(4), 909–921.
- Kitanidis, P. K. and Vomvoris, E. G. 1983. A geostatistical approach to the inverse problem in groundwater modeling (steady state) and one-dimensional simulations. *Water Resour. Res.* 19(3), 677–690.
- Kitanidis, P. K. 1991. Orthonormal residuals in geostatistics: Model criticism and parameter estimation. *Math. Geol.* 23(5), 741–758.
- Kitanidis, P. K. 1993. Generalized covariance functions in estimation. *Math. Geol.* 25(5), 525–540.
- Kitanidis, P. K. 1995. Quasilinear geostatistical theory for inversing. *Water Resour. Res.* 31(10), 2411–2419.
- Kitanidis, P. K. 1996. Analytical expressions of conditional mean, covariance, and sample functions in geostatistics. *Stoch. Hydrol. Hydraul.* 10(4), 279–294.

- Kitanidis, P. K. 1997. *Introduction to Geostatistics: With Applications in Hydrogeology*. Cambridge University Press, Cambridge, 249pp.
- Mantoglou, A. and Wilson, J. L. 1982. The turning bands method for simulation of random fields using line generation by a spectral method. *Water Resour. Res.* 18(5), 1379–1394.
- Matheron, G. 1971. *The Theory of Regionalized Variables and Its Applications*. Ecole de Mines, Fontainbleau, France.
- Neuman, S. P. 1980. Adjoint-state finite element equations for parameter estimation, in *Finite Elements in Water Resources: Proceedings of Third International Conference*. Eds. S. W. Wang, C. V. Alonso, C. A. Brebbia, W. G. Gray, and G. F. Pinder. Springer-Verlag, New York.
- Rao, C. R. 1973. *Linear Statistical Inference and its Applications*. John Wiley & Sons, New York.
- Rendu, J.-M. 1981. *An Introduction to Geostatistical Methods of Mineral Evaluation*. South African Institute of Mining and Metallurgy, Johannesburg.
- Rivoirard, J. 1994. *Introduction to Disjunctive Kriging and Non-Linear Geostatistics*. Oxford University Press, New York.
- Robin, M., Gutjahr, A., Sudicky, E., and Wilson, J. L. 1993. Cross-correlated random field generation with the direct Fourier transform method. *Water Resour. Res.* 29(7), 2385–2397.
- Rubin, Y. and Dagan, G. 1987a. Stochastic identification of transmissivity and effective recharge in steady groundwater flow, 1. Theory. *Water Resour. Res.* 23(7), 1185–1192.
- Rubin, Y. and Dagan, G. 1987b. Stochastic identification of transmissivity and effective recharge in steady groundwater flow, 2. Case study. *Water Resour. Res.* 23(7), 1193–1200.
- Shapiro, S. S. and Wilk, M. B. 1965. An analysis of variance test for normality (Complete Samples). *Biometrika*, 52, 591–610.
- Sun, N.-Z. 1994. *Inverse Problems in Groundwater Modeling*. Kluwer, Norwell, MA.
- Thiebaux, H. J. and Pedder, M. A. 1987. *Spatial Objective Analysis: With Applications in Atmospheric Science*. Academic Press, London.
- Tompson, A. F. B., Ababou, R., and Gelhar, L. W. 1989. Implementation of the three dimensional turning bands random field generator. *Water Resour. Res.* 25(10), 2227–2243.
- Townley, L. R. and Wilson, J. L. 1985. Computationally efficient algorithms for parameter estimation and uncertainty propagation in numerical models of groundwater flow. *Water Resour. Res.* 21(12), 1851–1860.
- Wagner, B. J. and Gorelick, S. M. 1989. Reliable aquifer remediation in the presence of spatially variable hydraulic conductivity: From data to design. *Water Resour. Res.* 25(10), 2211–2225.
- Yeh, W. W.-G. 1986. Review of parameter identification procedures in groundwater hydrology: The inverse problem. *Water Resour. Res.* 22(1), 95–108.
- Yeh, J. T.-C., Jin, M., and Hanna, S. 1996. An iterative stochastic inverse method: Conditional effective transmissivity and hydraulic head fields. *Water Resour. Res.* 32(1), 85–92.

Further Information

Cressie (1993) is a comprehensive reference on methods of statistical estimation of spatial functions. Journel and Huigbrechts (1978) and Rendu (1981) describe geostatistics as applied in mining, including methods for the estimation of volume averages. Isaaks and Srivastava (1989) is another reference on basic geostatistics. Software is found in Deutsch and Journel (1992). Kitanidis (1997) is an introduction to estimation of spatial functions with applications in hydrogeology. de Marsily (1986) has a chapter on geostatistics including universal kriging and cokriging. A two-part paper (ASCE, 1990) reviews applications of geostatistics in hydrogeology. Statistical groundwater mechanics are covered in Dagan (1989) and Gelhar (1993). Rivoirard (1994) covers nonlinear geostatistics. Sun (1994) covers inverse methods in hydrogeology including methods for estimation and sensitivity analysis (computation of derivatives) in steady and unsteady state. Cooley and Naff (1990) describe the least squares approach to inverse modeling.

17

Groundwater Contaminants

17.1	Introduction.....	17-1
17.2	Physicochemical Characteristics of Water	17-2
17.3	Factors Affecting Aqueous Solubility	17-3
	pH and Chemical Composition • Sorption to Solids • p_e • Temperature	
17.4	Inorganic Constituents	17-14
	Source Water Composition • Source Water:Soil Interactions	
17.5	Organic Constituents	17-19
	Background Organics • Nonaqueous Phase Liquids • High-Profile Organic Contaminants	
17.6	Radionuclides	17-24
	Radioactive Decay • Definitions and Units • Important Radionuclides and Their Sources • Radioactive Waste Disposal • Remediation	
17.7	Particulates	17-29
	Glossary.....	17-30
	References	17-31
	Further Information	17-32

Ernest R. Blatchley, III
and John E. Thompson
Purdue University

17.1 Introduction

Webster's New Collegiate Dictionary defines the verb "contaminate" as a process by which a material is made "inferior or impure by admixture" or made "unfit for use by the introduction of unwholesome or undesirable elements". While both of these definitions are accurate, they do not fully represent the connotation with which the word is used. For purposes of this chapter, the term "contaminant" will generally refer to an undesirable constituent that is introduced directly or indirectly as a result of human activity. Furthermore, the identification of groundwater contaminants and their effects will require knowledge of pre-existing or background conditions.

Groundwater contaminants exist in many forms, and contaminant classification schemes can be based on any of several physicochemical characteristics. For example, contaminants may be classified based on their preference for association with the aqueous phase or with particles. Contaminant distinction based on phase preference is important as the phase(s) a contaminant associates with can affect its transport behavior and toxicology. Furthermore, the form taken by a contaminant can also affect

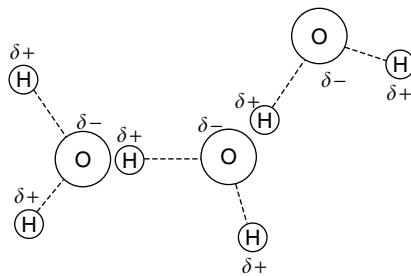


FIGURE 17.1 Schematic representation of the structure of the water molecule and resultant hydrogen bonding. The symbols $\delta+$ and $\delta-$ represent the positive and negative ends of the dipole, respectively. Alternatively, these symbols may be viewed as markers of areas of relatively low and high electron density, respectively.

the choice of treatment processes that may be implemented to remediate a contaminated area. In general, the physicochemical characteristics of groundwater contaminants and the surrounding aquifer will play critical roles in determining their fate, transport, and effects.

The goal of this chapter is to define the nature and behavior of groundwater contaminants. Important physicochemical characteristics will be defined and described in terms of their potential effects on contaminants. The major contaminant classes will be examined and examples of important contaminants within each class will be provided.

17.2 Physicochemical Characteristics of Water

Water molecules are comprised of two hydrogen atoms and one oxygen atom. The oxygen atom, being more electronegative than hydrogen, preferentially attracts electrons, thereby leading to a nonuniform electron distribution around the water molecule. Together with the two pairs of unbound electrons carried by the oxygen atom, the water molecule displays an asymmetrical structure (see Figure 17.1). The hydrogen atoms form covalent bonds with the oxygen atom at an angle of 105°.

The structure of water leads to several properties that are critical to its use as a resource and to the behavior of contaminants. Water is a highly polar molecule, with the negative end of the dipole being associated with the oxygen atom, and two strongly positive regions near the hydrogen atoms. One result of this structure is the ability of water molecules to attract each other through the formation of hydrogen bonds. These bonds, which are weak relative to covalent bonds, allow water to display properties that are somewhat anomalous based on its molecular size and composition. For example, water has a higher boiling point than hydrogen sulfide (H_2S) even though sulfur, which lies immediately below oxygen on the periodic table, is heavier than oxygen. As a result, water exists predominantly as a liquid at room temperature and pressure, whereas H_2S exists predominantly in the gas phase under the same conditions. The polarity of water is critical to its ability to function as a solvent. The dipolar nature of water allows it to associate with positively or negatively charged species; in turn, this property allows water to separate (dissociate) ionizable constituents, leading to dissolution.

The ability of water to function as a solvent is not limited to ionizable constituents. Essentially all compounds have measurable solubility in water. However, compounds that do not form ions will express affinity for water molecules via another mechanism. An important example is the ability of water to dissolve polar molecules. For example, if one examines the dipole moments of molecules in a homologous series of organic acids, decreasing polarity and aqueous solubility are observed as the molecule size increases. In essence, with the addition of each methyl group, the effects of the nonpolar group become less important as the electrons can distribute themselves over a larger molecule. Alternately, one could view the molecule as becoming more aliphatic as the chain length increases. Regardless of the viewpoint, the effect of polarity on water solubility is clear. In general terms, molecular solubility in water increases with polarity.

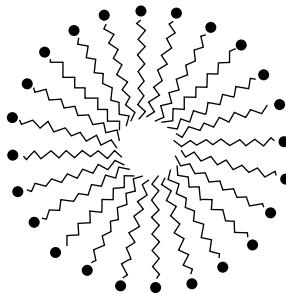


FIGURE 17.2 Two-dimensional schematic representation of a surfactant micelle in aqueous solution. The hydrophilic head is indicated by the filled circle (●); this end of the surfactant molecule will associate preferentially with the polar solvent (water). The nonpolar tails will be repelled by the water and allow the surfactant molecules to agglomerate into the micelles.

Natural or anthropogenic additions to water may also affect the solubility of constituents for water. For example, surfactants and detergents can dramatically increase the solubility of nonpolar compounds. In the case of surfactants, this can be accomplished through the formation of micelles, which exist as essentially a phase within a phase, often referred to as a “pseudo-phase” (see Figure 17.2). Micelles are agglomerations of organic molecules in which the polar ends of the surfactant molecules are directed outward, to associate with the bulk aqueous phase, and the relatively nonpolar ends of these molecules form the micelle core, yielding a relatively nonpolar region that promotes solubilization of nonpolar constituents. Detergents often serve as sequestering agents for Ca^{2+} and Mg^{2+} (the primary constituents of chemical hardness), thereby preventing interactions between hardness constituents and surfactants.

17.3 Factors Affecting Aqueous Solubility

The aqueous solubility of a contaminant compound will influence the transport, fate, and toxicity of that compound in a groundwater system. As such, it is important to define the characteristics of the system that can affect solubility. The primary factors that contribute to aqueous solubility are defined and reviewed below.

17.3.1 pH and Chemical Composition

Often referred to as a “master variable” in describing water composition, no single factor plays a more universal role in defining the characteristics of an aqueous system than pH. Water undergoes autoionization, resulting in the production of H^+ and OH^- .



In reality, ions such as H^+ and OH^- do not exist in solution, at least not as written. As described previously, ions are surrounded by water molecules. Therefore, the notation H^+ is really a shorthand representation for the hydrated proton that could be more accurately represented by $n(\text{H}_2\text{O})\text{H}^+$, where the coefficient n represents the number of waters of hydration. For purposes of this chapter, shorthand notation will be used to represent H^+ , OH^- , and all other ions. The strict definition of pH is as follows:

$$\text{pH} = -\log_{10}(\text{H}^+) \quad (17.2)$$

where (H^+) is defined as the activity of the hydrogen ion. For solutions of low ionic strength, activity may be closely approximated by molar concentration.

TABLE 17.1 Temperature Dependence of the Ion Product of Water (K_w)-Temperature (°C) pK_w

0	14.9435
5	14.7338
10	14.5346
15	14.3463
20	14.1669
24	14.0000
25	13.9965
30	13.8330
35	13.6801
40	13.5348
45	13.3960
50	13.2617
55	13.1369
60	13.0171

Note: Values as tabulated in the *CRC Handbook of Chemistry and Physics, 59th Edition*. 1978. CRC Press, Boca Raton, FL. With permission.

As the activity of molecular water (solvent) in aqueous solution is essentially constant, it is (by convention) excluded from the formal representation of the equilibrium constant for autoionization. Therefore, the equilibrium constant for reaction (17.1) is defined by:

$$K_w = (H^+)(OH^-) \quad (17.3)$$

K_w is also referred to as the ion product of water. For most acid-base reactions, equilibrium conditions in homogeneous aqueous solutions are established rapidly. However, it should be noted that the values of acid-dissociation constants, including K_w , are temperature-dependent (see Table 17.1). Under some circumstances, variations in K_w can have measurable consequences in a system. For example, in a typical mid-latitude groundwater system, with a characteristic temperature of roughly 13°C, the condition of neutrality (equal concentrations of H^+ and OH^-) will occur at a pH of approximately 7.2, whereas at room temperature, neutrality corresponds to pH = 7.0. In a physical sense, pH is a representation of the availability of hydrogen ions (protons) in solution; analytical determinations of pH are based on measurement of electrochemical potential using a hydrogen-ion selective electrode that has been previously calibrated using standard solutions.

Solution pH has a dramatic effect on the characteristics of solubilized compounds that can donate or accept protons (acids and bases, respectively). Some compounds have the ability to act as both acids and bases. Many examples of these so-called "amphoteric" compounds can be found, none more important than water itself. Since K_w has been accurately quantified as a function of temperature, knowledge of pH and temperature yields knowledge of both (H^+) and (OH^-).

The effect of pH on aqueous chemistry is not limited to the autoionization of the water molecule; pH will determine the speciation of all acids and bases in solution. An important example in groundwater chemistry is the distribution of carbonate species. Gaseous carbon dioxide (CO_2) will dissolve in water to yield aqueous CO_2 , followed by hydration to yield carbonic acid (H_2CO_3) that are often represented jointly by the fictitious species $H_2CO_3^*$. The equilibrium partitioning of gaseous and aqueous CO_2 are defined by Henry's law:

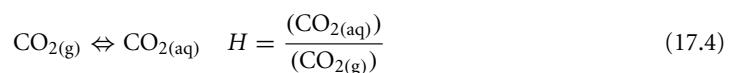


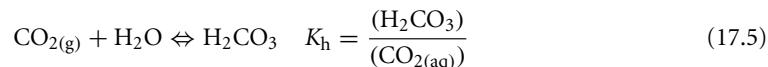
TABLE 17.2 Temperature Dependence of Equilibrium Constants for the Carbonate System

T (°C)	− log ₁₀ H	pK _{a,1}	pK _{a,2}
5	1.20	6.52	10.56
10	1.27	6.46	10.49
15	1.34	6.42	10.43
20	1.41	6.38	10.38
25	1.47	6.35	10.33
40	1.64	6.30	10.22
60	1.80	6.30	10.14

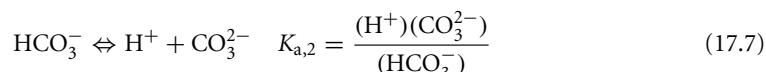
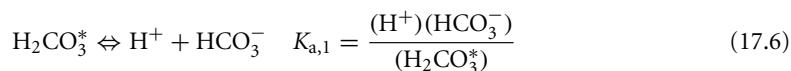
Source: Data from Larson, T. E. and Buswell, A. M. 1942. *Journal, AWWA*, **34**, 1664. With permission.

As activity (concentration) in the gas and liquid phases can be expressed in many units, many forms of Henry's law exist within the literature.

CO₂ hydration is defined by the following reaction:

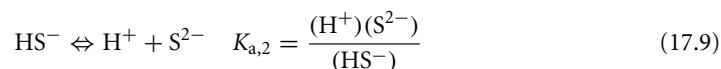
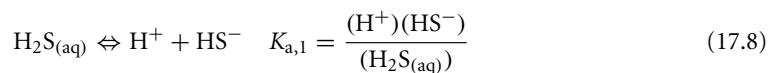


The equilibrium distribution of CO_{2(aq)} and H₂CO₃ is such that the vast majority of the dissolved CO₂ attributed to these species is found in the form of CO_{2(aq)}. However, the hydration reaction described by Equation 17.5 provides a direct link for CO₂ to the carbonate system. Therefore, the hypothetical compound H₂CO₃^{*} is used to represent the combination of these two compounds. Carbonic acid can donate two protons and is therefore referred to as a diprotic acid:



The equilibrium constants that describe these reactions are all functions of temperature. A compilation of these constants for the range of temperatures relevant to this subject can be found in Table 17.2. The information presented in Equations 17.4 through 17.7 and Table 17.2 reveal that with knowledge of pH and the concentration of any of the CO₂-based constituents, it is possible to quantify the concentrations (activities) of all CO₂-based constituents. As will be discussed later, the pH-dependence of the solubility of many metals in groundwater systems is closely related to the distribution of carbonate species.

Acid-base behavior determines the physicochemical characteristics of many other compounds. For example, H₂S is a relatively volatile, toxic compound. It is a product of some anaerobic biochemical processes, and can yield the sulfide ion (S²⁻), which can play an important role in determining the solubility and mobility of metals. The acid/base behavior of the diprotic acid H₂S can have important implications regarding the fate of sulfide, as well as many other aqueous constituents:



Dissociation of aqueous H₂S leads to the production of ionic species that have high solubilities relative to their parent compound. Furthermore, the sulfide ion (S²⁻) forms precipitates with many metals that have

TABLE 17.3 Acid Dissociation Constants for Common Ionizable Groundwater Constituents at 25°C

Compound (formula)	Conjugate Base (formula)	pK _a
Hydrochloric acid (HCl)	Chloride (Cl ⁻)	~−3
Sulfuric acid (H ₂ SO ₄)	Bisulfate (HSO ₄ ⁻)	~−3
Nitric acid (HNO ₃)	Nitrate (NO ₃ ⁻)	−1
Bisulfate (HSO ₄ ⁻)	Sulfate (SO ₄ ²⁻)	1.9
Phosphoric acid (H ₃ PO ₄)	Dihydrogen phosphate(H ₂ PO ₄ ⁻)	2.1
Acetic acid (CH ₃ COOH)	Acetate (CH ₃ COO ⁻)	4.7
Hydrogen sulfide (H ₂ S)	Bisulfide (HS ⁻)	7.1
Dihydrogen phosphate (H ₂ PO ₄ ⁻)	Hydrogen phosphate(HPO ₄ ²⁻)	7.2
Hydrogen cyanide (HCN)	Cyanide (CN ⁻)	9.2
Ammonium (NH ₄ ⁺)	Ammonia (NH ₃)	9.3
<i>o</i> -Silicic acid (Si(OH) ₄)	Silicate (SiO(OH) ₃ ⁻)	9.5
Silicate (SiO(OH) ₃ ⁻)	(SiO ₂ (OH) ₃ ²⁻)	12.6
Bisulfide (HS ⁻)	Sulfide (S ²⁻)	~17

Source: Data from Stumm, W. and Morgan, J. J. 1996. *Aquatic Chemistry, 3rd Edition*, John Wiley & Sons, New York. With permission.

extremely low solubility in water. The acid-dissociation constants for the H₂S system are listed together with those of many other important acid/base pairs in Table 17.3. These constants reveal the importance of pH in determining the characteristics and behavior of many compounds that display acid–base behavior in aqueous systems, such as groundwater.

For many of the same reasons that the term “pH” is often used to describe hydrogen-ion activity, the more general term “pC” is often used to describe the activity of a given constituent, where pC is defined as the negative log₁₀ of the activity of that constituent. An important application for this term is in the development of pC–pH diagrams, which are used to illustrate acid–base behavior over a broad range of conditions. These diagrams are constructed based on knowledge of the equilibrium relationships for a given system and the total concentration of the compound(s) which display(s) acid/base behavior.

An example of such a diagram is given for the carbonate system in Figure 17.3 for water exposed to the ambient atmosphere at 25°C, a so-called “open” system. Under the circumstances described in Figure 17.3, the constraint of equilibrium with atmospheric CO₂ dictates that the concentration of H₂CO₃* be fixed at approximately 10^{−5} molar, regardless of pH. At a pH between the first and second dissociation constants (i.e., pK_{a,1} < pH < pK_{a,2}), the majority of the carbonate present in the system will be found in the form of bicarbonate ion (HCO₃⁻). At pH > pK_{a,2} carbonate ion (CO₃²⁻) will represent the predominant species. Under some circumstances, equilibration with atmospheric CO₂ cannot be assumed, as will often be the case with groundwater. Under these circumstances (i.e., a “closed” system), the pC–pH diagram will be defined by the acid/base equilibria presented in Table 17.2 and the total amount of inorganic carbon (dissolved carbonate species) in solution. Figure 17.4 provides an example of the pH dependence of carbonate composition for a “closed” system, wherein exposure of the water to the ambient atmosphere is prevented. Since acid/base equilibria are established rapidly in most cases,¹ pC–pH diagrams represent a powerful tool for evaluating speciation of aqueous constituents that display acid/base behavior. The details of the development of these and other pC–pH diagrams can be found in texts that deal with the subject of aquatic chemistry (e.g., Stumm and Morgan, 1996).

Other aqueous constituents are also subject to pH effects. Principal among these are the metals that generally exist in their oxidized forms in aqueous solutions containing measurable quantities of dissolved

¹Note that for acid/base distributions that rely on equilibration with a gas, the rate of equilibration between the gas and liquid phases will generally be much slower than the rate at which acid/base equilibria are reached in solution. Therefore, the assumption of the applicability of Henry’s law (which defines gas–liquid partitioning at equilibrium), may not always be valid.

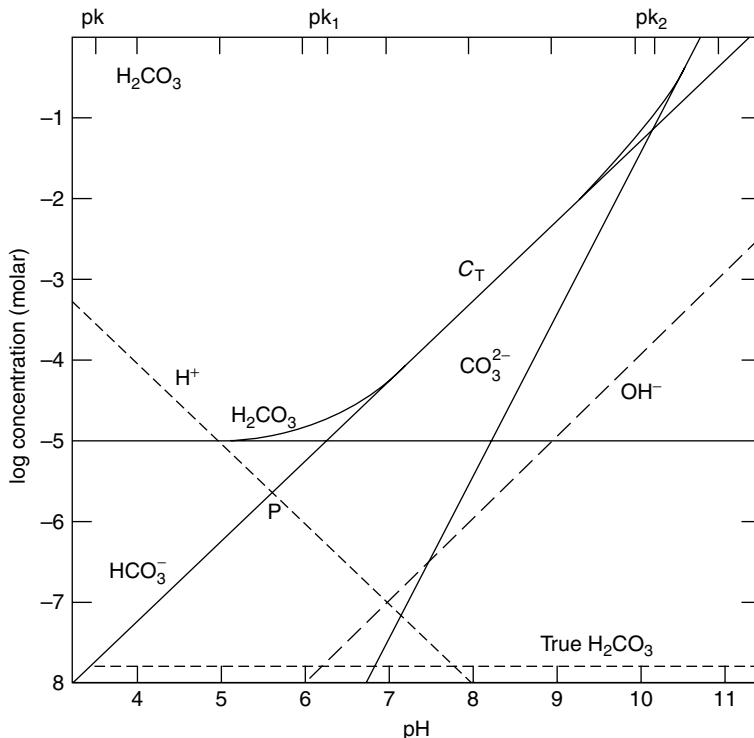


FIGURE 17.3 pC-pH diagram for the carbonate system under conditions of equilibrium with the atmosphere and 25°C. (From Stumm, W. and Morgan, J. J. 1981. *Aquatic Chemistry*, 2nd Edition, John Wiley & Sons, New York. With permission.)

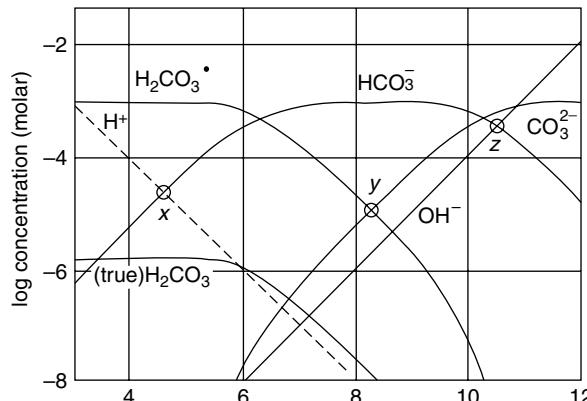


FIGURE 17.4 pC-pH diagram for the carbonate system without atmospheric exposure at 25°C. (From Stumm, W. and Morgan, J. J. 1981. *Aquatic Chemistry*, 2nd Edition, John Wiley & Sons, New York. With permission.)

oxygen. The majority of these compounds exist as polyvalent cations and related complexes. Important examples among background constituents include calcium and magnesium, which are found in the forms of Ca^{2+} and Mg^{2+} , respectively, and related complexes. Ca^{2+} and Mg^{2+} will both form complexes and precipitates with hydroxide, bicarbonate, and carbonate. These anionic constituents are referred to as ligands in the complexation and precipitations reactions. As the distribution of all three of these commonly found ligands are dependent on pH, the aqueous solubility of calcium and magnesium are also closely linked to pH. Similar logic can be used to describe the pH-dependent solubility behavior of many other

metals in solution. A detailed discussion of this topic is beyond the scope of this text, but good summaries can be found in textbooks relating to the subject of aquatic chemistry. An excellent compilation of the equilibrium constants for complex formation and precipitation can be found in the reference series by Martell and Smith (1974–1989).

As in the case of acid–base reactions, the reactions that result in formation of aqueous complexes in a homogeneous system (i.e., systems comprised only of an aqueous phase) are so rapid that in most circumstances equilibrium distributions can be assumed. However, the formation of precipitates and the reverse reaction (dissolution) involve at least two phases and equilibrium conditions do not always apply. Several factors contribute to the sometimes slow rates with which equilibria are achieved in these multiphase systems including transport limitations between phases and slow rates of crystallization among solids. Therefore, equilibrium calculations of solubility based on the presence of precipitates are not always valid and generally result in errors when compared to actual conditions.

The pC–pH diagram can also be used as a tool for evaluation of precipitate formation. One example of how pC–pH diagrams may be used for this purpose is the so-called predominance area diagram, which provides an indication of the conditions under which a particular solid phase (precipitate) is expected to predominate. These diagrams can be used to predict the form of a solid and the limits of solubility under aqueous conditions that will allow a metal to form precipitates with more than one ligand. An example of such a diagram is presented in Figure 17.5 for reduced iron, or Fe(II) at a total system concentration (C_t) of $10^{-5} M$. Fe(II) will form precipitates in the presence of both hydroxide and carbonate. As both ligands are likely to be present in groundwater systems, it is sometimes important to evaluate their combined effects on metal solubility. At pH < 10.5, the majority of precipitated Fe(II) in this example is expected to be present in the form of $\text{FeCO}_3(s)$. As pH increases above this value, more hydroxide ion becomes available and $\text{Fe(OH)}_2(s)$ is expected to predominate. Under either condition, it is likely that both forms of Fe(II) will be present, but one form will be expected to predominate. Figure 17.5 also defines the concentration of Fe(II) and related complexes that would be expected to be present in solution, in equilibrium with the solid phases, as a function of pH. The limit of solubility is defined by the lower borders of the solid regions within the predominance area diagram.

The characteristics and behavior of solids in solution are strongly related to pH; in particular, the net surface charge displayed by a particle in solution is largely determined by its composition and the solution

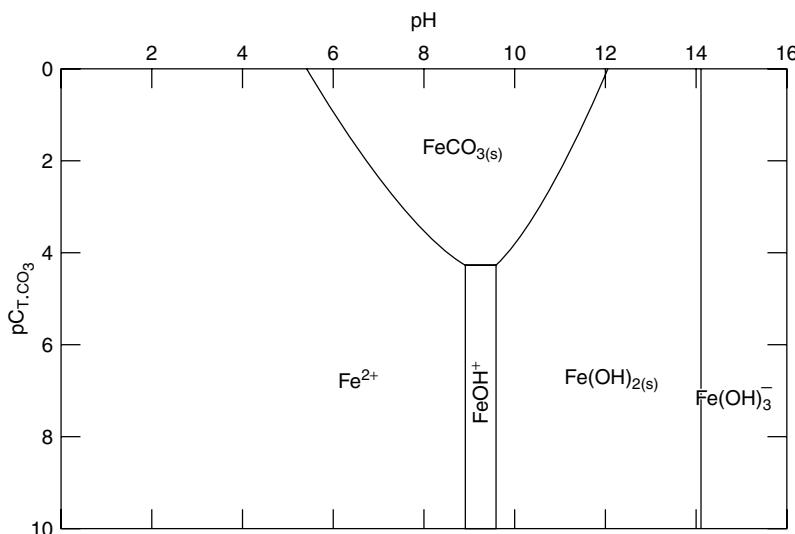


FIGURE 17.5 Predominance-area diagram for Fe(II) under conditions of exposure to carbonate and hydroxide, and with a total iron concentration of $10^{-5} M$. (From Snoeyink, V. L. and Jenkins, D. 1980. *Water Chemistry*, John Wiley & Sons, New York. With permission.)

pH. At least three mechanisms exist whereby solids generally express a surface charge in solutions. The first involves surface groups that display many of the same characteristics as dissolved constituents. For example, silica-based particles (e.g., sand) will have silanol groups; silanol groups display amphoteric behavior in aqueous solution:



The charge expressed by this surface group will be determined by the availability of protons to adsorb to the site. Therefore, the charge expressed by this surface group will depend on the pH of the solution to which it is exposed. Many other surface groups on particles express amphotericism; furthermore, this behavior is not limited to inorganic surface groups.

A similar mechanism by which particles obtain surface charge is the adsorption of ionic constituents. The similarity stems from the fact that the behavior of amphoteric groups may be viewed as the equivalent of adsorption or desorption of protons. Similarly, other ionic constituents generally adsorb to these sites resulting in a corresponding change of the surface charge.

A third mechanism by which particles express surface charge is isomorphous replacement of metals within crystal structures. For example, silica sand is commonly assumed to have a monomeric structure of SiO_2 ; however, imperfections in this structure, perhaps by the inclusion of metals with valence which is different than Si, results in a net surface charge.

Collectively, these processes allow particles in contact with liquid water to express a net surface charge. This charge is balanced by the migration of ions in solution to produce a region of local charge imbalance within the solution immediately adjacent to the particle : water interface. This region is often referred to as the electrical double-layer and plays an important role in governing the behavior of many constituents in groundwater systems.

The effect of pH in determining surface charge is profound. Titrimetric procedures exist whereby changes in surface charge may be quantified as a function of pH. These procedures allow for the determination of the zero point of charge (zpc), or the pH at which net surface charge is zero. A compilation of zpc values for some common particles is found in Table 17.4. At a pH above the zpc in a system where protons are the only adsorbable ion, the surface will express a net negative charge, whereas at a pH below the zpc, a net positive surface charge is expressed. From the values presented Table 17.4, it is evident that at near-neutral pH, most particles in solution express a net negative surface charge; however, important exceptions to this generalization can be found. This fact is important in that particles with like charges will express a repulsive electrostatic force, whereas those with opposite surface charge will experience

TABLE 17.4 Zero Points of Charge (pH_{zpc}) for Selected Inorganic Solids

Compound	pH_{zpc}
SiO_2	2.0
Montmorillonite	2.5
$\delta\text{-MnO}_2$	2.8
Kaolinite	4.6
$\alpha\text{-Al(OH)}_3$	5.0
Fe_3O_4	6.5
$\alpha\text{-FeOOH}$	7.8
CuO	9.5

Source: Data from Stumm, W. and Morgan, J. J. 1981. *Aquatic Chemistry, 2nd Edition*, John Wiley & Sons, New York. With permission.

an attractive electrostatic force. The magnitude of the electrostatic forces between particles is a strong function of separation distance. Particles in suspension will experience other forces that also are strong functions of separation distance (e.g., London and van der Waal's forces). The sum of all such forces will govern the stability of particulate suspensions.

Surface charges expressed by particles and soil media must be considered in an evaluation of the fate and transport of ionic and particulate contaminants. The fate of particles in soils has been the subject of considerable numerical and experimental research dealing with the process of sand filtration. The results of these research efforts have generated numerical models that can accurately represent the movement of particles in soil media under some conditions. Those interested in modeling particle transport in groundwater systems should refer to the text by Clark (1996), which provides an excellent critical review of the pertinent literature on the subject.

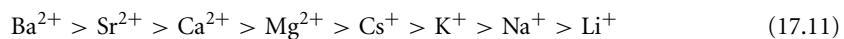
17.3.2 Sorption to Solids

In the general case, compounds present in groundwater systems will have the opportunity to associate with the aqueous phase and the solid phase(s). The partitioning of a constituent among these phases will depend on the characteristics of the constituent and the phases themselves.

Two main categories of association among aqueous and solid phases exist. The first involves the accumulation of material at a water:solid interface and is termed adsorption. The second involves the intermingling of solute molecules with the molecules of the solid phase and is referred to as absorption. Absorption may be viewed as the dissolution of an aqueous constituent in a solid solvent. The distinction between the two processes is often vague and in a large number of cases, such a distinction is nearly impossible to make. Furthermore, in many cases both adsorption and absorption play a significant role in determining the partitioning behavior of groundwater constituents among the solid and liquid phases. For purposes of this chapter, these terms will be collectively referred to as sorption.

Three general types of sorption can be described for groundwater systems. All three involve the formation of bonds between the constituent and the solid phase. The first type, physisorption, involves the formation of relatively weak physical bonds such as those attributable to London and van der Waal's forces. Under some conditions, the interactions between a solute and a solid surface may be more accurately described as a chemical bond, with the process being referred to as chemisorption. In chemisorption, the bond strength between a solute and the solid surface may approach that of a covalent bond. The third type of interaction involves electrostatic interactions between surface groups and the solute; this type of surface:solution interaction plays a crucial role in ion exchange.

Physisorption and ion exchange (which may be viewed as a special case of physisorption) tend to be readily reversible processes, as the forces that are responsible for the attraction of a solute to the solid are relatively weak. In the case of ion exchange, the factors that govern the affinity of a solute for the solid phase are comparatively well-defined. For ions of similar size, the valence of the ion will govern the magnitude of the coulombic force between the surface and solute; ions of higher valence will show a greater affinity for the solid surface. For ions of similar valence, the hydrated radius of the ion will dictate solid affinity, with ions having small hydrated radii showing preference over those with large radii. Interestingly, hydrated radius is inversely related to (unhydrated) ionic radius for elements of the same family on the periodic table. Therefore, ion exchange selectivity order may be defined for a particular solid medium. A representative example of a selectivity order is presented below for some common cations:



This order implies that at similar concentrations, ions will replace those ions that lie to their right in the selectivity order. As an example, Ca^{2+} will replace Na^+ under most conditions. The specific selectivity order displayed by a soil medium may vary from that defined in Equation 17.11 due to situation-specific circumstances. Furthermore, the reader should recognize that when solute concentrations are not similar,

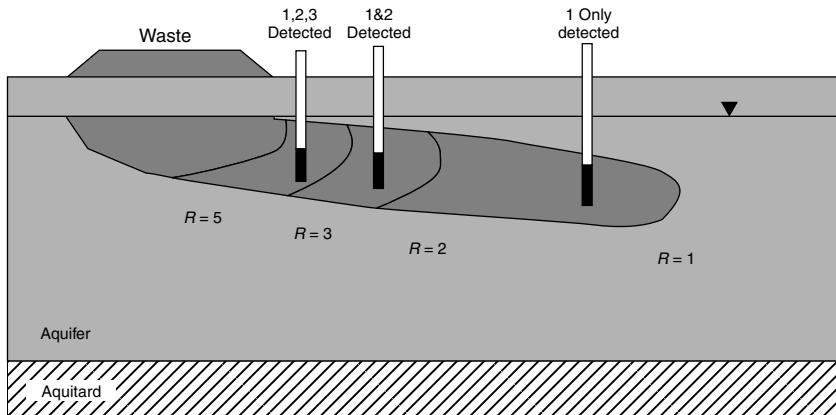


FIGURE 17.6 Schematic representation of the behavior of a mixture of pollutants (compounds 1, 2, 3, and 5) in a groundwater system. All pollutants are assumed to be introduced from the same spill and are assumed to be present initially in roughly equal amounts. The affinity of the constituents for the soil matrix is defined by the order $5 > 3 > 2 > 1$. (From US Environmental Protection Agency. 1989. *Transport and Fate of Contaminants in the Subsurface*, EPA/625/4-89/019. With permission.)

the tendency of ions with relatively high solution-phase concentration to replace ions associated with the solid phase will increase.

The importance of sorptive processes in determining the fate of groundwater contaminants may be illustrated by an example. Imagine a chemical spill comprised of four constituents: 1, 2, 3, and 5, all at similar concentrations. Furthermore, we will assume that all four constituents compete for the same sorption sites within the soil matrix and that the affinity of the constituents is defined by $5 > 3 > 2 > 1$. All four compounds will move through the soil medium by advection, but their movement will be retarded by interactions with soil surfaces. As constituent 5 shows the highest affinity for the soil matrix, it will be able to replace constituents 3, 2, and 1. Therefore, on average, constituent 5 will spend more time in association with the soil matrix than constituents 3, 2, or 1, and will therefore move through the matrix more slowly. A schematic representation of the movement of these constituents through this hypothetical system is provided in Figure 17.6. The separation of constituents 1, 2, 3, and 5 based on their affinity for the stationary soil matrix is referred to as chromatography. This process is used on a much smaller scale and under more controlled conditions in the family of chromatographic analytical procedures, which are extremely important in the analysis of groundwater contaminants. Important analytical techniques that employ these principles include gas chromatography (GC), liquid chromatography (LC), and ion chromatography (IC). Many variations on these processes exist.

17.3.3 $p\epsilon$

Much in the same way that pH is used as a master variable in describing the availability of protons, $p\epsilon$ is a master variable that is used to describe the availability of electrons for reaction. The analogy may also be carried further in that much as H^+ does not exist in solution, free electrons do not exist in solution. The mathematical definition of $p\epsilon$ is as follows:

$$p\epsilon = -\log_{10}(e^-) \quad (17.12)$$

In aqueous systems, $p\epsilon$ provides an indication of the tendency for electrons to be transferred. Under conditions of low $p\epsilon$, the activity of aqueous electrons is high, and there is significant pressure being exerted on aqueous constituents to undergo reduction. Conversely, under conditions of high $p\epsilon$, electron activity is low and the tendency is for oxidation reactions to be promoted. Oxidation and reduction

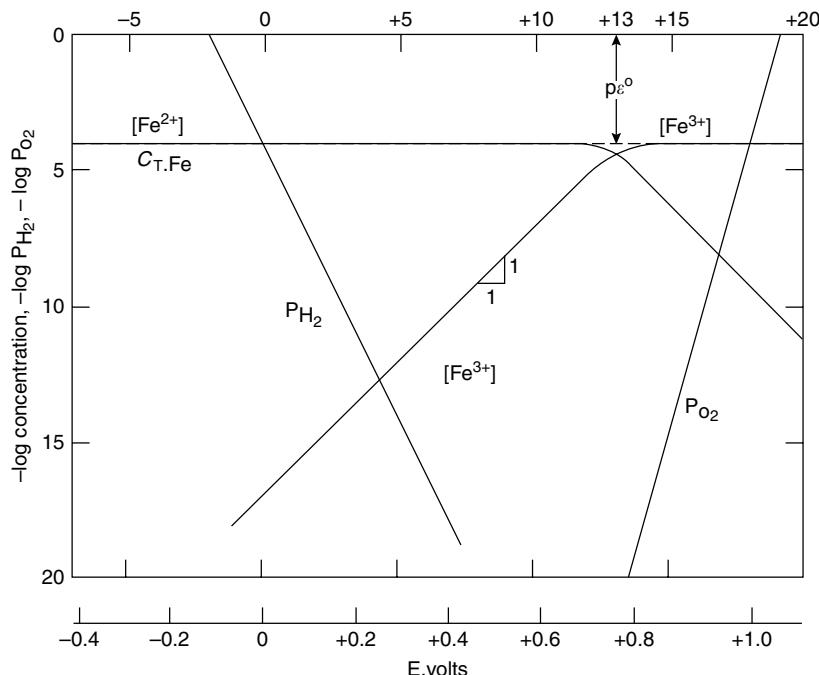


FIGURE 17.7 $p\epsilon$ - pC diagram for iron at a total dissolved iron concentration of $10^{-4} M$ and $pH = 2$. (From Snoeyink, V. L. and Jenkins, D. 1980. *Water Chemistry*, John Wiley & Sons, New York. With permission.)

reactions cannot take place in isolation (free electrons do not remain in solution for any significant period of time). Therefore, oxidation and reduction reactions must be coupled.

As in the case of pH, graphical tools exist to describe equilibrium conditions using $p\epsilon$ as a master variable. One such tool is the so-called $p\epsilon$ - pC diagram. Figure 17.7 provides an example of such a diagram for a system containing dissolved iron at a total iron concentration of $10^{-4} M$ and a pH of 2. Under conditions of low $p\epsilon$, the system will be characterized as a reducing environment. In the case of the Fe(II)/Fe(III) system, this suggests that Fe(II) will be the dominant form. When $p\epsilon$ increases, a more oxidizing environment will exist. For the iron system at pH = 2 and at $p\epsilon > 13$, Fe(III) is expected to be the predominant form.

The $p\epsilon$ - pC diagram is conceptually similar to the pC -pH diagram for description of acid/base equilibria. However, the $p\epsilon$ - pC diagram is somewhat limited in its ability to present these equilibria in the sense that only a single pH can be represented. Another graphical tool for presentation of redox equilibria is the $p\epsilon$ -pH diagram. Figure 17.8 provides an example of such a graph for the sulfur system at a total dissolved sulfur concentration of $10^{-2} M$ (Stumm and Morgan, 1981). This diagram provides an illustration of the conditions under which given sulfur species are expected to predominate. Though many similarities exist between pH and $p\epsilon$ and the equilibria they can be used to describe, a significant difference between equilibrium calculations based on acid/base transfer reactions and those corresponding to oxidation/reduction reactions is that whereas equilibrium conditions can often be assumed to apply for reactions involving proton transfer (i.e., acid/base reactions), the same cannot be said for reactions which involve electron transfer (i.e., oxidation/reduction reactions). Therefore, equilibrium calculations for redox reactions will not always yield an accurate representation of actual conditions in an aqueous system; however, these calculations do represent a powerful tool for defining the conditions one would expect to find if an infinite amount of time were allowed for the system to stabilize.

Compilations of aqueous redox chemical equilibria can be found in the literature. An excellent summary of the methods used to conduct redox equilibrium calculations and graphical representations of these equilibria can be found in Stumm and Morgan (1996).

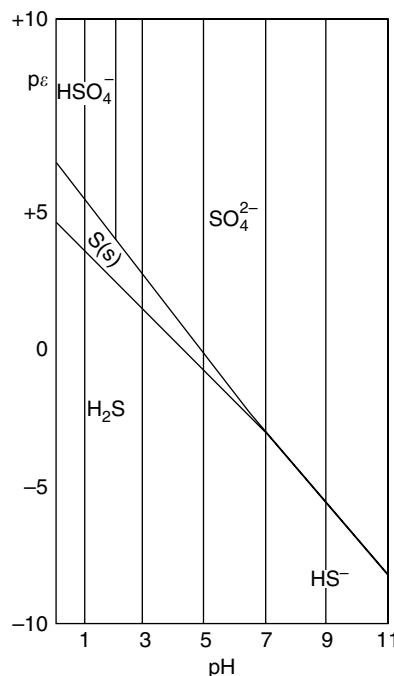


FIGURE 17.8 $p\epsilon$ -pH diagram for sulfur at a total sulfur concentration of 10^{-2} M . (From Stumm, W. and Morgan J. J. 1981. *Aquatic Chemistry, 2nd Edition*, John Wiley & Sons, New York. With permission.)

17.3.4 Temperature

Thermodynamic considerations play an important role in evaluating the rates and equilibria of chemical reactions. In general, reaction rates will increase with temperature due to the increased availability of thermal energy. This energy will promote collisions between molecules by increasing rates of diffusion and will reduce energy barriers to reaction progress. Equilibrium conditions are achieved in a system when the rates of forward and reverse reactions are equal. If a given reaction is endothermic (i.e., heat is “consumed” by the reaction), then an increase in temperature would be expected to promote the forward reaction and shift the equilibrium to the right (an increase in “product” formation). Examples of the effects of heat on groundwater contaminants include changes in the equilibria and kinetics of precipitation and dissolution of minerals, and changes in the rates by which biological transformation processes proceed.

In most groundwater situations involving *in situ* contamination, temperature can often be assumed to be constant at a value of 1–2°C greater than the annual average ambient air temperature for the region (Freeze and Cherry, 1979). For example, the groundwater temperature in the midwestern United States (at depths greater than approximately 10 m, where diurnal and seasonal variations in air temperature have minimal effect) can generally be assumed to be approximately 13°C. Under these conditions, the role of temperature in governing contaminant transport and behavior is more of a buffer than a variable. However, those who deal with groundwater contaminants should be aware of the potential for thermal effects to influence system behavior. For example, shallow contamination in areas subject to substantial variations in climate (i.e., within the frost zone) generally display variations in behavior as a function of the season. Similarly, other thermal sources and sinks within groundwater systems may play a role in determining the local behavior of contaminants. And finally, many remediation strategies call for removal of contaminants for *ex situ* treatment. Contaminants exposed to these conditions will obviously be subjected to a less stable thermal environment than those that remain *in situ*. Consequently, the behavior of groundwater contaminants may be different in the *ex situ* environment than *in situ*.

17.4 Inorganic Constituents

The compounds that comprise the inorganic fraction of groundwaters are heavily influenced by interactions between source water and soils. As source water moves through a soil, chemical reactions taking place at the soil : water interface will change the composition of both phases. In the previous section, the basic characteristics of these reactions were reviewed. Based on the working definition of “groundwater contaminants” provided in the introduction, the inorganic constituents of primary interest are those that have been introduced to a groundwater system as a result of human activity. Therefore, it is necessary to define background conditions in an unpolluted groundwater system to understand the importance of these reactions relative to inorganic groundwater contaminants.

17.4.1 Source Water Composition

The accumulation of atmospheric moisture as a result of evapotranspiration may be viewed as a large-scale distillation process. However, the composition of precipitation, in the form of rain or snow, is not accurately represented by the composition of distilled water one might find in a laboratory. In particular, rain and snow will react with atmospheric constituents to determine their composition prior to reaching the earth’s surface. A summary of measured compositions in rain and snow is provided in Table 17.5.

Many atmospheric gases display substantial solubility in water. The typical composition of gases in a “clear” atmosphere (which also must be defined in relative terms as a result of human activities) is summarized in Table 17.6. Normally, atmospheric precipitation can be assumed to be at equilibrium

TABLE 17.5 Measurements of the Chemical Composition of Rain and Snow

Constituent	Location				
	1	2	3	4	5
SiO ₂	0.0		1.2	0.3	
Al (III)	0.01				
Ca ²⁺	0.0	0.65	1.2	0.8	3.3
Mg ²⁺	0.2	0.14	0.7	1.2	0.36
Na ⁺	0.6	0.56	0.0	9.4	0.97
K ⁺	0.6	0.11	0.0	0.0	0.23
NH ₄ ⁺	0.0				0.42
HCO ₃ ⁻	3		7	4	0.0
SO ₄ ²⁻	1.6	2.18	0.7	7.6	6.1
Cl ⁻	0.2	0.57	0.8	17	2.0
NO ₂ ⁻	0.02		0.0	0.02	
NC ₃ ⁻	0.1	0.62	0.2	0.0	2.2
Total dissolved solids	4.8		8.2	38	
pH	5.6		6.4	5.5	4.4

Note: All values are presented in units of mg/L, except pH.

1. Snow, Spooner Summit, U.S. Highway 50, Nevada (east of Lake Tahoe), altitude 7100 ft, Nov. 20, 1958. J.H. Feth, S.M. Rogers, and C.E. Roberson, *Chemical Composition of Snow in the Northern Sierra Nevada and Other Areas*. U.S. Geological Survey Water Supply Paper 1535J, 1964, 39 pp.

2. Average composition of rain from August 1962 to July 1963 at 27 points in North Carolina and Virginia. A.W. Gambell and D.W. Fisher, *Chemical Composition of Rainfall, Eastern N. Carolina and Southeastern Virginia*: U.S. Geological Survey Water Supply Paper 1535K, 1964, 41 pp.

3. and 4. Rain, Menlo Park, CA, 7:00 P.M. Jan. 9 to 8:00 A.M. Jan. 10, 1958. Whitehead, and J.H. Feth, *Chemical Composition of Rain, Dry Fallout, and Bulk Precipitation at Menlo Park, CA, 1957–1959*, *J. Geophys. Res.*, 69:3319–3333 (1964).

5. Station 526U, Belgium, European Atmospheric Chemistry Network. Average of 180 samples. L. Granat, On the Relation Between pH and the Chemical Composition in Atmospheric Precipitation, *Tellus*, 24, 550–556 (1972).

Source: From Snoeyink, V. L. and Jenkins, D. 1980. *Water Chemistry*, John Wiley & Sons, New York. With permission.

with atmospheric gases, as defined by Henry's law. Of the constituents listed in Table 17.6, several play significant roles in determining the chemistry of rain and snow. Primary among these is carbon dioxide (CO_2). As described previously, CO_2 will dissolve in water to form carbonic acid, which can dissociate to form carbonic acid (H_2CO_3), bicarbonate (HCO_3^-), and carbonate (CO_3^{2-}) ions. The carbonate system plays an important role in governing groundwater composition through precipitation and complexation reactions. The carbonic acid cycle also plays a major role in determining the pH and buffering capacity of precipitation. Specifically, water at equilibrium with atmospheric CO_2 will have a pH of approximately 5.7. This value is often used as benchmark in the evaluation of the effects of atmospheric constituents that participate in the formation of acidic precipitation — precipitation with $\text{pH} < 5.7$ is often assumed to be under the influence of strong acids.

The influence of atmospheric composition on groundwater chemistry goes far beyond CO_2 . Examples of other atmospheric constituents that play important roles in the subsurface environment are provided below.

Oxygen (O_2) will dissolve in water to yield a liquid-phase concentration of approximately 8 to 10 mg/L, depending on altitude and temperature. In a groundwater system, the availability of oxygen (and other constituents) will play a major role in determining the extent and type (i.e., aerobic, anoxic, or anaerobic) of microbiological activity that can take place. Microbiologically mediated chemical transformations can play a significant role in determining groundwater composition; this fact is exploited in many soil and groundwater remediation strategies (see Chapter 31 and Chapter 36).

The oxides of nitrogen and sulfur (NO_x and SO_x), which are released to the atmosphere as a part of many natural processes (e.g., volcanic activity, microbiological activity), can react in the atmosphere to yield strong acids, such as HNO_3 and H_2SO_4 . In an unpolluted atmosphere, the concentrations of these constituents are relatively low and do not play an important role in determining precipitation chemistry.

TABLE 17.6 Typical Gas-Phase Composition in a “Clean” Atmosphere
(Based on a Total Pressure of 1.0 atm)

Gas	Partial Pressure (atm)
N_2	0.781
O_2	0.209
H_2O	0.028
Ar	0.0093
CO_2	0.0003
Ne	1.8×10^{-5}
He	5.2×10^{-6}
CH_4	1.5×10^{-6}
Kr	1.1×10^{-6}
CO	$(0.6-1) \times 10^{-6}$
SO_2	1×10^{-6}
N_2O	5×10^{-7}
H_2	5×10^{-7}
O_3	$(0.1-1.0) \times 10^{-7}$
Xe	8.7×10^{-8}
NO_2	$(0.05-2) \times 10^{-8}$
Rn	6×10^{-20}

Source: B. A. Mirtov, Gaseous composition of the atmosphere and its analysis. *Akad. Nauk. SSSR, Inst. Prikl. Geofiz. Moskva* (translated by the Israel Program for Scientific Translations, published in Washington, US Dept. of Commerce, Office of Technical Services, 1961, 209 pp. With permission.).

However, combustion of fossil fuels in mobile and stationary sources can result in large increases in atmospheric NO_x and SO_x concentration. The spatial scale on which these combustion processes are carried out allows them to influence rain and snow chemistry over geographic areas that span international boundaries.

Radon (Rn) is a noble gas, which means that its electron structure is such that it tends to be quite inert with respect to conventional chemical reactions. However, radon has several naturally occurring isotopes, some of which have unstable nuclear structures, and therefore express radioactivity. In the natural atmosphere, the concentration of Rn is quite low and does not play a significant role in determining precipitation chemistry.

Groundwater composition will not be strictly governed by the composition of atmospheric gases; the subsurface generally contains gases at partial pressures that are markedly different from those of the atmosphere. For example, locally intense microbiological activity can substantially alter the partial pressures of gases such as CO_2 and O_2 . The reader should be aware of the potential for subsurface conditions to change gas-phase composition, and the resulting composition of groundwater with which it comes into contact. However, atmospheric composition provides a reference point in the analysis of groundwater composition and contamination.

Rain and snow are also effective scavengers of atmospheric particulate matter. Naturally occurring dust particles and aerosols will contribute to the inorganic composition of rain and snow. These particles are probably responsible for the presence of metals such as sodium (Na), potassium (K), calcium (Ca), and magnesium (Mg) in precipitation.

In a “polluted” atmosphere, the concentrations of gaseous and particulate constituents may be markedly different from those of a clean atmosphere. For comparison, Table 17.7 provides a listing of atmospheric composition in a clean atmosphere and in a polluted atmosphere, based on the trace constituents. It is interesting to note that most of these compounds are formed and can be found in unpolluted air, but the concentrations of many oxidizing compounds (e.g., ozone), acid-forming gases (e.g., oxides of nitrogen and sulfur), and particles can be substantially higher in a polluted atmosphere than in an unpolluted atmosphere.

The nature of atmospheric contamination will be specific to an area. Therefore, the composition of a polluted atmosphere cannot be quantitatively generalized and local information is needed. However, for a given atmospheric composition, increases in atmospheric constituent concentrations attributable to anthropogenic activities can be expected to translate to parallel changes in rainwater and snow chemistry

TABLE 17.7 Typical Concentration Ranges of Trace Atmospheric Constituents in Clean and Polluted Atmospheres

Species (formula)	Concentration (ppbv)	
	Clean Troposphere	Polluted Troposphere
Sulfur Dioxide (SO_2)	1–10	20–200
Carbon Monoxide (CO)	120	1,000–10,000
Nitric Oxide (NO)	0.01–0.05	50–750
Nitrous Oxide (N_2O)	0.1–0.5	50–250
Ozone (O_3)	20–80	100–500
Nitric Acid (HNO_3)	0.02–0.3	3–50
Ammonia (NH_3)	1	10–25
Formaldehyde (HCHO)	0.4	20–50
Formic Acid (HCOOH)		1–10
Nitrous Acid (HNO_2)	0.001	1–8
Peroxyacetyl Nitrate ($\text{CH}_3\text{C}(\text{O})\text{O}_2\text{NO}_2$)		5–35
Non-Methane Hydrocarbons	500–1200	

Source: From Seinfeld, J. H. 1986. *Atmospheric Chemistry and Physics of Air Pollution*, John Wiley & Sons, New York. With permission.

in downwind areas. Those who work in the groundwater arena should be aware of the link between groundwater behavior and the surrounding environment, including the atmosphere.

17.4.2 Source Water: Soil Interactions

Percolation of rainwater and snow into a soil formation is responsible for the existence of groundwater. As water moves through soil, a complex set of reactions take place that alter the chemical composition of both phases (solid and liquid). The general categories of reactions have already been reviewed: acid/base, complexation, precipitation/dissolution, and oxidation/reduction. The local implications of these reactions will depend on the specific compositions of the solid and liquid phases that come into contact. For example, many geologic formations consist largely of calcium- and magnesium-based minerals, such as limestone. Dissolved CO₂ (and other acids) will interact with these minerals to promote equilibration between the phases. This results in mineral dissolution and an increase in the liquid-phase concentration of the metals that make up the dissolved minerals. In groundwater systems, the presence of Ca²⁺ and Mg²⁺ are important not only in terms of the dietary needs they serve to people who consume water from a groundwater source, but also in terms of industrial applications. Chemical hardness, defined as the sum of the concentrations of all polyvalent cations in an aqueous system, is almost exclusively attributable to Ca²⁺ and Mg²⁺ in most groundwaters. The constituents of chemical hardness cause problems of fouling and scale formation in industrial systems, and form stable complexes with detergents, thereby increasing detergent usage. Mineral dissolution will also lead to increases in the concentrations of inorganic constituents.

A summary of the major, minor, and trace constituents, according to the designations defined by Freeze and Cherry (1979), in groundwater is provided in Table 17.8. As can be seen from this table, the ranges of constituent concentrations are quite broad. Not surprisingly, groundwater composition is highly variable between sites, and is largely determined by the differences in geology among the sites.

Inorganic groundwater contaminants may come in the form of increases in the concentrations of the constituents listed in Table 17.8 above their naturally occurring concentrations, or the introduction of constituents that otherwise would not be detectable in the system at all. A number of factors can contribute to the changes in solution chemistry in these systems, including changes in atmospheric composition, changes in soil chemistry, and human activities that directly alter the physical or chemical conditions of a soil matrix.

The introduction of strong acids (e.g., H₂SO₄ and HNO₃) to groundwater systems from atmospheric sources can have a profound effect on groundwater composition. These acids reduce the pH of source waters (rain and snow) and promote dissolution of the soil matrix. In contaminated soils, dissolution is not limited to naturally occurring minerals. For example, many military sites have extensive heavy-metal contamination in their soils as a result of firing range activities. The heavy metals at these sites are often dominated by lead (Pb), but other elements are also present including zinc (Zn), copper (Cu), and antimony (Sb). The solubility of these metals in their zero-valent states is quite limited. Therefore, for these metals to represent a potential source of groundwater contaminants, they must generally undergo oxidation. A number of mechanisms exist whereby these metals may be oxidized in a soil including microbiological activity, contact with chemical oxidants, and galvanic interactions with other metals. As a result of these and other processes, zero-valent metals are transformed, usually at slow rates, to their more mobile oxidized forms, where dissolution by aqueous constituents can become important.

Among heavy metals, lead has generally been the subject of the most extensive and comprehensive research, probably because of the large number of potential sources and its general ubiquity. Other examples of potential sources of lead in soils and groundwater include lead-based paints and leakage of lead-based fuels from underground storage tanks. While both of these products are now out of production in the United States, important reservoirs of contamination (or at least potential contamination) still exist. Furthermore, these compounds are still being produced in many areas of the world.

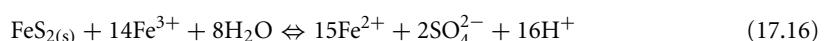
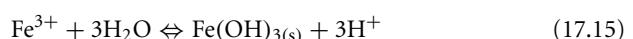
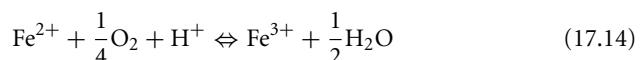
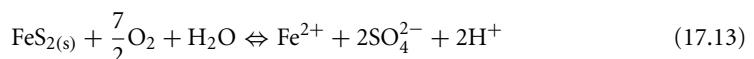
Another source of inorganic contamination is mining and related activities. In the past, a common mining technique involved the application of liquids to promote selective chemical extraction of valuable mineral constituents. These practices that are no longer in use *in situ*, sometimes created significant local

TABLE 17.8 Major, Minor, and Trace Constituents as defined by Freeze and Cherry (1979)

<i>Major Constituents (> 5 mg/L)</i>	
Bicarbonate	Magnesium
Calcium	Silicon
Chloride	Sodium
Carbonic acid	Sulfate
<i>Minor Constituents (0.01–10 mg/L)</i>	
Boron	Nitrate
Carbonate	Potassium
Fluoride	Strontium
Iron	
<i>Trace Constituents (< 0.1 mg/L)</i>	
Aluminum	Molybdenum
Antimony	Nickel
Arsenic	Niobium
Barium	Phosphate
Beryllium	Platinum
Bismuth	Radium
Bromide	Rubidium
Cadmium	Ruthenium
Cerium	Scandium
Cesium	Selenium
Chromium	Silver
Cobalt	Thallium
Copper	Thorium
Gallium	Tin
Germanium	Titanium
Gold	Tungsten
Indium	Uranium
Iodide	Vanadium
Lanthanum	Ytterbium
Lead	Yttrium
Lithium	Zinc
Manganese	Zirconium

soil and groundwater contamination problems, some of which are quite long-lasting, and still exist today. Modern mining techniques are substantially improved in terms of their potential to contaminate, but are not exempt from the ability to cause groundwater problems.

The presence of mineral acidity in the form of elevated proton activity (low-pH) leads to the dissolution of many precipitates, including those that involve hydroxide, carbonate, sulfide, and phosphate ligands. The mineral composition of many soils incorporates large amounts of these materials, which themselves act as buffers against changes in pH. Therefore, the presence of mineral acidity in soils and groundwater is somewhat unusual. However, processes do exist that can create locally low pH. One prominent example of such a process is the oxidation of pyrite (FeS_2) and other iron sulfide crystalline forms. As will be discussed below, the process is often linked to mining activities, and the resulting low-pH waters from these operations are commonly referred to as “acid mine drainage.” The stoichiometry of the process is described as follows:



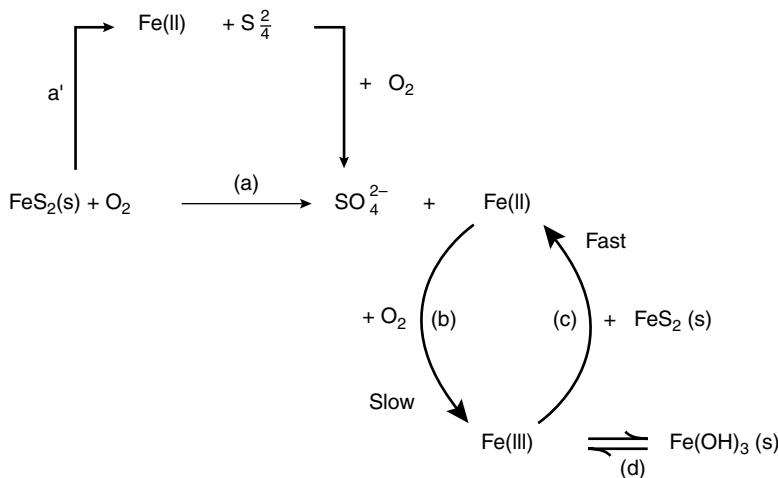


FIGURE 17.9 Mechanistic description of the processes responsible for pyrite oxidation, as hypothesized by Temple and Delchamps (1953).

Reactions 17.13 through 17.16 do not accurately represent the mechanism by which pyrite oxidation leads to acidification. Temple and Delchamps (1953) hypothesized the mechanism illustrated in Figure 17.9 to describe the process and its rate limitations. Under the low-pH conditions that would be expected in waters subject to these reactions, the rate of Fe^{2+} oxidation is extremely slow. However, several species of bacteria are known to catalyze this reaction, including *Thiobacillus thiooxidans*, *Thiobacillus ferrooxidans*, and *Ferrobacillus ferrooxidans*. These microorganisms, which are categorized as microaerophiles, require oxygen to catalyze Reaction 17.14, though they can survive under conditions of low oxygen concentration. The introduction of oxygen, often as a result of mining activities, initiates the process by dramatically increasing the rate of reaction (Equation 17.14). Once the process is started, the presence of oxygen is no longer a requirement for the development of mineral acidity. Therefore, sealing of mines that have experienced this problem does not represent an acceptable remediation approach.

The generation of acid mine drainage is often associated with coal mining activities as pyrite-bearing minerals are commonly found in conjunction with coal deposits. However, pyrite is also found in other mineral deposits, and similar acidification problems have been linked to acid mine drainage within copper and uranium mining facilities.

Another source of inorganic contamination is storage of salt for municipal operations, such as road deicing. Large salt piles, largely comprised of sodium chloride (NaCl), are often stored with little or no weather protection. Given the solubility of NaCl and related compounds in water, these situations provide a tremendous potential for salt introduction to a groundwater system. While the constituents of road salts do not generally represent an acute toxicity problem toward humans, they can increase salt concentrations to the point where a groundwater is not fit for human consumption. Additionally, the removal of sodium and potassium salts from groundwaters is difficult because they are so soluble in water. Their affinity for water makes them difficult and expensive to separate them from the aqueous phase.

17.5 Organic Constituents

17.5.1 Background Organics

In uncontaminated soil/groundwater systems, the predominant organic constituents are humic substances. These compounds, which represent the products of natural decay processes from plant and animal matter, do not have a well-defined chemical structure. Therefore, humic substances are identified

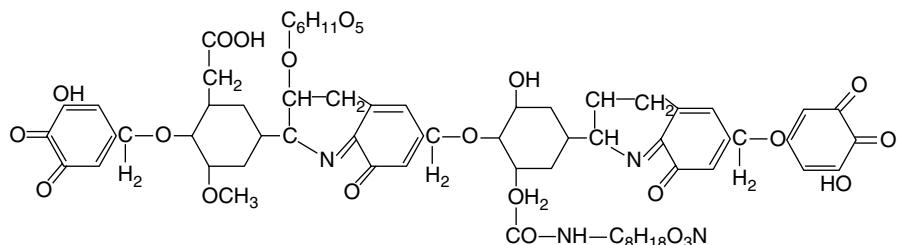


FIGURE 17.10 Representation of the structure of humic acid. (From Schnitzer, M. and Khan, S. U. 1972. *Humic Substances in the Environment*, Marcel Dekker, New York. With permission.)

and differentiated according to a broad set of operational definitions. Three classes of humic substances exist:

1. Fulvic acid — constituents that are soluble at any pH.
2. Humic acid — constituents that are insoluble at low-pH, but otherwise soluble.
3. Humin — constituents that are insoluble in water at any pH.

In most groundwaters, the concentrations of humic substances (humic and fulvic acids) are quite low, normally less than several milligrams per liter.

As humic substances are operationally defined, no unique structural representation of these compounds can be defined. Humic substances generally take the form of large heteroatomic organic polymers. The composition and structure of humic substances will differ depending on the source material and conditions under which they are generated. However, some generalizations have been developed regarding the functional groups that tend to exist within humic substances. For example, the structure presented in Figure 17.10 has been hypothesized to be representative of a monomer of a humic acid (Schnitzer and Khan, 1972). One of the most significant features of these hypothesized structural representations is the presence of oxygenated functional groups. These groups are believed to be largely responsible for the ability of humic substances to complex metals.

Many classes of synthetic organic compounds have been identified as groundwater contaminants, including pesticides, petroleum products, and (halogenated) solvents. Though it is not practical to list all possible organic groundwater contaminants, a number of their important characteristics can be illustrated with examples from the aforementioned categories.

The mobility of groundwater contaminants is largely determined by their affinity for soil particles. Generally speaking, contaminants with high affinity for soil particles preferentially associate with the soil matrix, and therefore move relatively slowly (see Figure 17.6). Among organic compounds, relative affinity for soil vs. water can be evaluated using the octanol : water partitioning coefficient (K_{ow}). K_{ow} provides a measure of the equilibrium distribution of a specific compound between octanol and water phases in direct contact. K_{ow} values for a wide variety of important compounds, including many pesticides are presented in Table 17.9. Large values of K_{ow} are indicative of compounds that are likely to preferentially associate with organic constituents (because of their relatively low polarity), such as those found in many soils; small values indicate a preference for the more polar aqueous phase. Common sources of hydrophobic organics, such as those listed in Table 17.9, include land application, either directly or by incorporation with soil amendments (such as process sludges) and surface spills.

17.5.2 Nonaqueous Phase Liquids

Compounds that display extremely low aqueous solubility can exist as a separate liquid phase in groundwater systems, if present in sufficient quantities. In groundwater systems, these contaminant phases are referred to as nonaqueous phase liquids (NAPLs), and their behavior will be fundamentally different from that of the bulk aqueous phase. In the case of a NAPL with density less than the surrounding water, the

TABLE 17.9 Octanol–Water Partition Coefficients (K_{ow}) for Selected Soil and Groundwater Contaminants

Compound	$\log_{10} K_{ow}$
<i>n</i> -Octane	5.18
Acetone	−0.24
Benzene	2.13
Toluene	2.69
Xylenes	3.12–3.18
Phenol	1.45
Naphthalene	3.36
Polychlorinated Biphenyls	4.09–8.23
2,3,7,8-Tetrachlorodibenzo- <i>p</i> -dioxin	6.64
Lindane	3.78
Dieldrin	5.48
<i>p</i> - <i>p</i> '-DDT	6.36
Parathion	3.81
Malathion	2.89
Atrazine	2.56
RDX	0.87

Source: From Schwarzenbach, R. P., Gschwend, P. M., and Imboden, D. M. 1993. *Environmental Organic Chemistry*, John Wiley & Sons, New York. With permission.

nonaqueous phase is referred to as a light NAPL (LNAPL), and will generally be found at or near the phreatic surface (see Figure 17.11). Fluctuations in the height of the phreatic surface can enhance LNAPL dispersion. LNAPL spills often are attributable to leaking underground fuel tanks (LUFTs) used for storage of petroleum hydrocarbons. Cleanup activities focus on location of the LNAPL phase and removal, by pumping if possible. Hydrocarbons that cannot be recovered by these methods may be subjected to other remediation methods, including soil vapor extraction, *in situ* bioremediation, or a combination.

Low-solubility organic compounds with density greater than water can exist as a dense NAPL (DNAPL) phase; their behavior in the subsurface environment is fundamentally different from that of LNAPLs. DNAPL compounds, because of their high density, tend to sink in groundwater systems. Often these compounds have lower viscosity than water — as a result, they tend to sink according to a pattern known as viscous fingering (see Figure 17.12). As they sink, they leave a trail of residual DNAPL compounds among the solids (soil particles) and liquid (water) in their wake. A DNAPL phase will continue to sink until it completely disperses among the soil and water phases, or it reaches an impermeable boundary, such as a clay lens or bedrock (see Figure 17.12). DNAPL pools can be extremely difficult to locate and are often difficult (or impossible) to remediate using existing technologies. Furthermore, as DNAPLs are often comprised of halogenated solvents, microbiological activity is slow to bring about changes in their molecular structure, and the products of microbial degradation involving these compounds are sometimes more toxic than their parent compounds. Possibly the most common examples of DNAPL compounds are trichloroethylene (TCE) and perchloroethylene (PCE). As in the case of LNAPLs, the source of DNAPL spills is often leakage or failure of underground storage tanks.

17.5.3 High-Profile Organic Contaminants

Within roughly the last decade, several important classes of organic groundwater contaminants have been identified. The discoveries of these contaminants have been attributed to improvements in analytical capabilities, changes in industrial practices, as well as changes in efforts to identify and quantify contaminants in groundwater settings.

The text below describes two high-profile groups of contaminants that have emerged as potentially serious threats to groundwater supplies. These two classes of compounds have physical, chemical, and

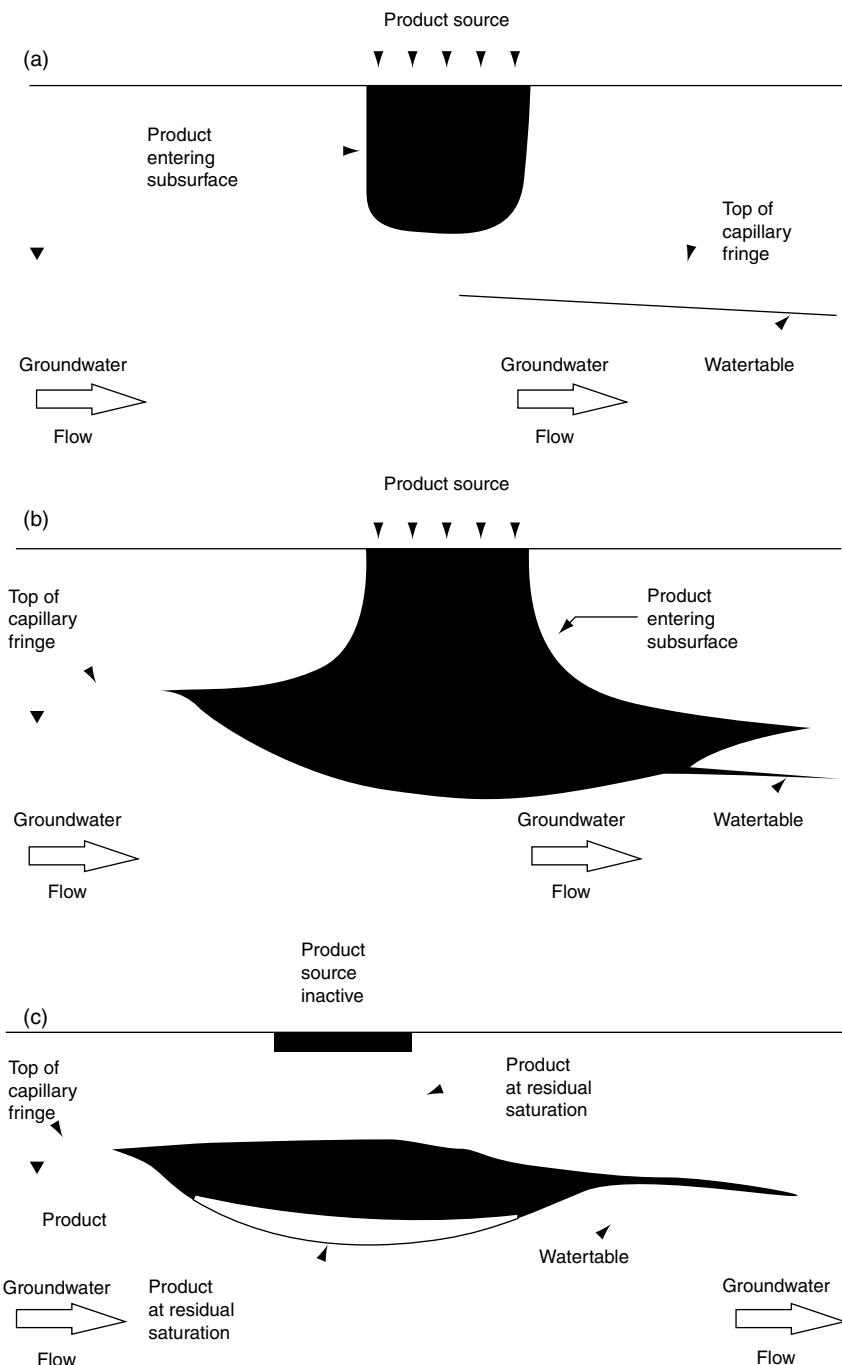


FIGURE 17.11 Schematic illustration of the behavior of LNAPL compounds. Under some conditions (a), the mass of an LNAPL spill is insufficient to allow penetration to the capillary fringe. With additional compound introduction (b), LNAPL product will reach the water table and begin to spread, though the compound will not penetrate far beyond the phreatic surface. If the source of LNAPL is eliminated (c), removal of LNAPL will allow for “rebound” of the water table. (From US Environmental Protection Agency. 1989. *Transport and Fate of Contaminants in the Subsurface*, EPA/625/4-89/019. With permission.)

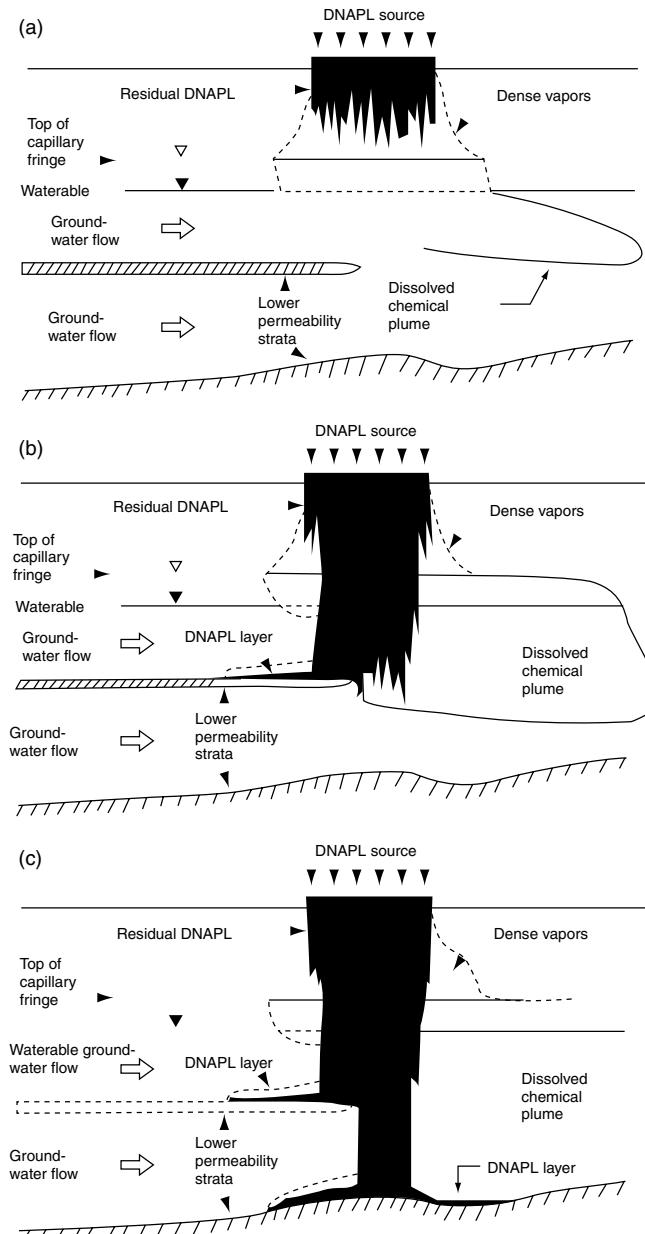


FIGURE 17.12 Schematic illustration of the behavior of DNAPL compounds. Under some conditions (a), the mass of a DNAPL spill is insufficient to allow penetration of the intact DNAPL to the capillary fringe; vertical movement of the DNAPL is by viscous fingering. With additional input of the compound (b) and (c), DNAPL product will reach the water table and continue to move vertically until it reaches an impermeable boundary. (From U.S. Environmental Protection Agency. 1989. *Transport and Fate of Contaminants in the Subsurface*, EPA/625/4-89/019. With permission.)

toxicological characteristics that have created important problems relative to the actions that can be taken to address groundwater contamination problems involving them.

17.5.3.1 N-Nitrosamines

These represent a family of compounds that are believed to be extremely potent carcinogens in humans. Notable among these compounds is *N*-nitrosodimethylamine (NDMA). NDMA is a common

contaminant in the manufacture of some rocket fuels, and has been discovered at high concentrations in the vicinity of some rocket fuel production facilities (Mitch et al., 2003). However, NDMA is also found in many food items and in tobacco smoke. NDMA can also be formed as a by-product of halogen-based water and wastewater disinfection processes; therefore, groundwater contamination by NDMA has been observed in the vicinity of groundwater recharge zones, where halogenated wastewater effluents are used as the water source for recharge.

Two general classes of reaction mechanism have been identified as sources of NDMA formation: nitrosation and oxidation of unsymmetrical dimethylhydrazine. Several reaction pathways have been identified that fit into these reaction mechanisms.

The US EPA has defined a cleanup level of 0.7 ng/L for NDMA in groundwater, based on a 10^{-6} increased lifetime cancer risk. However, no Federal regulatory standard for NDMA has been identified (Najm and Trussell, 2001). Treatment of NDMA-contaminated water is complicated by the physico-chemical characteristics of the compound. In particular, NDMA has a water solubility of roughly 4 g/L, a relatively low Henry's law constant (2.6×10^{-4} atm M^{-1} at 20°C), and is fairly hydrophilic. Therefore, it is not amenable to treatment by conventional processes such as gas–liquid transfer, or sorption. To date, the most effective treatment strategy appears to be UV-based photolysis (Mitch et al., 2003).

17.5.3.2 Fuel Oxygenates

The 1990 amendments to the Clean Air Act mandated use of fuel oxygenates as a means of reducing air pollution (CO and O₃ formation potential) in urban areas (Shih et al., 2004). Based on these amendments, fuels sold in many urban areas for automobiles were reformulated or oxygenated by the addition of ethers or alcohols. Although no specific fuel oxygenate was specified or mandated by regulation, methyl *tert*-butyl ether (MTBE) is the most widely used compound in the United States because of low cost and generally favorable engineering characteristics (Shih et al., 2004; Deeb et al., 2003).

MTBE is highly soluble in water, it has a relatively low Henry's law constant, and it has a low affinity for most soils. Therefore, it is not easily removed from water by conventional physico/chemical treatment processes, and it tends to migrate rapidly in aquifers. MTBE and related compounds have low taste and odor thresholds; therefore, groundwater contamination at low concentrations can result in the loss of water potability (Froines et al., 1998). MTBE has been linked to several other acute health problems, including potential adverse effects within the central nervous system and the respiratory system. It has been identified as a carcinogen in rats and mice and has been identified as a possible human carcinogen; however, the issue of carcinogenicity in humans remains unresolved, as conflicting reports appear in the literature.

The greatest source of MTBE contamination in groundwater supplies is believed to be LUFTs (Shih et al., 2004). Leaks from underground fuel tanks are quite common, and the tanks themselves are often located in close proximity to wells used for public water supplies. Complicating this situation is the fact that the characteristics of MTBE (and other fuel oxygenates) are such that they tend to migrate faster and farther than most other petroleum-derived fuel constituents in the subsurface environment.

17.6 Radionuclides

17.6.1 Radioactive Decay

Each chemical element in the periodic table has a fixed number of protons, but can exist with a variety of neutrons. Elemental isotopes are nuclei with the same number of protons, but different numbers of neutrons. Radionuclides are typically written with the chemical symbol and a superscript to the left of the symbol that denotes the atomic mass (total number of protons and neutrons) of that isotope. For example, ¹²C and ¹⁴C are isotopes of the element carbon; each has 6 protons, but are associated with 6 and 8 neutrons, respectively. Elements that undergo radioactive decay as a result of an unstable nuclear structure are referred to as radionuclides. As part of the decay process, radionuclides release ionizing radiation in the form of alpha particles (α), beta particles (β), and gamma radiation (γ).

Alpha particles are composed of two protons and two neutrons. They interact with matter through electrostatic interactions between their positive charge and the negative charge of orbital electrons in the medium through which they travel. Because of this, α particles deposit large amounts of energy in a very small travel distance; the mean travel distance of a 4.7 million electron volt (MeV) α particle, such as that emitted from ^{237}Np , is only 3.3 cm through air (Knoll, 1989).

Beta decay results in the production of high-energy electrons or positrons. Interactions of β particles result in deposition of the β particle energy in the medium through which it travels. Beta particles can also undergo electron-nuclear reactions as well. Beta particles lose energy at a lower rate as compared to α particles, and therefore have greater penetrating power. Positrons are the antiparticle of the electron. Once a positron has deposited its energy in the medium, it will undergo an annihilation reaction with an electron, producing two 0.511 MeV photons. Electrons are expelled from the nucleus in the conversion of a neutron to a proton, while a positron is emitted for a proton to neutron conversion. Weakly interacting particles known as (anti)neutrinos are also produced in β decay, but are not the focus of this discussion as they have an extremely small interaction probability with matter.

Gamma rays are high-energy photons released by excited nuclei as they transition from an excited state in a radioactive decay process. The γ photons are usually emitted in conjunction with an α or β decay process, and interact primarily with atomic electrons through photoelectric absorption, Compton scattering, and pair production. Unlike α and β particles, which slow down gradually via Coulombic forces, γ radiation interacts by partial or complete energy transfer to atomic electrons in sudden, discrete interactions. As γ rays deposit energy in a medium, they transfer energy to atomic electrons, producing high-energy electrons in the medium that interact in the same manner as β decay high-energy electrons (Knoll, 1989). Gamma radiation has a much larger penetrating power relative to α or β radiation; the mean free path of a 0.66 MeV photon, such as that emitted by ^{137}Cs , is approximately 100 m in air.

Each of the types of radiation described above is known as ionizing radiation. That is, they are capable of transferring enough energy to atomic electrons to allow the electrons to escape from the atom. As radiation interacts with the medium through which it travels, it creates a path of excitation (promotion of electron to a higher energy shell) and ionization in its wake.

The rate of radioactive decay of a group of atoms is proportional to the number of atoms of that isotope that are present in the group. This principle can be represented mathematically by defining a decay constant, λ , as the proportionality constant for decay. Equation 17.17 relates the time rate of change of the number of nuclei to the decay constant.

$$\frac{dN}{dt} = -\lambda N \quad (17.17)$$

This equation yields a simple analytical solution shown in Equation 17.18.

$$N = N_0 e^{-\lambda t} \quad (17.18)$$

Using Equation 17.18, it is possible to determine the amount of time for half of a radionuclide to decay, as indicated by Equation 17.19. This time is referred to as the half-life of an isotope, $t_{1/2}$.

$$t_{1/2} = \frac{\ln(2)}{\lambda} \quad (17.19)$$

As an element decays, the parent nucleus is transformed into a daughter nucleus, which may or may not be stable. If the daughter nucleus is unstable, it can then be considered a parent nucleus with an associated stable or unstable daughter nucleus. In this way, decay chains from a parent nucleus can involve numerous decays and radioactive emissions.

17.6.2 Definitions and Units

With these equations defined, it is necessary to understand the units that are commonly used to describe radionuclide activity. The activity, A , of a radioactive substance is defined as the number of decays of that radionuclide per unit time; the mathematical definition of activity is shown in Equation 17.20.

$$A = \lambda N \quad (17.20)$$

Activity is typically reported in units of Curies (Ci), where each Ci is defined as 3.7×10^{10} disintegrations per second. The SI unit for activity is the Becquerel (Bq), defined as one disintegration per second. It is important to note that a short half-life (high λ) radionuclide produces much higher activity relative to a long half-life (low λ) radionuclide when they are present in equimolar concentrations.

It is sometimes necessary to be able to calculate the amount of energy deposited in a medium exposed to radiation to analyze ionizing radiation effects. The amount of energy deposited per unit mass is referred to as dose, D , and is measured in Grays (Gy). A Gray is defined as one Joule of energy deposited per kg of material and is equivalent to 100 rads (radiation absorbed dose).

The biological damage done by ionizing radiation to organisms (e.g., humans) is called the dose equivalent, H , and is related to the dose by a quality factor, Q (see Equation 17.21).

$$H = DQ \quad (17.21)$$

Dose equivalent is measured in Sieverts (Sv) using the product of the quality factor and dose in Grays (rems, or roentgen equivalent man, are also dose equivalent units where one rem is 0.01 Sv). The quality factor is related to the linear energy transfer (LET) of the radiation passing through the organism. High LET radiation does greater damage than lower LET radiation, and this is reflected by a higher quality factor for external α and β radiation as compared to γ radiation. Radiation energy is typically reported in units of millions of electron volts (MeV), while LET is typically given in units of kiloelectron volts per centimeter.

17.6.3 Important Radionuclides and Their Sources

There exist a variety of sources of radionuclides that can be dissolved in groundwater. Sources are typically categorized as natural or anthropogenic. Natural sources of radionuclides include formation in the atmosphere caused by cosmic irradiation, and presence in mineral deposits of long-lived radioisotopes of heavy elements and their associated daughters. Anthropogenic sources include nuclear power plants, nuclear weapons explosions and testing, medical, research, industrial, and pharmaceutical radionuclide applications, and mining wastes. A number of specific radionuclides are known to represent potential problems in groundwater systems. Each is described below.

17.6.3.1 Radon (^{222}Rn)

The only long-lived isotope of radon is ^{222}Rn , which undergoes α decay with a half-life of 3.82 d. This radionuclide is produced by decay of ^{226}Ra , and is often associated with deposits that have significant uranium concentrations (^{226}Ra and ^{222}Rn are part of the ^{238}U decay chain). ^{222}Rn represents a significant health concern in some homes, entering as a dissolved species in drinking water or directly from soil as a gas. Additional radiation is emitted from daughter products of ^{222}Rn as it decays to a stable element. ^{222}Rn is a naturally occurring radioisotope, and can be significant for uranium mining in terms of worker safety and waste disposal (Fetter, 1993).

17.6.3.2 Radium (^{226}Ra and ^{228}Ra)

^{226}Ra ($t_{1/2} = 1600$ yr) is a daughter of ^{230}Th , part of the ^{238}U decay chain, and decays subsequently to ^{222}Rn . As ^{228}Ra ($t_{1/2} = 5.8$ yr) is a daughter of ^{232}Th , both radium radionuclides are associated with uranium- and thorium-containing deposits. Radium is more soluble than either thorium or uranium,

and can therefore be transported in higher concentrations with an associated higher activity, particularly when considering the relatively short $t_{1/2}$ of ^{228}Ra . As radium is associated with uranium deposits, it is therefore a potentially important problem in mining wastes (Freeze and Cherry, 1979).

17.6.3.3 Uranium (^{238}U and ^{235}U)

Uranium in mineral deposits is composed of 99.3% ^{238}U ($t_{1/2} = 4.5 \times 10^9$ yr) and 0.7% ^{235}U ($t_{1/2} = 7.1 \times 10^8$ yr). As ^{235}U is fissile (split by low-energy neutrons), it is used as fuel for nuclear power reactors worldwide. Both uranium isotopes are long-lived and part of long decay chains. Each has numerous radioactive daughter nuclei that produce emissions before decaying to stable states. Uranium is typically mined in open pits or underground and concentrated to 65 to 70% uranium via flotation, dissolution, and solvent extraction processes. Further purification to U_3O_8 and enrichment to increase the fraction of the ^{235}U isotope can achieve desired levels of ^{235}U , with 3 to 4% being typical for nuclear power reactors (Choppin et al., 1995).

17.6.3.4 Neptunium (^{237}Np)

^{237}Np can be found in nuclear waste as the product of nuclear reactions (in a nuclear power reactor) converting ^{238}U to ^{237}U , which then β decays to ^{237}Np . It has a $t_{1/2}$ of 2.2×10^6 yr and undergoes α decay to ^{233}Pa , with associated daughter decays from that element and others. ^{237}Np is of great concern in high-level and spent nuclear fuel disposal as it is only slightly sorbed (i.e., high mobility) in many groundwater environments and has a long half-life (USDOE, 1995).

17.6.3.5 Tritium (^3H)

Tritium undergoes β decay with a half-life of 12.43 yr, producing ^3He as a daughter element. As tritium can combine with hydrogen and oxygen to form tritiated water, it poses a significant threat in terms of groundwater transport. Tritium is especially useful in delineating groundwater recharge occurring since the 1963 spike of tritium released by the United States and Soviet Union prior to the implementation of an above-ground testing treaty (Fritz et al., 1991). Tritium concentrations in natural waters have increased nearly tenfold since the advent of the nuclear age (Fetter, 1993). Tritium is formed naturally in the atmosphere when high-energy neutrons bombard ^{14}N , producing ^{12}C and ^3H . It is also produced in large quantities at nuclear power plants, and in nuclear weapons testing and production.

17.6.3.6 Carbon (^{14}C)

Like tritium, ^{14}C can be used for dating purposes, in this case for organic material. It decays by β emission to ^{14}N , and has a $t_{1/2}$ of 5715 yr. ^{14}C labeling of organic compounds is often done to provide sensitive detection methods for experiments with organic compounds. As it can be oxidized to soluble or gaseous $^{14}\text{CO}_2$, or form a variety of organic compounds, ^{14}C can pose a significant transport threat as a radionuclide. It can be produced naturally in the atmosphere when low-energy neutrons bombard ^{14}N , producing ^{14}C and ^1H . Anthropogenic sources include nuclear power plants, nuclear weapons testing, and research laboratories.

17.6.3.7 Fission Fragments (^{90}Sr , ^{99}Tc , ^{129}I , and ^{137}Cs)

Fission fragments are radioactive materials produced by fission of heavy isotopes such as ^{235}U . The fission reaction typically splits the heavy atom into two (sometimes three) lighter isotopes, with associated neutron, neutrino, and γ radiation emissions as well. The lighter isotopes are produced primarily in the atomic mass range of 80 to 155. Although a variety of fission fragments are produced, ^{90}Sr , ^{99}Tc , ^{129}I , and ^{137}Cs are of primary concern due to the significant amounts that are produced and their relatively long half-lives.

^{90}Sr decays by β emission to ^{90}Y with a $t_{1/2}$ of 29.1 yr. ^{90}Y then decays to stable ^{90}Zr via β emission. Strontium is a prominent fission fragment with no associated γ emissions. In high-level nuclear waste and spent fuel, it is one of the highest activity radioisotopes for the first 200 yr after fission ceases. It can be particularly dangerous when ingested by humans because it is known to accumulate in or near the bones, resulting in greater irradiation of sensitive bone marrow.

^{99}Tc undergoes β decay to ^{99}Ru with a $t_{1/2}$ of 2.1×10^5 yr. In addition to being a fission fragment, ^{99}Tc is also associated with medical and research wastes where it is used in diagnostic testing. ^{129}I is an important fission fragment as it β decays to stable ^{129}Xe with a long $t_{1/2}$ of 1.7×10^7 yr. It is also used in medical diagnostic tests and can therefore be found in low-level waste as well as spent fuel.

^{137}Cs emits γ and β radiation while decaying to stable ^{137}Ba and has a $t_{1/2}$ of 30.3 yr. Along with ^{90}Sr , ^{137}Cs is one of the high-activity isotopes in high-level nuclear waste for the first 200 yr after fission ceases.

17.6.4 Radioactive Waste Disposal

There are four different classes of radioactive wastes defined in United States regulations. *High-Level Waste* (HLW) is defined as radioactive waste produced in the reprocessing of nuclear fuel to recover usable uranium and plutonium. *Spent Nuclear Fuel* (SNF) is nuclear fuel that has been removed from a power reactor but has not been reprocessed. *Transuranic Waste* (TRU) consists of wastes from nuclear power or weapons production operations with atomic number greater than 92 (uranium), being primarily α emitters with long half-lives; it is typically associated with nuclear weapons production, fuel fabrication, and reprocessing. *Low-Level Waste* (LLW) is high volume, low radioactivity waste produced by industry, consumer products, research and medical labs, government, and nuclear power plants (Tang and Saling, 1990).

The United States does not reprocess commercial nuclear fuel, although Japan and most of Europe have chosen this option. As a result, US wastes associated with the nuclear power industry are primarily in the form of SNF, but with significant amounts of HLW and TRU from government (defense) related activity. The following discussion will focus on radioactive waste disposal options being pursued in the United States, but different strategies are being considered all over the world.

17.6.4.1 HLW and SNF Disposal

Ultimate disposal of HLW and SNF is unlikely to occur in a timely fashion, though the ultimate responsibility lies with the Federal government to solve the problem. At present, three Federal agencies within the United States participate in the management and disposal of HLW and SNF: the US EPA, the US Nuclear Regulatory Commission, and the US Department of Energy. The policies that define the approach to be used for HLW disposal are defined in the Nuclear Waste Policy Act of 1982.

The current US plan for ultimate disposal (storage) of HLW and SNF is to place it in the deep geologic repository at Yucca Mountain in Nevada. The repository will rely on an engineered barrier system (EBS) combined with site characteristics (geology) to contain the waste. A deep repository has the advantages of providing the best safety, flexibility, retrievability, convenience, maintenance, and feasibility relative to other disposal options (Faure, 1991). The US Department of Energy is currently developing plans for construction of the repository at Yucca Mountain (DOE, 2005a). The entire HLW and SNF disposal plan is far behind schedule due to technical difficulties, public fear and protests, and resistance from the host state. Because of this, there are plans to place the waste temporarily in monitored, retrievable storage (MRS). Once the Yucca repository is complete, waste will be transferred from the MRS to the final repository.

17.6.4.2 TRU Disposal

Transuranic wastes have been held in interim storage since 1974 while an ultimate disposal solution was pursued. The wastes will be placed in the Waste Isolation Pilot Plant (WIPP), located in a salt bed repository in New Mexico. Site characterization was completed in 1978 and an environmental impact statement filed in 1980. Full construction started in 1984, and the site began operations in March 1999 (DOE, 2005b). TRU disposal plans are notable because it is the only successful radioactive waste disposal program in the United States that has not been significantly delayed or halted.

17.6.4.3 LLW Disposal

LLW disposal is the responsibility of the states, with technical assistance from the federal government. To alleviate some of the difficulty of each state constructing an LLW landfill, the states have formed compacts where numerous states can contribute waste to a single site. Much like federal HLW disposal efforts, the

states are well behind schedule in meeting their planned opening dates for disposal sites. Currently, three low-level radioactive waste disposal facilities are operating in the United States. These facilities are located at Barnwell, SC; Hanford, WA; and Clive, UT (NRC, 2004).

17.6.5 Remediation

Remediation of radionuclide contaminated sites can be a costly and difficult endeavor. The chemical behavior of an atom in the environment is dictated by its atomic structure (not nuclear structure). Therefore, different isotopes of the same element behave identically in terms of chemical reactions. The essential difference between isotopes is that unstable elements will undergo radioactive decay. Because of this, radionuclide remediation techniques consist of the same methods that are used for inorganics and have been described previously in this chapter.

A fundamental difference between radionuclide and inorganic wastes is that radionuclides do not require direct contact to cause damage to an organism; radiation provides its own transport mechanism to reach a target. Because of this, providing for public and worker safety at a site that is contaminated by radioactive wastes may be more difficult to accomplish than at sites without such contamination. In addition, remediation techniques can concentrate a radionuclide (e.g., ion exchange or activated carbon) to such an extent as to create a hazardous radiation field; appropriate shielding and transportation arrangements of the concentrate medium must be considered.

It is sometimes possible to use the unstable characteristics of radionuclides to facilitate remediation. For radionuclides with short half-lives, it may only be necessary to contain the waste for an appropriate period of time before the unstable nuclei have decayed below the background radiation. In this manner, it is possible to use the decay properties of the nucleus to facilitate safe disposal of some radionuclide wastes. The idea of transmutation (bombarding radioactive waste with neutrons or protons) has been proposed as a way to stabilize radioactive nuclei, but has been discarded primarily because of high costs associated with this process.

17.7 Particulates

The definition of a particle is arbitrary and will depend on the application. For purposes of this presentation, particulate matter will refer to constituents that exist as a distinct solid phase, either within or apart from the groundwater itself. Though no absolute distinction is possible to define particles based on size, constituents with a size of 10 nm or greater may be considered to display particle behavior in groundwater system. Figure 17.13 provides a graphical representation of the spectrum of particle sizes that may be found in groundwaters, some important examples, and an indication of the physicochemical processes by which these constituents may be separated from the aqueous phase.

Microorganisms play an important role as particulate contaminants in groundwater systems. Important categories of microbial groundwater contaminants include viruses, bacteria, and protozoan cysts. Based on the arbitrary definition given above, viruses are representative of the smallest particles that are likely to be present in groundwater systems. Their small size facilitates mobility in groundwater systems; however, even though viruses are smaller than the pore size openings common to groundwater systems, their movement in aquifers is not unimpeded by soil particles. Like all constituents in groundwater systems, movement is largely governed by association with soil grains. Viruses display surface characteristics (large surface:volume ratio, net surface charge) that allow them to associate with soil grains. In the general case, groundwater supplies are assumed to be subject to possible contamination by pathogenic viruses, and disinfection processes for public water supplies are designed to address this source of contamination. Virus movement within aquifers is strongly influenced by electrostatic interactions with soil surfaces and Brownian motion. Microorganisms from the other categories (e.g., bacteria and protozoan cysts) will also move in groundwater systems; however, their movement within an aquifer is likely to be largely influenced by size exclusion and inertial processes, in contrast with the phenomena that govern movement of virus

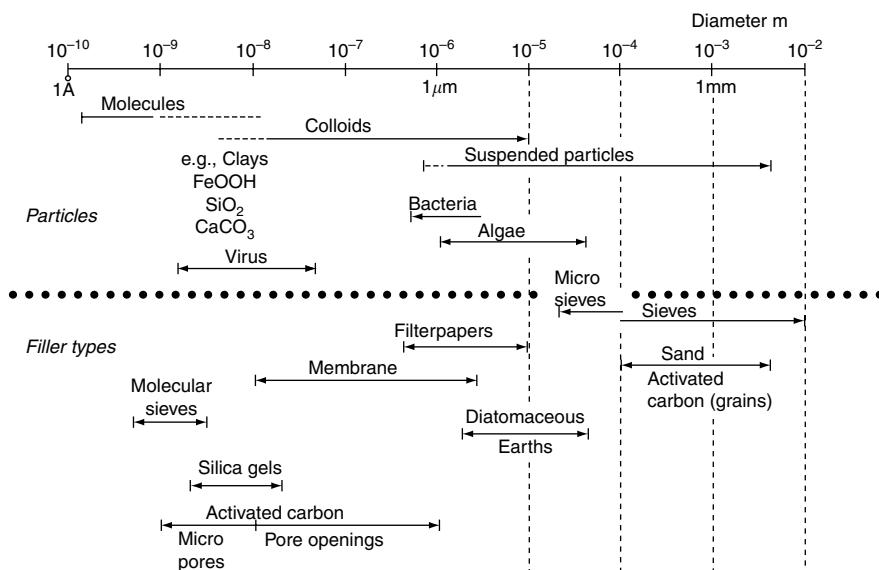


FIGURE 17.13 Spectrum of particle sizes that may be observed in natural waters, including groundwaters. Physical processes for separation of particles of a given size from the aqueous phase are also indicated. (From Stumm, W. 1977. *Environ. Sci. Technol.*, 11, 1066. With permission.)

particles. Groundwater supplies that are under the direct influence of surface waters are subject to more stringent disinfection requirements for potabilization due to the possible presence of protozoan cysts such as *Giardia lamblia* and *Cryptosporidium parvum*, which are generally more difficult to inactivate than either bacteria or viruses by conventional chemical disinfectants (e.g., chlorine).

Other, nonmicrobial colloidal particles can also exist and be transported in groundwater systems. Movement of these particles represents an important mechanism of transport for constituents that associate strongly with them. Important examples of compounds for which groundwater transport is enhanced by the presence of colloids include hydrophobic organic compounds, metals, and microorganisms.

Glossary

Acid/Base Reaction: A reaction in which proton transfer takes place.

Amphotericism: Ability of a solid surface or solute to display the characteristics of an acid or base.

Chemisorption: Association of solute (adsorbate) molecules with a solid (sorbent) as a result of relatively strong chemical bonds, similar in strength to covalent bonds.

Dense, Nonaqueous Phase Liquid (DNAPL): (Organic) liquids that are more dense than water, often being comprised of halogenated organic solvents.

Groundwater Contaminant: Groundwater constituent, often viewed as undesirable, introduced to the subsurface either directly or indirectly as a result of human activity.

High Level Waste (HLW): Radioactive waste produced in the reprocessing of nuclear fuel.

Humic Substances: Stable organic products of the natural decay of plants and animals.

Ion Exchange: Association of charged solutes (ions or complexes) with a solid (sorbent) as a result of electrostatic interactions.

Light Nonaqueous Phase Liquid (LNAPL): (Organic) liquids that are less dense than water, often being comprised of hydrocarbons or hydrocarbon mixtures, such as those that result from petroleum refining.

Low Level Waste (LLW): High volume, low radioactivity waste produced from many industrial and commercial processes.

Micelle: Pseudo-phase agglomeration of (organic) surfactant molecules in which polar “ends” of the surfactant molecules are directed outward and the nonpolar “ends” are directed inward.

Nonaqueous Phase Liquid (NAPL): Usually comprised of organic compounds having extremely low aqueous solubilities.

NO_x: Nitrogen oxides — common atmospheric constituents produced by a number of natural and anthropogenic processes, including combustion of fossil fuels.

Oxidation/Reduction (Redox) Reaction: Reaction in which electron transfer takes place.

Particle: Constituents that exist as a distinct solid phase within or apart from a solvent; oftentimes particles are operationally defined as constituents greater than a given size (e.g., 10 nm).

p ϵ : Negative log₁₀ of electron (e^-) activity; a representation of electron availability in solution.

pH: Negative log₁₀ of proton (H^+) activity; a representation of proton availability in solution.

Physisorption: Association of solute (adsorbate) molecules with a solid (sorbent) as a result of relatively weak physical bonds.

Polarity: A relative term describing the distribution of electrons (or electronic charge) around a molecule. Highly polar molecules generally contain atoms of widely varying electronegativity.

Radionuclides: Unstable elemental isotopes that release ionizing radiation.

SO_x: Sulfur oxides — common atmospheric constituents produced by a number of natural and anthropogenic processes, including combustion of fossil fuels.

Transuranic Waste (TRU): Wastes from nuclear power or weapons production having an atomic number greater than 92.

Zero Point of Charge (zpc): pH at which net surface charge is zero.

References

- Choppin, G., Rydberg, J., and Liljenzin, J. 1995. *Radiochemistry and Nuclear Chemistry*, 2nd Edition, Butterworth-Heinemann Limited, Oxford, UK.
- Clark, M. M. 1996. *Transport Modeling for Environmental Engineers and Scientists*, John Wiley & Sons, New York.
- CRC Handbook of Chemistry and Physics, 59th Edition. 1978. R. C. Weast and M. J. Astle (Eds.), CRC Press, Boca Raton, FL.
- Deeb, R. A., Chu, K., Shih, T., Linder, S., Suffet, I., Kavanaugh, M. C., and Alvarez-Cohen, L. 2003. MTBE and Other Oxygenates: Environmental Sources, Analysis, Occurrence, and Treatment, *Environmental Engineering Science*, **20**, 5, 433–447.
- Department of Energy. 2005a. *Yucca Mountain Project*, <http://www.ocrwm.doe.gov/ymp/index.shtml>.
- Department of Energy. 2005b. *Waste Isolation Pilot Plant*, <http://www.wipp.ws/>.
- Faure, G. 1991. *Principles and Applications of Inorganic Geochemistry*, Macmillan Publishing Company, New York.
- Fetter, C. 1993. *Contaminant Hydrogeology*, Macmillan Publishing Company, New York.
- Freeze, R. and Cherry, J. 1979. *Groundwater*, Prentice-Hall, Englewood Cliffs, NJ.
- Fritz, S. J., Drimmie, R. J., and Fritz, P. 1991. Characterizing shallow aquifers using tritium and ¹⁴C: Periodic sampling based on tritium half-life, *Applied Geochemistry*, **6**, 17–33.
- Froines, J. R., Collins, M., Fanning, E., McConnell, R., Robbins, W., Silver, K., Kun, H., Mutialu, R., Okoji, R., Taber, R., Tareen, N., and Zandonella, C. 1998. *An Evaluation of the Scientific Peer-Reviewed Research and Literature on the Human Health Effects of MTBE, its Metabolites, Combustion Products and Substitute Compounds*, <http://tsrtp.ucdavis.edu/mtberpt/vol2.pdf>.
- Knoll, G. 1989. *Radiation Detection and Measurement*, 2nd Edition, John Wiley & Sons, New York.
- Larson, T. E. and Buswell, A. M. 1942. Calcium carbonate saturation index and alkalinity interpretations, *Journal, AWWA*, **34**, 1664.
- Martell, A. E. and Smith, R. M. 1974–1989. *Critical Stability Constants*, Volumes 1–6, Plenum Press, New York.

- Mitch, W. A., Sharp, J.O., Trussell, R.R., Valentine, R.L., Alvarez-Cohen, L., and Sedlak, D.L. 2003. N-Nitrosodimethylamine (NDMA) as a drinking water contaminant: A review, *Environmental Engineering and Science*, **20**, 5, 389–404.
- Najm, I. and Trussell, R. R. 2001. NDMA formation in water and wastewater, *Journal, American Water Works Association*, **93**, 2, 93–99.
- Newman, A. 1994. Low-level radioactive wastes, high-level risk? *Environmental Science & Technology*, **28**, 11, 488A.
- Nuclear Regulatory Commission. 2004. *Locations of Low-Level Waste Disposal Facilities*, <http://www.nrc.gov/waste/llw-disposal/locations.html>.
- Schnitzer, M. and Khan, S. U. 1972. *Humic Substances in the Environment*, Marcel Dekker, New York.
- Schwarzenbach, R. P., Gschwend, P. M., and Imboden, D. M. 1993. *Environmental Organic Chemistry*, John Wiley & Sons, New York.
- Seinfeld, J. H. 1986. *Atmospheric Chemistry and Physics of Air Pollution*, John Wiley & Sons, New York.
- Shih, T., Rong, Y., Harmon, T., and Suffett, M. 2004. Evaluation of the impact of fuel hydrocarbons and oxygenates on groundwater resources, *Environmental Science & Technology*, **38**, 42–48.
- Snoeyink, V. L. and Jenkins, D. 1980. *Water Chemistry*, John Wiley & Sons, New York.
- Stumm, W. 1977. Chemical interaction in particle separation. *Environmental Science Technology* **11**, 1066.
- Stumm, W. and Morgan, J. J. 1981. *Aquatic Chemistry*, 2nd Edition, John Wiley & Sons, New York.
- Stumm, W. and Morgan, J. J. 1996. *Aquatic Chemistry*, 3rd Edition, John Wiley & Sons, New York.
- Tang, Y. and Saling, J. 1990. *Radioactive Waste Management*, Hemisphere Publishing Company, New York.
- Temple, K. L. and Delchamps, E. W. 1953. Autotrophic bacteria and the formation of acid in bituminous coal mines. *Applications in Microbiology*, **1**, 255.
- US Department of Energy (USDOE). 1995. *Total System Performance Assessment: An Evaluation of the Potential of the Yucca Mountain Repository*, prepared by TRW for USDOE, B00000000-01717-2200-00136, Rev. 01, Nov., 1995.
- US Environmental Protection Agency. 1989. *Transport and Fate of Contaminants in the Subsurface*, EPA/625/4-89/019.

Further Information

Clark (1996) is a reference text in which transport phenomena are presented in terms of environmental processes. US EPA (1989) provides basic information on many of the fundamental phenomena that govern the fate and transport of subsurface contaminants. Freeze and Cherry (1979) is perhaps the most frequently cited reference text on the subject of groundwater, and includes an excellent presentation of many important aspects of groundwater chemistry. Martell and Smith (1974–1989) represents a critical evaluation of stability (equilibrium) constants for complexation and precipitation reactions involving organic and inorganic aqueous constituents. This series of six volumes represents the most comprehensive summary of data on the subject. Snoeyink and Jenkins (1980) is an intermediate-level text on the subject of aquatic chemistry. Stumm and Morgan (1996) is the third edition of this advanced text on the subject of aquatic chemistry and is perhaps the most frequently cited reference on the subject.

18

Nonreactive Contaminant Transport in the Saturated Zone

18.1	Introduction.....	18-1
18.2	Hierarchical Heterogeneity.....	18-2
18.3	Role of Multiple Scales	18-2
18.4	Dispersion	18-3
18.5	Eulerian and Lagrangian Perspectives	18-3
18.6	Classical Deterministic Models	18-4
18.7	Alternate Approaches	18-4
	Statistical Mechanical • Continuous Time Random Walks •	
	Fractional Advection-Dispersion • Stochastic Perturbation	
18.8	Locally Stationary Systems	18-11
18.9	Spreading vs. Dilution	18-12
18.10	Experiments	18-14
	Lab Experiments • Field Experiments	
	Glossary.....	18-15
	Acknowledgments.....	18-16
	References	18-16
	Further Information	18-18

Natalie Kleinfelter and
John H. Cushman
Purdue University

18.1 Introduction

The modeling of transport in porous media continues to be an important and challenging problem with applications in groundwater pollution, bioremediation of soils, petroleum engineering, food science engineering, drug delivery systems, and blood flow in tissues to name a few. Over the past several decades the need for a multidisciplinary approach to this problem has been recognized and collaborative work is becoming the rule rather than the exception. A combination of hydrologists, mathematicians, chemists, microbiologists, engineers, and physicists is needed to maintain progress in the modeling and prediction of flow and transport in porous media for both nonreactive and reactive solutes. This chapter focuses on the nonreactive case while the subsequent chapter examines the reactive case.

We employ the language commonly used in hydrology, but the reader should keep in mind that the transport problem is not limited to this discipline. The current understanding of transport of tracers (dissolved or suspended) is examined with emphasis on passive tracers. Due to the natural variability of geologic formations, that is, heterogeneity, the problem is extremely complex. Early theories assumed homogeneity of the medium but later this assumption was found to be inadequate for natural formations; now, sophisticated stochastic models and approaches to account for natural heterogeneity are the rule. The mathematical and physical challenges involved in model development, the design of new experiments to determine parameters and verify the models, plus the need for better numerical methods to solve the resulting equations in reasonable computational time are all necessary components for new theories.

We begin with a discussion of different types of heterogeneities and the importance of scale, as well as a discussion of the most important driving processes that lead to dispersion. Models can be developed from two perspectives, Eulerian or Lagrangian and models may be either deterministic or stochastic. The classical deterministic model is presented along with a variety of advanced models. These models are subsequently compared and the advantages and disadvantages discussed. The chapter concludes with a brief discussion on laboratory and field-scale experiments.

18.2 Hierarchical Heterogeneity

The concept of homogeneity is inherently linked to the concept of the scale of observation and the existence of a representative elementary volume (REV) (in the sense of Bear, 1972). Porous media may appear homogeneous at one scale and heterogeneous at another, and hence, the scale of observation is critically important. Heterogeneous media vary widely in structure and function, but are invariably hierarchical.

Hierarchical heterogeneity can be classified as either *structural* or *functional*. In a structural hierarchy, the porous media can be decomposed into successively nested, interacting, physical subunits. Functional hierarchy relates to the transport processes within the medium. Often the two types of hierarchy are intimately connected.

One can further break structural (functional) hierarchy into *discrete* and *continuous* types. In a discrete hierarchical medium, a finite number of nested structural subunits (subprocesses) exist as opposed to a continuous hierarchy where no clear-cut decomposition is in evidence and the number of subunits goes to infinity. In discrete hierarchical media, every level has a specific set of state variables of relevance. Examples include the pore-scale velocity at the pore scale and the Darcy velocity in a homogenized system. As one moves up the hierarchy, only portions of the information from the level below are retained by the system; this loss of information leads to the appearance of new constitutive variables at each scale. The windows associated with the experimental apparatus must correspond to the discrete hierarchical structure of the porous medium to decouple the scales and propagate information over scales.

Continuous hierarchy is associated with “self-organization” and is evident in long-range correlation among the dynamic or static processes. One example is a fractal porous medium whose properties have infinite correlation lengths (Wheatcraft and Tyler, 1988). Information transfer in such media cannot be decomposed as in the discrete media, but can be scaled (compressed). The constitutive variables for such a system are often wave vector (dual to spatial displacement) and frequency (dual to time elapsed) dependent and appear in nonlocal theories.

In the following section we discuss the measurement and theoretical issues associated with multiple scales.

18.3 Role of Multiple Scales

It is important to understand the role played by multiple scales when developing models and hence also in the design of experiments. Heterogeneity may evolve over many orders of magnitude of space and time and, depending on the nature of the problem, different models may be required at different scales. The requisite

physics can be rigorously investigated using upscaling techniques such as mixture theory, mathematical homogenization, stochastic perturbation tools, and central limit theorems (Cushman et al., 2002). The assumptions involved in the mathematical analysis depend on scale separation and scale of observation.

The scale of the window of observation must be accounted for in the models. The attribute that one wants to measure or describe must be well defined at that scale. When one measures an attribute it is often a “point” measurement, where a “point” is defined by the appropriate scale for the problem. Changing the resolution or changing the scale may radically change the value measured and care should be taken in measurement to ensure that the appropriate attribute and the appropriate resolution are used to provide a meaningful value for the model. When the window of observation or the resolution increases, plateaus in measured attributes may exist indicating a discrete hierarchy, while at other times they will not, indicating an evolving (continuous) heterogeneity. In the case of discrete hierarchies, constitutive relations will be local in the asymptotic limit while in the continuous case they will be nonlocal. Discussions of measurement techniques for hierarchical porous media can be found in Cushman (1990, 1997).

18.4 Dispersion

Dispersion is a measure of solute spreading from the mean concentration and can be influenced by a variety of factors. The scale of observation and the processes of interest are key in determining the dominant dispersive processes. Several types of dispersion are listed below:

1. *Diffusion driven dispersion* is the result of particle interactions at the pore scale. The scale at which this is significant will depend on the particle size as well as the observation window. Diffusion takes place regardless of the presence of a porous structure, but may be affected by the porous structure.

2. *Heterogeneity driven dispersion* is that most commonly discussed in porous media literature. The porous matrix causes velocity fluctuations to occur over a variety of scales. Variations in the hydraulic conductivity, possibly over several orders of magnitude (Gelhar, 1986) lead to velocity fluctuations which in turn affect transport of a dissolved solute. Heterogeneous systems naturally display larger velocity fluctuations than homogeneous systems and hence greater mixing. The characterization of this is usually evident in the macrodispersivity tensor defined at the Darcy or higher scales. In general, larger scales lead to larger macrodispersivity tensors.

3. *Viscosity driven dispersion* occurs when two immiscible liquids have different viscosities. Viscous fingering leads to dispersive mixing that would not have occurred without the presence of viscous forces and surface tension. Parlange has done significant work on fingering (Liu et al., 1994; Selker et al., 1996).

4. *Buoyancy induced dispersion* occurs when the fluids or constituents have varying density. Examples include: Rayleigh–Benard convection where fluids of different temperatures drive mixing; air bubbles in water; and mixtures of DNAPLs (dense nonaqueous phase liquids) and LNAPLs (light nonaqueous phase liquids).

5. In other situations the transported particles themselves may enhance dispersion; two examples include microbial and colloidal systems. The transport of microbes in a fluid may have increased dispersion due to microbial motility (Berg, 2000; Bonilla and Cushman, 2002). Charged or magnetic colloids may also be affected by the presence of externally applied electric or magnetic fields.

18.5 Eulerian and Lagrangian Perspectives

Transport may be modeled from two different perspectives, Eulerian and Lagrangian or, in some cases, a mixture of these. Models can be continuous or discrete. The perspective depends on the mathematical approach one wishes to employ which in turn may be dictated by the situation as well as available data. In the Lagrangian approach, a particle is referenced from its initial location and a description of its evolution is developed. Hence in the Lagrangian approach the viewer observes the world from the location of a particle rather than a fixed location in space and uses the particle pathline or its statistics to construct

models. The Eulerian perspective describes the physics and attribute evolution at a fixed location in space. Translation between the two modeling perspectives is always possible in principle, but not always possible in practice without simplifying assumptions. Additionally, it is difficult to transfer data from one perspective to the other.

18.6 Classical Deterministic Models

Most classical deterministic models for transport employ Darcy's law in the flow problem to obtain the velocity field which can then be utilized with Fick's law and mass conservation of a solute. These conservation laws rely on a discrete scale separation (e.g., Darcy vs. pore). The classical deterministic advection-dispersion equation (ADE) can be derived from rigorous approaches that allow for specific characterization of the material parameters (conductivity, dispersivity). These approaches include homogenization (Sanchez-Palencia, 1980), generalized Taylor–Aris–Brenner (Shapiro and Brenner, 1988), volume averaging in the sense of Quintard and Whitaker (1987), mixture theory and hybrid mixture theory (Cushman et al., 2002), and renormalization (Creswick et al., 1992). A number of these tools are reviewed in Cushman et al. (2002). These models revolve around assumptions made at either the microscale (pore scale) and/or the macroscale (in this case the Darcy scale) to yield the constitutive theory. Theories that lead to more complex and generally non-local models for flow and transport will be discussed in the subsequent section. In this section we present the deterministic ADE without derivation.

The localized mass balance equation for a constituent is

$$\frac{\partial(nC)}{\partial t} + \nabla \cdot \mathbf{F} = nr \quad (18.1)$$

where C is the concentration, n is the porosity (volume fraction in mixture theory), \mathbf{F} is the mass flux, and r is the rate of solute mass produced per unit volume. Since this chapter considers only nonreactive solutes, $r = 0$ unless there is mass added from an exterior source. In general however, $r = r(\mathbf{x}, t)$ for a conservative tracer and $r = r(C, \mathbf{x}, t)$. By employing an expression for the flux that includes macroadvection as well as local dispersion, the ADE is

$$\frac{\partial(nC)}{\partial t} + \nabla \cdot (n\mathbf{V}C) - \nabla \cdot (n\mathbf{D} \cdot \nabla C) = nr. \quad (18.2)$$

\mathbf{D} is the local (symmetric) dispersion tensor. Equation 18.2 can be regarded as an equation for the average concentration on a scale large enough so that the effects of individual pores are not observed. At scales smaller than the REV this will not hold due to variations in the pore-scale velocity and similarly, in heterogeneous systems, at scales much larger than the REV this equation breaks down. Often \mathbf{D} is taken as the effective dispersion tensor or the macrodispersivity tensor at the larger scales. After solving the flow problem, with the definition of the velocity depending on the scale of measurement and interest, the transport equation is solved.

As the deterministic ADE has been found to break down for many natural and laboratory media, we examine some alternate deterministic approaches and stochastic approaches in the next two sections that lead to nonlocal generalizations of the ADE.

18.7 Alternate Approaches

18.7.1 Statistical Mechanical

Generalized hydrodynamics utilizes statistical mechanical tools with appropriate probability densities to examine the evolution of the concentration in space and time. This approach was developed by Cushman and Ginn (1993) and is described in detail in Cushman (1997).

In a statistical mechanical approach, the N particles of the solute are labeled $1, \dots, N$ and this collection of particle locations by $\mathbf{X} = (\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_N)$. One then examines the probability, f , of particles having a certain displacement. The concentration, C , is defined by

$$C(\mathbf{x}, t) = \sum_{i=1}^N \int_{V^N} \delta(\mathbf{x} - \tilde{\mathbf{X}}_i) f(\tilde{\mathbf{X}}) d\tilde{\mathbf{X}} \quad (18.3)$$

with δ the Dirac delta function, V is the volume of possible displacements for a particle, i the particle label, and tilde indicates particle displacement from the initial time. It is assumed that the probability measure for the ensemble is conserved in time and that the mean increment of the displacement depends on time lag alone and not on the particle label. If the mean velocity is not dependent on space, then the latter assumption holds even if the velocity is nonsteady. The model can be used for a wide class of media that have no scale separation. The resulting equation for the concentration is

$$\begin{aligned} \frac{\partial C}{\partial t} &= -\nabla \cdot (\mathbf{V}C) + \nabla_x \cdot \int_0^t \int_{\mathbf{R}^3} \mathbf{D}_1(\mathbf{y}, t, \tau) C(\mathbf{x} - \mathbf{y}, t - \tau) d\mathbf{y} d\tau \\ &\quad + \nabla_x \cdot \int_0^t \int_{\mathbf{R}^3} \mathbf{D}_2(\mathbf{y}, t, \tau) \cdot \nabla_{\mathbf{x}-\mathbf{y}} C(\mathbf{x} - \mathbf{y}, t - \tau) d\mathbf{y} d\tau \end{aligned} \quad (18.4)$$

The first term on the right-hand side results from mean convection where \mathbf{V} is the average velocity of the particles. The second and third terms are nonlocal terms containing information from the convective and dispersive fluxes, respectively. This equation is exact with the appropriate values of the dispersion tensors (from a deterministic modeling viewpoint).

This general model will reduce to the ADE in the limit of long time and large distances which can be seen by a Taylor series expansion in Fourier-Laplace space. The dispersion tensor \mathbf{D}_2 resembles the classical dispersion tensor, but is a generalization since it is non-constant and shows up in an integral. The term \mathbf{D}_1 plays the role of a nonlocal convective velocity. If these two dispersion tensors depend only on τ and \mathbf{x} , which occurs when mean displacement depends on time lag alone, then Equation 18.4 is called “convolution-Fickian.”

Particle tracking experiments (Section 18.10.1) have been performed to examine this theory and calculate the dispersion tensor when \mathbf{D}_1 is zero. In the large spatial limit (hence long time) the dispersion tensor \mathbf{D}_2 may go to the expected value of the effective dispersion tensor as estimated by the velocity covariance (Figure 18.1). This indicates experimentally that the classical ADE is valid for some media at large scales, but at smaller scales a more general approach is always necessary.

18.7.2 Continuous Time Random Walks

Like the previous method, Continuous Time Random Walk (CTRW) models may be used for highly heterogeneous systems. One can consider the medium as a random network where the cloud of particles executes a series of transitions with variable velocity. Since the CTRW is for highly heterogeneous systems, averaging over an REV is not useful and the effective volume contributing to particle flux (hence change in concentration) can be allowed to vary over the system. The main ideas and connections to the ADE are presented below. Additional details and information on the CTRW can be found in a variety of references (e.g., Scher and Montroll, 1975; Metzler and Klafter, 2000; Berkowitz et al., 2002; Dentz and Berkowitz, 2003; Cortis and Berkowitz, 2004).

A “Master Equation” is used to express the change of concentration in time via the transition rate from one spatial location to another. This master equation represents the transition density for trajectories of

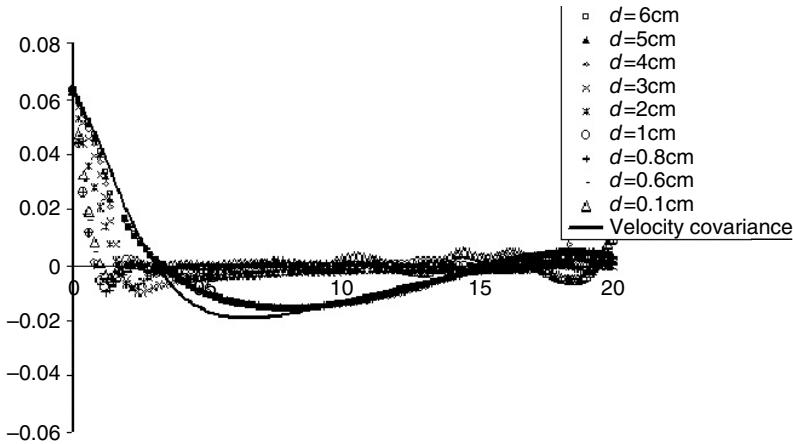


FIGURE 18.1 The generalized dispersion tensor for the transverse direction of the homogeneous porous medium compared to the velocity covariance. Here $k = 2\pi/d$ where k is the magnitude of the wave vector dual to the wave vector for the transverse direction. Increasing d corresponds to increasing the scale of observation for the system. In the Fickian limit the time integral of the velocity covariance is the dispersion tensor. As the spatial scale increases, the generalized dispersion tensor approaches the value of the classical dispersion tensor. (Figure 20 in Moroni and Cushman, 2001, *Phys. Fluids* 13, 81–91.)

a solute particle. An ensemble average over all possible paths is used to develop the Generalized Master Equation (GME):

$$\frac{\partial C(\mathbf{s}, t)}{\partial t} = - \sum_{\mathbf{s}'} \int_0^t \phi(\mathbf{s}' - \mathbf{s}, t - t') C(\mathbf{s}, t) dt' + \sum_{\mathbf{s}'} \int_0^t \phi(\mathbf{s} - \mathbf{s}', t - t') C(\mathbf{s}', t) dt' \quad (18.5)$$

where $C(\mathbf{s}, t)$ is the concentration, ϕ is a function of $\psi(\mathbf{s}, t)$ which is the probability rate per unit time for displacement \mathbf{s} in time t . The GME is equivalent to the CTRW

$$R(\mathbf{s}, t) = \sum_{\mathbf{s}'} \int_0^t \psi(\mathbf{s} - \mathbf{s}', t - t') R(\mathbf{s}', t') dt'$$

with

$$C(\mathbf{s}, t) = \int_0^t \left[1 - \int_0^t \sum_{\mathbf{s}'} \psi(\mathbf{s}, t') dt' \right] R(\mathbf{s}, t') dt' \quad (18.6)$$

where $R(\mathbf{s}, t)$ is the probability a walker arrives at \mathbf{s} at time t . The term in brackets is the probability a walker remains at a fixed site.

The GME does not separate terms based on advection and dispersion, but this can be accomplished by assuming relatively homogeneous regions exist and concentration varies slowly over a finite length scale. Additionally one defines the “effective velocity”, \mathbf{v}_ψ , and “dispersion tensor”, Φ_ψ , as:

$$\mathbf{v}_\psi \equiv \sum_{\mathbf{s}} \phi(\mathbf{s}, t) \mathbf{s} \quad \text{and} \quad \Phi_\psi \equiv \sum_{\mathbf{s}} \phi(\mathbf{s}, t) \frac{1}{2} \mathbf{s} \mathbf{s}^\top \quad (18.7)$$

yielding the form of the GME

$$\frac{\partial C(\mathbf{s}, t)}{\partial t} = \sum_{\mathbf{s}'} \int_0^t dt' \left[\phi(\mathbf{s}' - \mathbf{s}, t - t') (\mathbf{s}' - \mathbf{s}) \cdot \nabla C(\mathbf{s}, t') \right. \\ \left. + \phi(\mathbf{s} - \mathbf{s}', t - t') \frac{1}{2} (\mathbf{s}' - \mathbf{s})(\mathbf{s}' - \mathbf{s}) : \nabla \nabla C(\mathbf{s}, t') \right] \quad (18.8)$$

The colon represents the tensor product. This equation has the same form as Equation 18.4 with the integral in Equation 18.4 replaced by the summation in Equation 18.8. Thus this CTRW is a special case of the generalized hydrodynamics approach. If ψ has finite first and second moments in t , then an expansion can be performed to relate the GME to the ADE. Additionally, it can reduce to equations that are fractional in time or fractional in space (FADE) which will be discussed in the next section. As mentioned, most of these models are variants of the generalized model in Section 18.7.1.

18.7.3 Fractional Advection-Dispersion

The Fractional Advection-Dispersion Equation (FADE) generalizes the ADE and creates nonlocality through the presence of fractional (spatial) derivatives. Although it has been shown that both the generalized hydrodynamics approach (Cushman and Ginn, 2000) and the CTRW approach (Berkowitz and Scher, 2001; Berkowitz et al., 2002) can be simplified using certain assumptions to obtain the FADE, it is worthwhile to treat the FADE (Meerschaert et al., 1999; Benson et al., 2000, 2001) as a separate topic for discussion due to the extensive theoretical, experimental and numerical background that has been developed for this approach. In particular, the one-dimensional FADE has useful closed-form solutions (Benson et al., 2001).

The FADE was developed for advective-dispersive transport in flow fields that are nonuniform on multiple scales. If the tails of a concentration plume are power laws, that is, “heavier” than the standard Gaussian process, anomalous dispersion is taking place and the FADE is often applicable. In multidimensional form, the FADE is given by

$$\frac{\partial C}{\partial t} = -\mathbf{V} \cdot \nabla C + \mathbf{D} \nabla_M^\alpha C \quad (18.9)$$

where \mathbf{V} is the drift velocity, \mathbf{D} is the dispersion tensor that helps describe the spread of the process, and ∇_M^α is the fractional derivative that specifies the weight of the tails of the plume (α) and M , the skewness of the movements (a random measure on the unit sphere). The fractional derivative is defined by its Fourier transform and hence incorporates nonlocality into the model. This can be seen easily in one-dimension, with a uniform M , the fractional derivative has a real space representation:

$$\frac{d^\alpha f(x)}{d(\pm x)} = \frac{(\pm 1)^n}{\Gamma(n - \alpha)} \frac{d^n}{dx^n} \int_0^\infty \xi^{n-\alpha-1} f(x \mp \xi) d\xi \quad (18.10)$$

where n is the smallest integer greater than α . Through this expression one sees that nonlocality is incorporated into the FADE through the fractional derivative. If $\alpha = 2$, the plumes are Gaussian and the fractional derivative with a uniform mixing measure reduces to the standard Laplacian and thus the ADE is recovered. The FADE is closely connected to Lévy flights as the equation can be used to describe the transition density for such a flight.

To model transport, one needs to estimate the order of the fractional derivative and constant \mathbf{D} . Solutions to Equation 18.9 are obtained via Laplace or Fourier transforms. Benson et al. (2001) has examined the FADE for experimental data from the MADE site (Figure 18.2). The data is from large-scale natural gradient tracer tests with strong heterogeneity. Both *a priori* and *a posteriori* parameter estimation

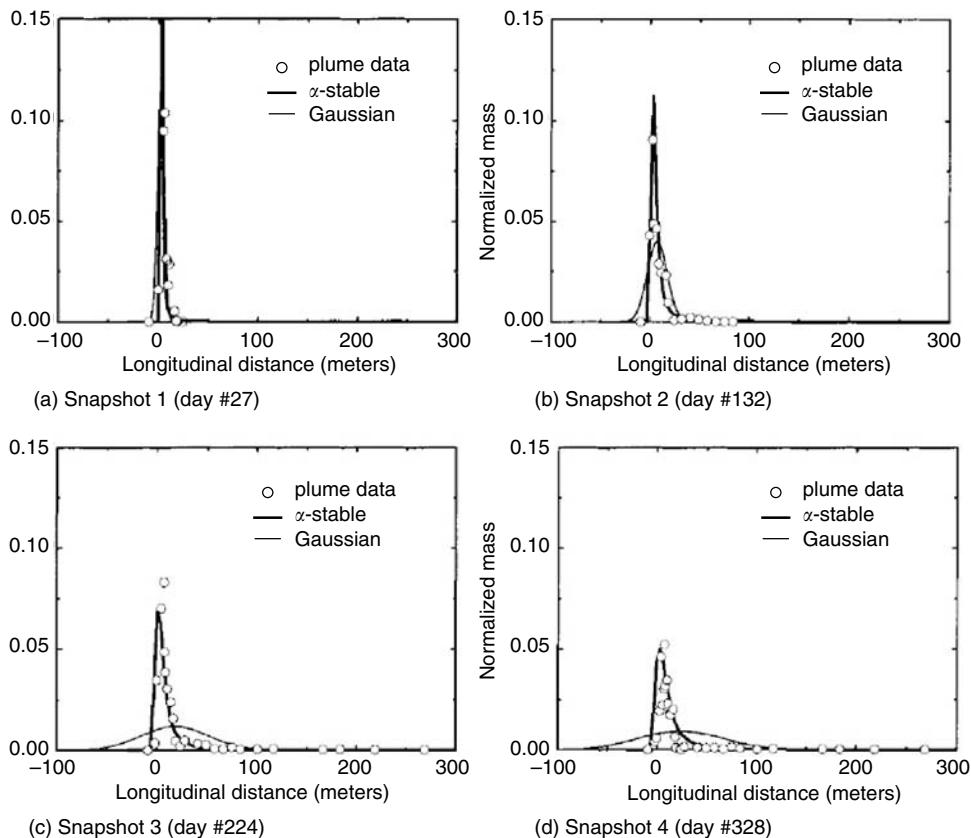


FIGURE 18.2 MADE tracer test plume comparisons compared to the predictions from the FADE(α -stable) and the ADE. The measured heavy tail plumes are much better suited to prediction from the FADE than the ADE. Analytic solutions for the FADE and ADE were obtained by numerical integration. (Figure 4 from Benson et al., 2000, *Water Resour. Res.* 36, 1413–1423.)

were investigated and it was found that although the ADE fails to model the conservative tracer accurately, the FADE concentration profiles are more accurate over all spatial and temporal scales than classical models. An alternate way the parameter α can be determined is from Lagrangian data via the finite size Lyapunov exponent (Kleinfelter et al., 2005).

18.7.4 Stochastic Perturbation

Stochastic tools have been extensively employed to model flow and transport in porous media over the past three decades. While all methods cannot be discussed in detail here, we will highlight some of the main approaches and some simple resulting equations. Stochastic analysis assumes we have incomplete knowledge of the system and can incorporate this uncertainty into the model through probability distributions of material parameters of the system (i.e., conductivity). In stochastic analysis of transport, the concentration is assumed to be a random function of space and time and can be examined spatially in terms of its mean, variance, and higher moments.

18.7.4.1 Fourier Transform Method and Green's Function Approaches

Consider the traditional ADE with C and V random functions,

$$\frac{\partial C}{\partial t} + \nabla \cdot (VC) - \nabla \cdot (\mathbf{D} \cdot \nabla C) = 0 \quad (18.11)$$

and decompose C and \mathbf{V} into their mean and fluctuating parts,

$$C = \bar{C} + c \quad \text{and} \quad \mathbf{V} = \bar{\mathbf{V}} + \mathbf{v}. \quad (18.12)$$

It is assumed the conductivity field is spatially random resulting in a stochastic velocity and thus randomness in C . The overbar, which indicates an average or expectation, is the average over all realizations of porous formations consistent with the specific conductivity distribution.

As a simple example of this method, assume an infinite domain with a diagonal dispersion tensor, constant diagonal elements, and constant mean velocity in the direction of mean flow (call this direction x_1). Using Equation 18.12, the mean for Equation 18.11 can be rewritten as

$$H\bar{C} = -\nabla \cdot \bar{\mathbf{v}} \quad (18.13)$$

and the mean removed equation (Equation (18.12) minus (18.13)) written as

$$Hc = -\mathbf{v} \cdot \nabla \bar{C} - \nabla \cdot (\mathbf{v}c) + \nabla \cdot (\bar{\mathbf{v}}c) \quad (18.14)$$

where the operator H is given by

$$H = \frac{\partial}{\partial t} + \bar{V}_1 \frac{\partial}{\partial x_1} - \nabla \cdot (\mathbf{D} \cdot \nabla) \quad (18.15)$$

with the assumption that the velocity is divergence free $\nabla \cdot \mathbf{v} = 0$. One can use Green's functions to get the equation for the concentration fluctuation:

$$\begin{aligned} c(\mathbf{x}, t) = & \int_0^t \int_{\mathbb{R}^3} B(\mathbf{x} - \mathbf{y}, t - \tau) [-\mathbf{v} \cdot \nabla \bar{C}(\mathbf{y}, \tau) - \nabla \cdot (\mathbf{v}(\mathbf{y}) \bar{C}(\mathbf{y}, \tau))] \\ & + \nabla \cdot (\bar{\mathbf{v}}(\mathbf{y}) c(\mathbf{y}, \tau)) d\mathbf{y} d\tau \end{aligned} \quad (18.16)$$

with

$$B(\mathbf{x}, t) = \prod_{i=1}^3 \exp \left[-\frac{(x_i - \bar{V}_i t)^2}{4D_{ii}t} \right] (4\pi D_{ii}t)^{-1/2}. \quad (18.17)$$

Notice that Equation 18.13 contains a forcing function on the RHS, and Equation 18.16 can be used to estimate this term. Multiply Equation 18.16 by \mathbf{V} and then take the average. Assume that $\partial v_i(\mathbf{y}) v_j(\mathbf{y}) c(\mathbf{y}, \tau) / \partial y_i$ is small compared to the other terms and that the velocity covariance function is stationary. If one further assumes that the gradient of the mean concentration varies slowly relative to the remainder of the integrand, it may be removed and the resulting equation is:

$$\mathbf{q} = \overline{\mathbf{V}(\mathbf{x}, t) c(\mathbf{x}, t)} = - \left[\int_0^t \int_{\mathbb{R}^3} B(\mathbf{y}, t - \tau) \mathbf{R}(\mathbf{y}) d\mathbf{y} d\tau \right] \cdot \nabla \bar{C} = -\mathbf{D}'(t) \cdot \nabla \bar{C} \quad (18.18)$$

Equation (18.18) is called quasi-Fickian and \mathbf{R} is the velocity covariance. The mass balance is $\partial C / \partial t = -\nabla \cdot \mathbf{q}$. Classical Fickian dispersion is obtained when $\mathbf{D}'(t)$ is replaced by $\mathbf{D}'(\infty)$. This is the classical result

of Gelhar et al. (1979) obtained through spectral integral representation. If the flux is convolution-Fickian, then the mass balance takes the form

$$\frac{\partial \bar{C}}{\partial t} + \bar{V}_1 \frac{\partial \bar{C}}{\partial x_1} - \nabla \cdot (\mathbf{D} \cdot \nabla \bar{C}) - \int_0^t \int_{\mathbb{R}^3} \mathbf{D}'(\mathbf{x} - \mathbf{y}, t - \tau) \cdot \nabla_y \bar{C} d\mathbf{y} d\tau = 0 \quad (18.19)$$

where $\mathbf{D}'(\mathbf{x}, t) = B(\mathbf{x}, t)\mathbf{R}(\mathbf{x})$ is the generalized dispersion tensor. Equation 18.18 can be written as an algebraic expression by converting it to Fourier–Laplace space, solved and then numerically inverted back to real space.

The above discussion can be generalized to a finite domain and Equation 18.16 and Equation 18.17 will still be obtained, but boundary conditions can be imposed and the solution for c would look the same, but the B would now be a Green's function with the integral over appropriate domain and include boundary integrals. Neuman has pioneered work in this area (Hsu et al., 1996).

18.7.4.2 Lagrangian Perturbation Approach

Next we discuss a Lagrangian approach pioneered initially by Simmons (1982) and generalized by Dagan (1994) and Dagan et al. (1992). In the Lagrangian perturbation approach, one again assumes the classical ADE holds with steady, divergence free, Darcy velocity and the mean flow aligned in the x_1 direction and add in the condition of no dispersion. Label the trajectory of a fluid ‘particle’ by \mathbf{Y} originating at the spatial point $\mathbf{y} = \mathbf{y}^0$ at $t = t^0$. \mathbf{Y} satisfies

$$\frac{d\mathbf{Y}}{dt} = \mathbf{V}(\mathbf{Y}) \quad (18.20)$$

where $\mathbf{y} = \mathbf{Y}(t; t^0, \mathbf{y}^0) = \mathbf{Y}(t - t^0; \mathbf{y}^0)$. Change to the natural coordinates

$$\tau(\mathbf{y}; \mathbf{y}^0) = t - t^0 \quad \eta(\mathbf{y}; \mathbf{y}^0) = X_2(\tau; \mathbf{y}^0) \quad \xi(\mathbf{y}; \mathbf{y}^0) = X_3(\tau; \mathbf{y}^0) \quad (18.21)$$

with positive velocity in the mean flow direction. The travel time, τ , to reach the control plane $t_1 = X_1(t - t^0; \mathbf{y}^0)$ is obtained by inverting Equation 18.20. The time, τ , is the “first passage time” for a particle to cross a control plane at x_1 . The coordinates (τ, η, ξ) satisfy

$$\begin{aligned} \frac{d\tau}{dy_1} &= \frac{1}{V_1(y_1, \eta, \xi)}; \quad \frac{d\eta}{dy_1} = \frac{d\eta}{d\tau} \frac{d\tau}{dy_1} = \frac{V_2(y_1, \eta, \xi)}{V_1(y_1, \eta, \xi)} \\ \frac{d\xi}{dy_1} &= \frac{d\xi}{d\tau} \frac{d\tau}{dy_1} = \frac{V_3(y_1, \eta, \xi)}{V_1(y_1, \eta, \xi)} \end{aligned} \quad (18.22)$$

subject to $\tau = 0$, $\eta = y_2^0$, and $\xi = y_3^0$ for $y_1 = y_1^0$. Hence for the steady velocity, $y_2 = \eta(y_1; \mathbf{y}^0)$ and $y_3 = \xi(y_1; \mathbf{y}^0)$ are equations of the streamlines. In this coordinate system along the above streamlines,

$$\frac{\partial C}{\partial t} + \frac{\partial C}{\partial \tau} = 0 \quad (18.23)$$

This equation has been solved to give

$$C(\mathbf{y}, t) = \int_{V^0} \frac{C_0(\mathbf{y}^0)}{V_1(y_1, \eta, \xi)} \delta(y_2 - \eta) \delta(y_3 - \xi) d\mathbf{y}^0 \quad (18.24)$$

where the integral is over the region where the initial concentration is nonzero.

To compare the results of the Fourier transform (Green's function) method and the Lagrangian perturbation methods, one can examine the centered spatial moments as long as d_{ij} is set to zero in the former method. The zeroth and first order moments can easily be seen to be equivalent, but the other moments must be evaluated numerically. If a model for fluctuating conductivity covariance is assumed, then comparisons can be made. It was found that at least up through the third moments the two methods agree. A more detailed discussion of the comparison can be found in Cushman (1997, Chapter 6).

18.8 Locally Stationary Systems

Many of the stochastic approaches require simplifying assumptions such as a stationary hydraulic conductivity field, steady state flow with no boundaries, or uniform mean velocity in space. Violation of any of these may lead to nonstationary groundwater flow. A novel method of moments was first developed by Lee and Cushman (1993) for inverse problems and subsequently for forward problems by Hu and coworkers (Wu et al., 2003a, 2003b; Zhang et al., 2000) for flow and solute transport in non-stationary, heterogeneous media with deterministic local boundaries. In the forward method, solute flux statistics are derived through Lagrangian perturbation and are expressed in term of the pdfs (probability density functions) of the particle travel time and transverse displacement. The pdfs are evaluated through the first two moments of travel time and transverse displacement using the assumed solute pdf form. Due to the complexity of the governing equations and the boundary initial conditions, numerical methods are used to solve the problem and the efficiency is greater than from Monte Carlo numerical simulation.

We follow the discussion of Cushman et al. (2002). Consider incompressible, saturated groundwater flow in an aquifer with spatially variable hydraulic conductivity, $K(x)$. The groundwater seepage velocity $\mathbf{V}(x)$ satisfies Darcy's Law, $\mathbf{V}(x) = -(K/n)\nabla h$ where h is the hydraulic head, and n is the effective porosity. $K(x)$ is assumed piecewise stationary because of small-scale boundaries, but globally nonstationary, and this in turn causes \mathbf{V} to be spatially nonstationary. A solute is injected with mass, M , with a known release rate $\phi(t)$ over an area A_0 . Let $\rho_0(\mathbf{a})$ be the area density of the solute mass at \mathbf{a} in A_0 . For $t > 0$, the solute plume is advected toward a (y,z) -control plane (CP). The (nonreactive) solute flux orthogonal to the CP at x over a sampling area $A(y)$ centered at $y(y,z)$ is

$$q(t, \mathbf{y}; x, A) = \frac{1}{A} \int_{A_0} \int_A \rho_0(\mathbf{a}) \phi(t - \tau) \delta(\mathbf{y}' - \boldsymbol{\eta}) d\mathbf{y}' d\mathbf{a} \quad (18.25)$$

Here $\tau \equiv \tau(x; \mathbf{a})$ is the travel time of the advective particle from \mathbf{a} to the control plane at x and $\boldsymbol{\eta} \equiv (\eta(x; \mathbf{a}), \xi(x; \mathbf{a}))$ is the transverse location of a particle passing through the control plane. τ and $\boldsymbol{\eta}$ in Equation 18.25 are random variables that are functions of the underlying velocity field. The expected value in the case of point sampling ($A \rightarrow 0$) is

$$\langle q(t, (x, \mathbf{y})) \rangle = \int_{A_0} \rho_0(\mathbf{a}) \phi(t - \tau) f_1[\tau = t, \boldsymbol{\eta} = \mathbf{y}] d\tau d\mathbf{a} \quad (18.26)$$

where $f_1[\tau, \boldsymbol{\eta}]$ is the joint pdf of a travel time τ for a particle from \mathbf{a} to reach x and the corresponding transverse displacement is $\boldsymbol{\eta}$ (recall τ and $\boldsymbol{\eta}$ are functions of x and \mathbf{a}). The average variance of the solute flux can be expressed with a two-dimensional joint density function. The averages of τ and $\boldsymbol{\eta}$ can be written as functions of the mean Eulerian velocities. Due to the nonstationarity of flow, a numerical approach needs to be implemented to obtain the means and covariances of τ and $\boldsymbol{\eta}$. If τ and $\boldsymbol{\eta}$ obey log-normal and normal distributions, then the joint pdf for one and two particle statistics needed for mean solute flux and covariance representations can be obtained. The input for the model is the hydraulic conductivity distributions specified by the means, variances and correlation lengths for various models

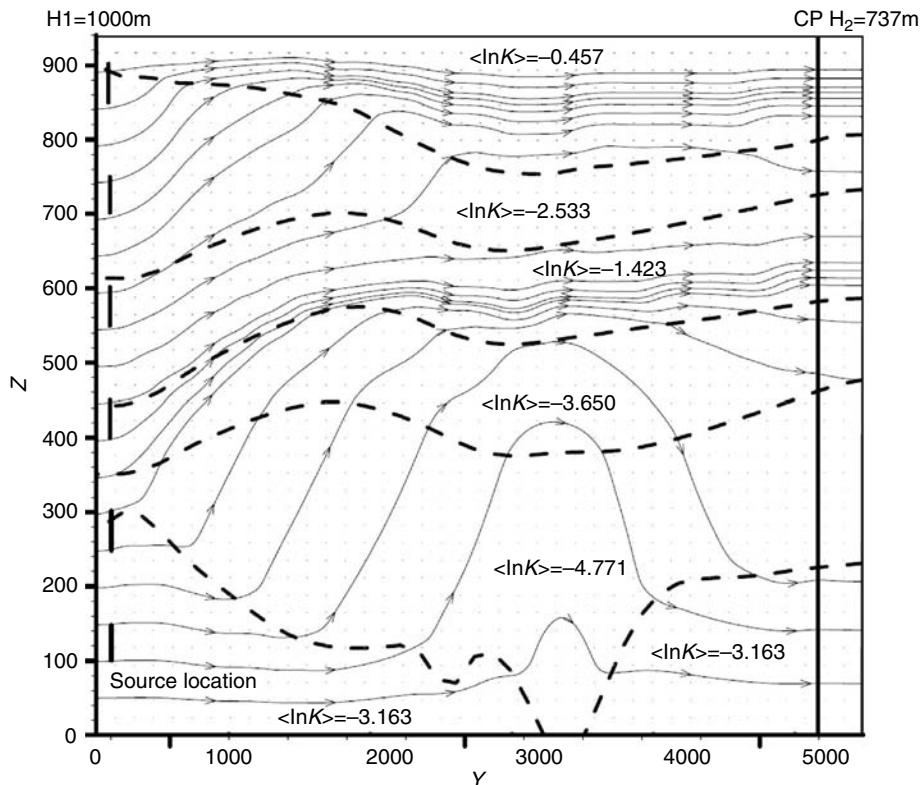


FIGURE 18.3 Nonstationary conductivity field for Yucca Mountain used in the stochastic convective nonstationary approach results shown in Figure 18.4. The six different source lines used are close to the left boundary with the control plane near the right boundary. The mean of the log of conductivity is shown in each layer and dashes represent the layers in the medium. Mean flow lines are shown and indicate the main flow is through the high-conductivity layers. More detail on the set-up and computational time can be found in Cushman et al. (2002). (Figure 1 from Cushman et al., 2002, *Adv. Water Resour.* 25, 1043–1067.)

(i.e., log-normal) and local boundary data. A sample conductivity field is shown in Figure 18.3 with the results of the mean solute discharge and variance for such a field shown in Figure 18.4.

A more general method for globally nonstationary, but piecewise stationary solute transport has been developed by Winter and Tartakovsky (2000) and Winter et al. (2002) by including random local boundaries. It is not likely that the exact boundaries for the changes in hydraulic conductivity will be known so the boundaries having uncertainties makes the model more realistic.

18.9 Spreading vs. Dilution

Since anomalous dispersion has become more widely modeled and evidenced in experiments, the difference between spreading and dilution has also become of interest. Spreading is connected to the stretching and deformation of the contaminant plume while dilution is connected to the change in the volume of fluid the solute occupies (Kitanidis, 1994). When dispersion coefficients are estimated, they give a macroscopic measure of the rate of spreading which may not necessarily correspond to rate of dilution. The peak concentration is one measure of the dilution since if the peak concentration is lower, the solute is more dilute overall. Kitanidis (1994) further quantified the notion of dilution by introducing the concepts of the “dilution index” and “reactor ratio” which are used to measure how uniformly the solute is dispersed in the plume.

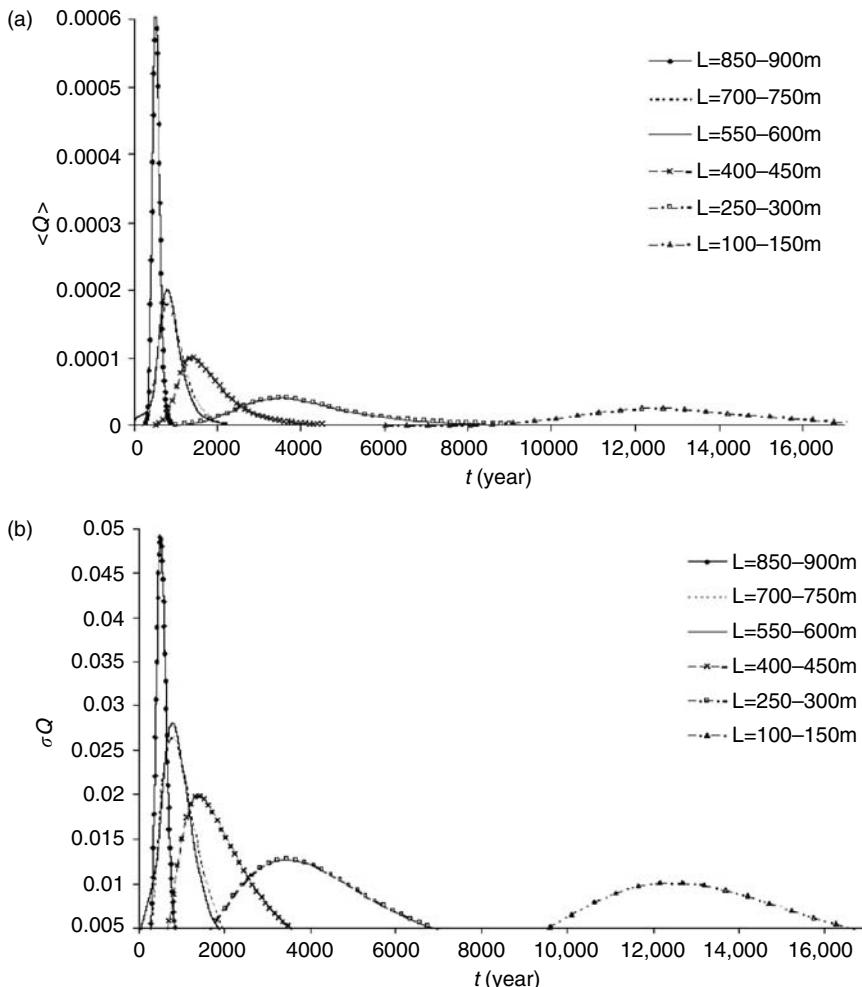


FIGURE 18.4 Q is the total solute flux over the control plane. (a) Expected solute breakthrough curves with the source line at different locations. (b) Variances about the mean predictions. (Figure 2 from Cushman et al., 2002, *Adv. Water Resour.* 25, 1043–1067.)

Consider the normalized concentration, $C(\mathbf{x}, t)$. The dilution index $E(t)$ is an entropy-like measure defined by

$$E(t) = \exp \left[- \int_V C(\mathbf{x}, t) \ln C(\mathbf{x}, t) dV \right] \quad (18.27)$$

The theoretical maximum of this quantity can be considered for systems and the dilution index compared to this maximum quantity, yielding the reactor ratio. The dilution index and reactor ratio have been thoroughly explored for unbounded domains via the advection-dispersion equation with constant coefficients as well as other systems. In addition, $E(t)$ has been related to the concentration variance (Kitanidis, 1994; Kapoor and Kitanidis, 1997; Pannone and Kitanidis, 1999).

Kapoor and Gelhar (1994) examined spreading vs. dilution in the Cape Cod bromide tracer data via peak concentration. They found that the macroscale dispersion tensor was not a very accurate predictor of the peak concentration (their measure of dilution) and that, at best, the macroscale dispersion tensor

can be used to describe the plume spread in terms of the second moment of concentration. Additionally, it was found through numerical simulations in a heterogeneous conductivity field that an n -fold increase in the covariance of the concentration in the mean flow direction does not result in an n -fold decrease in the peak concentration as would occur in a homogeneous conductivity field. A hundred fold decrease in covariance resulted in only a tenfold decrease in peak concentration. This indicates in heterogeneous aquifers that care should be taken in contamination transport in relating dilution and spreading.

18.10 Experiments

18.10.1 Lab Experiments

The majority of laboratory experiments involve the analysis from an Eulerian perspective by looking at solute concentration (breakthrough curves), porosity, pore-velocity, and hydraulic conductivity and employ a Fickian estimate for the dispersion tensors by using the velocity covariance (Harleman and Rumer, 1963; Klotz et al., 1980; Eidsath et al., 1983; Rashidi et al., 1996). An alternative method that allows greater freedom in analysis and more detailed information is Particle Tracking Velocimetry (PTV).

Particle Tracking Velocimetry has been utilized by Moroni and Cushman (2001) in the investigation of the generalized hydrodynamic approach (Section 18.7.1) and to study the validity of the ADE. PTV has the benefit that Lagrangian trajectories of the passive tracer are obtained for analysis. In a time stationary system one can then convert this information to an Eulerian frame.

To utilize PTV in a porous medium the matrix must be made invisible. This can be accomplished using a fluid and matrix that have matching refractive indices (matched-index system). One naturally occurring system is cryolite and water, however to develop a laboratory system another possibility is glycerol and Pyrex at 23°C in natural light. Moroni and Cushman (2001) used Pyrex beads and cylinders to create both homogeneous and heterogeneous media. Air bubbles, of an appropriate size, were used as the passive tracer (they have a different refractive index than the glycerol and Pyrex) and the flow of the bubbles in the glycerol was tracked with different flow velocities in the various media. Through reconstruction techniques, three-dimensional (3D) trajectories of the air bubble paths were followed and analyzed. A sample of 3D trajectories from the “homogeneous” medium is shown in Figure 18.5. It was found that all media have anomalous flow characteristics at some scale, even “homogeneous” media, (see Figure 18.5) and thus reinforces the notion that homogeneity is a scale dependent concept.

18.10.2 Field Experiments

Many large scale field experiments have been performed over the past several decades to examine and compare theories of macrodispersion. The largest such natural gradient experiments are the Borden, Cape Cod, and the MADE tests. A comprehensive discussion of field-scale dispersion experiments can be found in Gelhar (1992).

The Borden aquifer, located in Canada, is a sand quarry with medium and fine-grained sands. Chloride and bromide were used as the conservative tracers. Discussions of this test can be found in Sudicky (1986) and Freyberg (1986). Scale-dependent dispersion is present in this experiment and has mainly been studied using the ADE. The Cape Cod (Massachusetts) site is a sand and gravel aquifer with a low level of heterogeneity and both reactive and nonreactive tracers were used. The Cape Code site has been analyzed with stochastic theories and the first and second moments of the tracer plumes have fit well with these theories (Garabedian et al., 1991; Kapoor and Gelhar, 1994), but these theories have not held up well in the more heterogeneous case of the MADE site. The concept of spreading vs. dilution has also been investigated for the Cape Cod site through the covariances. The MADE site (Mississippi) is substantially more heterogeneous than the other sites and also involved the use of reactive and nonreactive tracers (Adams and Gelhar, 1992; Boggs et al., 1992; Benson et al., 2000). The FADE has been studied relative to this data as discussed in Section 18.7.3. Further discussion on the MADE site can be found in Chapter 26 (C. Zheng).

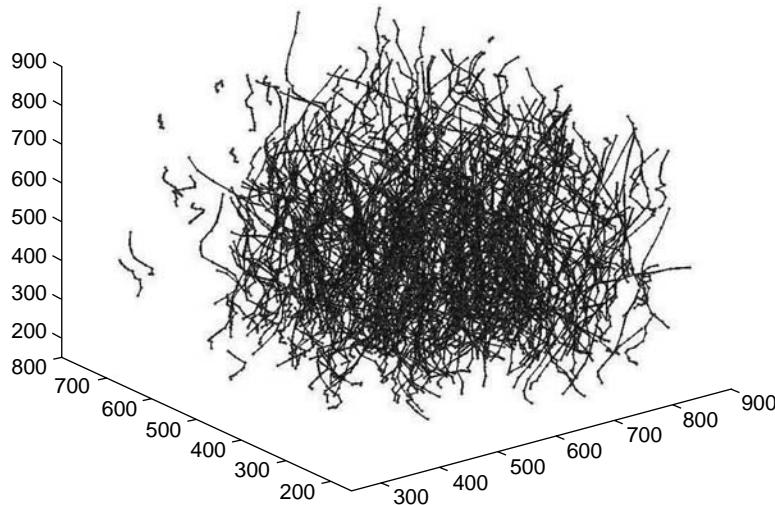


FIGURE 18.5 Example of trajectories obtained via Particle Tracking velocimetry through a bench scale porous medium. Trajectories are of length greater than 12 sec. (Figure 7 in Moroni and Cushman, 2001, *Phys. Fluids* 13, 81–91.)

These large-scale field experiments are providing a basis for the testing of new theories and a better understanding of when classical theories break down; it is expected that they will continue to play an important role in the modeling of transport in porous media.

Glossary

Passive tracers Solute that does not react with the porous medium or the suspending solution.

Conservative tracer Solute that maintains its mass in the fluid volume.

Heterogeneity The occurrence of variability of the values of field properties in space.

Stochastic model A mathematical model that involves spatially or temporally material parameters or a random forcing function.

Structural vs. functional hierarchy Structural hierarchy refers to nested physical subunits of the medium whereas functional hierarchy refers to nested processes.

Discrete vs. continuous hierarchy Discrete hierarchy involves a finite number of nested subunits (subprocesses) whereas in a continuous hierarchy the number of such subunits goes to infinity.

Pore scale The scale of the pore. The pore velocity is the velocity measured at this scale.

Darcy scale A scale large enough so that Darcy's law holds. The Darcy velocity is the velocity used at the Darcy scale.

Hydraulic conductivity The constant of proportionality in Darcy's law between the Darcy velocity and the gradient of the hydraulic head.

Macrodispersivity tensor The dispersivity tensor in the ADE when it is used on a Darcy scale or higher.

Pathline A particles trajectory in time.

Nonlocal constitutive law One that involves integral or higher order gradients.

Covariance The average of the product of a random variable at two different spatial points (example velocity covariance $C_v(x, y) = E[v(x)v(y)]$).

Steady flow A flow where the properties do not depend explicitly on time.

Stationarity May be temporal or spatial. Refers to dependence of a property only on the time lag or spatial displacement instead time and location. In a random variable the most common types are

strict stationarity (statistically homogeneous) and second order or weak stationarity. Strict stationarity means that the probability density function is invariant to spatial shifts. In weak stationarity the mean and variance are constant and the covariance depends only on the displacement vector between the two points and is symmetric.

Probability density function Function describing the probability of event. The sum (or integral) over all possible events must be 1.

Acknowledgments

This work was funded under the support of the National Science Foundation under Contract Nos 0310029-EAR, 0417555-EAR, and 0003878-EAR.

References

- Adler, P.M. 1992. *Porous Media: Geometry and Transports*. Butterworth/Heinemann.
- Adams, E.E. and Gelhar, L.W. 1992. Field study of dispersion in a heterogeneous aquifer 2. Spatial moments analysis. *Water Resour. Res.* 28, 3293–3307.
- Bear, J. 1972. *Dynamics of Fluids in Porous Media*, Elsevier.
- Benson, D.A., Schumer, R., Wheatcraft, S.W., and Meerschaert, M.M. 2001. Fractional dispersion, Levy motion and the MADE tracer tests. *Transp. Porous Media* 42, 211–240.
- Benson, D.A., Wheatcraft, S.W., and Meerschaert, M.M. 2000. The fractional-order governing equation of Levy motion. *Water Resour. Res.* 36, 1413–1423.
- Berg, H.C. 2000. Motile behavior of bacteria. *Phys. Today* 53, 24–29.
- Berkowitz, B., Klafter, J., Metzler, R., and Scher, H. 2002. Physical pictures of transport in heterogeneous media: Advection-dispersion, random-walk, and fractional derivative formulations. *Water Resour. Res.* 38, Art 1191.
- Berkowitz, B. and Scher, H. 2001. The role of probabilistic approaches to transport theory in porous media. *Transp. Porous Media* 42, 241–263.
- Boggs, J.M., Young, S.C., Beard, L.M., Gelhar, L.W., Rehfeldt, K.R., and Adams, E. 1992. Field study of dispersion in a heterogeneous aquifer 1. Overview and site description. *Water Resour. Res.* 28, 3281–3291.
- Bonilla, F.A. and Cushman, J.H. 2002. Effect of alpha-stable sorptive waiting times on microbial transport in microflow cells. *Phys. Rev. E* 66, 031915.
- Bouleau, N. and Leipingle, D. 1994. *Numerical Methods for Stochastic Processes*. Wiley, New York.
- Brenner, H. and Edwards, D.A. 1993. *Macrotransport Processes*. Butterworth-Heinemann, MA.
- Cortis, A. and Berkowitz, B. 2004. Anomalous transport in “classical” soil and sand columns. *Soil Sci. Soc. Am. J.* 68, 1539–1548.
- Creswick, R.J., Farch, H.A., and Poole, C.P. Jr. 1992. *Introduction to Renormalization Group Methods in Physics*. Wiley, New York.
- Cushman, J.H. (Ed.) 1990. *Dynamics of Fluids in Hierarchical Porous Media*. Academic Press.
- Cushman, J.H. 1997. *The Physics of Fluids in Hierarchical Porous Media: Angstroms to Miles*. Kluwer.
- Cushman, J.H. and Ginn, T.R. 1993. Nonlocal dispersion in media with continuously evolving heterogeneity. *Transp. Porous Media*, 13, 123–138.
- Cushman, J.H. and Ginn, T.R. 2000. Fractional advection-dispersion equation: A classical mass balance with convolution-Fickian flux. *Water Resour. Res.* 36, 3763–3766.
- Cushman, J.H., Bennethum, L.S., and Hu, B.X. 2002. A primer on upscaling tools for porous media. *Adv. Water Resour.* 25, 1043–1067.
- Dagan, G. 1989. *Flow and Transport in Porous Formations*. Springer-Verlag, New York.
- Dagan, G. 1994. An exact nonlinear correction to transverse macropdispersivity for transport in heterogeneous formations. *Water Resour. Res.* 30, 2699–2705.

- Dagan, G., Cvetkovic, V., and Shapiro, A. 1992. A solute flux approach to transport in heterogeneous formations, 1. The general framework. *Water Resour. Res.* 28, 1369–1376.
- Dentz, M. and Berkowitz, B. 2003. Transport behavior of a passive solute in continuous time random walks and multirate transfer. *Water Resour. Res.* 39, Art. 1111.
- Eidsath, A., Carbonell, R.G., Whitaker, S., and Herrmann, L.R. 1983. Dispersion in pulsed systems-III. Comparison between theory and experiments for packed beds. *Chem. Eng. Sci.* 38, 1803–1816.
- Freyberg, D. 1986. A natural gradient experiment on solute transport in a sand aquifer, 2. Spatial moments and the advection and dispersion of nonreactive tracers. *Water Resour. Res.* 22, 2031–2046.
- Garabedian, S.P., LeBlanc, D.R., Gelhar, L.W., and Celia, M.A. 1991. Large-scale natural gradient tracer test in sand and gravel, Cape Cod, Massachusetts 2. Analysis of spatial moments for a nonreactive tracer. *Water Resour. Res.* 27, 911–924.
- Gelhar, L.W., Gutjahr, A.L., and Naff, R.L. 1979 Stochastic analysis of macrodispersion in a stratified aquifer. *Water Resour. Res.* 15, 1387–1397.
- Gelhar, L.W. 1993. *Stochastic Subsurface Hydrology*. Prentice-Hall, NJ.
- Gelhar, L.W. 1986. Stochastic subsurface hydrology from theory to application. *Water Resour. Res.* 22, S135–S145.
- Gelhar, L.W. 1992. A critical review of data on field scale dispersion in aquifers. *Water Resour. Res.* 28, 1955–1974.
- Harleman, D. and Rumer, R.R. 1963. Longitudinal and lateral dispersion in an isotropic porous medium. *J. Fluid Mech.* 16, 385–394.
- Hsu, K.C., Zhang, D.X., and Neuman, S.P. 1996. Higher-order effects on flow and transport in randomly heterogeneous media. *Water Resour. Res.* 32, 571–582.
- Kapoor, V. and Gelhar, L.W. 1994. Transport in three-dimensionally heterogeneous aquifers: 2. Predictions and observations of concentration fluctuations. *Water Resour. Res.* 30, 1789–1801.
- Kapoor, V. and Kitanidis, P. 1997. Advection–diffusion in spatially random flows: Formulation of concentration covariance. In *Stochastic Hydrology and Hydraulics*, vol. 11, pp. 397–422, Springer-Verlag.
- Kitanidis, P. 1994. The concept of the dilution index. *Water Resour. Res.* 30, 2011–2026.
- Kleinfelter, N., Moroni, M., and Cushman, J.H. 2005. Application of a finite size Lyapunov exponent to particle tracking velocimetry in fluid mechanics experiments. *Phys. Rev. E* 72(5), Art. 056306, Part 2.
- Klotz, D., Selier, K.-P., Moser, M., and Neumaier, F. 1980. Dispersivity and velocity relationship from laboratory and field experiments. *J. Hydrol.* 45, 169–184.
- Lee, K.K. and Cushman, J.H. 1993. Multiscale adaptive estimation of the conductivity field using pump test data. *Stoch. Hydrol. Hyd.* 7, 241–254.
- Liu, Y.P., Steenhuis, T.S., and Parlange, J.Y. 1994. Closed-form solution for finger width in sandy soils at different water contents. *Water Resour. Res.* 30, 149–952.
- Meerschaert, M., Benson, D.A., and Bäumer, B. 1999. Multidimensional advection and fractional dispersion. *Phys. Rev. E* 59, 5026–5028.
- Metzler, R. and Klafter, J. 2000. The Random walk's guide to anomalous diffusion: A fractional dynamics approach. *Phys. Rep.* 339, 1–77.
- Moroni, M. and Cushman, J.H. 2001. Statistical mechanics with three-dimensional particle tracking velocimetry in the study of anomalous dispersion. 1. Experiments, *Phys. Fluids* 13, 81–91.
- Pannone, M. and Kitanidis, P. 1999. Large-time behavior of concentration variance and dilution in heterogeneous formations. *Water Resour. Res.* 35, 623–634.
- Quintard, M. and Whitaker, S. 1987. Ecoulement monophasique en milieu poreux: Effect des hétérogénéités locales. *J. Mec. Theo. Appl.* 6, 691–726.
- Rashidi, M., Peurrung, L., Tompson, A.F.B., and Kulp, T.J. 1996. Experimental analysis of pore-scale flow and transport in porous media. *Adv. Water Resour.* 19, 163–180.
- Sanchez-Palencia, E. 1980. *Non-Homogeneous Media and Vibration Theory*. Lecture Notes in Physics, 127, Springer-Verlag, New York.

- Scher, H. and Montroll, E.W. 1975. Anomalous transit time dispersion in amorphous solids. *Phys. Rev. B.* 12, 2455–2477.
- Selker, J.S., Steenhuis, T.S., and Parlange, J.Y. 1996. An engineering approach to fingered vadose pollutant transport. *Geoderma* 70, 197–206.
- Shapiro, H. and Brenner, H. 1988. Dispersion of a chemically reactive solute in a spatially periodic model of a porous medium. *Chem. Eng. Sci.* 43, 551–571.
- Simmons, C.S. 1982. A stochastic-convective transport representation of dispersion in one-dimensional porous-media systems. *Water Resour. Res.* 18, 1193–1214.
- Sudicky, E.A. 1986. Natural gradient experiment on solute transport in a sand aquifer: Spatial variability of hydraulic conductivity and its role in the dispersive process. *Water Resour. Res.* 22, 2069–2082.
- Wheatcraft, S.W. and Tyler, S.W. 1988. An explanation of scale dependent dispersivity in heterogeneous aquifers using concepts of fractal geometry. *Water Resour. Res.* 24, 566.
- Winter, C.L. and Tartakovsky, D.M. 2000. Mean flow in composite porous media. *Geophys. Res. Lett.* 27, 1759–1762.
- Winter, C.L., Tartakovsky, D.M., and Guadagnini, A. 2002. Numerical solutions of moment equations for flow in heterogeneous composite aquifers. *Water Resour. Res.* 38, 1055.
- Wu, J.C., Hu, B.X., and Zhang, D.X. 2003a. Applications of nonstationary stochastic theory to solute transport in multi-scale geological media. *J. Hydrol.* 275, 208–228.
- Wu, J.C., Hu, B.X., Zhang, D.X., and Shirley, C. 2003b. A three-dimensional numerical method of moments for groundwater flow and solute transport in a nonstationary conductivity field. *Adv. Water Resour.* 26, 1149–1169.
- Zhang, D.X. 2002. *Stochastic Methods for Flow in Porous Media: Coping with Uncertainties*. Academic Press.
- Zhang, D.X., Andricevic, R., Sun, A.Y., Hu, B.X., and He, G.W. 2000. Solute flux approach to transport through spatially nonstationary flow in porous media. *Water Resour. Res.* 36, 2107–2120.

Further Information

We have presented a brief outline of some of the methods and resources that have been employed in the study of conservative tracers. Zhang (2002) provides a good introduction to stochastic methods for the flow problem. An introduction to stochastic numerical methods can be found in Bouleau and Lépingle (1994). The books of Adler (1992), Brenner and Edwards (1993), Cushman (1990, 1997), Dagan (1989), and Gelhar (1993) are useful references on transport problems.

19

Reactive Contaminant Transport in the Saturated Zone: Review of Some Upscaling Approaches

19.1	Introduction.....	19-2
19.2	Physical, Chemical, and Microbiological Heterogeneity	19-4
19.3	Scale of Observation and Constitutive Concepts	19-5
19.4	Upscaling Methods for Chemical Heterogeneity	19-5
19.5	Linearity and Nonequilibrium Eulerian Stochastic Models	19-7
	Deterministic Nonequilibrium Sorption • Random Nonequilibrium Sorption and Equilibrium Microbial Decay • Numerical Results for Random K_d and No Degradation • Numerical Results with Random Chemical and Microbiological Conditions	
19.6	Upscaling Microbial Dynamics	19-11
19.7	Lagrangian Stochastic Models for Nonlinear Reactions	19-13
	Streamtube Formulation • Stochastic–Convective Averaging • Travel-Time Distribution Function • Some Reaction System Examples • Toward Handling Heterogeneity in Reactive Properties	
19.8	Connection between Approaches.....	19-22
	Definition of Symbols	19-23
	Glossary.....	19-24
	Acknowledgments.....	19-25
	References	19-25
	Further Reading.....	19-29

J. H. Cushman
Purdue University

T. R. Ginn
University of California at Davis

19.1 Introduction

We extend the discussion of the previous chapter to a setting wherein the dissolved chemical reacts with its environment. The catalogue of natural geochemical reactions and transformations associated with man-made contaminants that have entered the saturated zone is essentially limitless, but for the most part these reactions can be categorized according to speed, reversibility, and phase association. A relevant subset of the classification scheme of Rubin (1983) includes the four cases (1) reactions that are fast (equilibrium) and reversible involving aqueous–aqueous or (2) aqueous–solid phase interactions; and (3) reactions that are not fast (nonequilibrium) involving aqueous–aqueous or (4) aqueous–solid phase interactions. The terms “fast” and “not fast” are defined relative to the other transformation processes involved, including other reactions and convective and dispersive transport on the observation time scale. Aqueous–aqueous (Categories 1 and 3) are termed “homogeneous” and aqueous–nonaqueous (2 and 4) are termed “heterogeneous” (regarding phase; not to be confused with spatial variability). This scheme is useful because there is a distinct mathematical approach for each of these categories. The chemical mechanisms driving general reactions and the mass balance approaches for formulating representative transformation models are beyond the present scope, and are summarized at various levels in for instance Rubin (1983), Fetter (1993), Marsily (1986), Sposito (1994), and Lichtner et al. (1996). Here we will be restricted to some of the simpler (but representative) reaction mechanisms in categories 2 (sorption) and 4 (biodegradation), and focus on scaling issues associated with transport in heterogeneous porous media. We begin with a brief summary of reaction-transformation formulations and then describe concepts of heterogeneity, scale, and upscaling. The remainder of the chapter describes current methods for scaling plume behavior involving linear reactions (Eulerian approaches), and for scaling contaminant arrival behavior involving nonlinear reactions (Lagrangian approaches).

We start with a simple illustrative example of an equilibrium heterogeneous reaction, involving the single electrolyte NaCl. Dissolved chloride is a simple anion and generally moves in accordance with the local pore-water velocity, much as the water molecules themselves do (with some exception due to the repulsive electrostatic potential between the anion and the mineral surfaces in, for example, quartzitic or kaolinitic materials). Thus chloride is often reasonably treated as an inert or “passive” tracer using methods described in the previous chapter. Sodium, however, tends to be both transported and exchanged with other cations on the negatively charged mineral surfaces. That is, aqueous sodium reacts with the solid phase by cation-exchange, one of a number of mechanisms by which aqueous species may associate with the mineral surfaces. Other mechanisms include precipitation (e.g., Novak and Sevoulian, 1993), surface complexation (e.g., Dzombak and Morel, 1990), hydrophobic sorption of organic species to organic solids (Karickhoff, 1983), colloid filtration (Tien, 1976), and intergranular diffusion (Wood et al., 1990).

Reaction transformations are formally captured via augmenting the basic solute transport rate equation (Equation 13.1 of the previous chapter) with a generalized mass loss rate, $nr = [\partial C / \partial t]_{\text{transformation}}$, here written as mass transfer to the solid phase, with constant porosity n :

$$\frac{\partial C}{\partial t} + \frac{\partial(V_i C)}{\partial x_i} - \frac{\partial}{\partial x_i} \left(d_{ij} \frac{\partial C}{\partial x_j} \right) = \left(\frac{\partial C}{\partial t} \right)_{\text{transformation}} \equiv -\frac{\rho_b}{n} \frac{\partial S}{\partial t} \quad (19.1a)$$

where C is the aqueous concentration [M/L^3], S is the sorbed concentration per unit mass of solid [M/M], ρ_b is the bulk density of the aquifer material [M/L^3], n is the porosity, and d_{ij} is the symmetric dispersion tensor.

A majority of sorption and other heterogeneous transformations may be represented by specifying this mass transformation rate as a function of the departure of the system from an equilibrium state. The first-order form is

$$\frac{\partial S}{\partial t} = K_r(K_d C - S) \quad (19.1b)$$

where K_r is the desorption rate [1/T] and K_d is the equilibrium distribution coefficient [L³/M]. Equation 19.1b specifies a linear nonequilibrium (“not fast”) reversible mass transformation rate that approximates numerous sorption processes (of category 4) in natural media. Alternatively, one may express the right-hand side of Equation 19.1b as $[K_f C - K_r S]$, where $K_f = K_r K_d$ is the forward sorption rate [1/T]. Reactive solutes undergoing transport according to Equation 19.1 exhibit attenuation. If K_r is very large so that Equation 19.1b responds much more quickly than local conditions change (due to other mass fluxes in Equation 19.1a), then the “local equilibrium assumption” is invoked and the chemical partitioning is taken to occur instantaneously. That is, ironically, the reaction rate in Equation 14.1b is so fast that it occurs over a negligibly short time, and literally most of the time, Equation 14.1b is zero. Note that this is not to be confused with assuming steady state conditions for S ; S will change, instantaneously with changes in C . We may then use the equilibrium solution to Equation 19.1b, $S = K_d C$, as an algebraic relation for S in terms of C in the derivative on the right-hand side of Equation 19.1a (to express the dependence of S on C). Differentiating and grouping terms gives from Equation 19.1a

$$\frac{\partial C}{\partial t} + \frac{1}{R} \frac{\partial(V_i C)}{\partial x_i} - \frac{1}{R} \frac{\partial}{\partial x_i} \left(d_{ij} \frac{\partial C}{\partial x_j} \right) = 0 \quad (19.2)$$

where $R = 1 + \rho_b K_d / n$ is known as the retardation factor. When transport is governed by an equation involving equilibrium reactions such as Equation 19.2, the solute(s) are retarded (as opposed to attenuated), as is clear upon recognizing that R reduces the transport parameters V_i and d_{ij} .

Equation 19.1 and Equation 19.2 illustrate the simplest case of heterogeneous mass transformations of the nonequilibrium and equilibrium types. This example illustrates the different mathematical approaches for the two cases. For nonequilibrium reactions we specify a (“kinetic”) function for

$$\frac{\partial S}{\partial t} \equiv r(S, C)$$

and for equilibrium we specify the algebraic relation (termed “isotherm”) of the form $S = r(C)$, so that

$$\frac{\partial S}{\partial t} \equiv r'(C) \frac{\partial C}{\partial t}$$

and $r(C)$ becomes involved in the transport operator. More complex forms of r arise when the mass transformations are nonlinear in C and in S , when multiple processes are in effect, when the nonaqueous phase (for S) undergoes transformations, and when multiple interacting solutes are involved, to note only a few cases.

In the case of multiple interacting solutes and solid-phase species, an evolution equation is written for each component, and these equations are coupled through the (multiple) reaction terms appearing. Such systems present special issues requiring attention in their solution depending on the character of the reactions (cf. review in Yeh and Tripathi, 1989). When all interactions are at equilibrium, the transport equations are coupled with a set of algebraic equations relating mobile to immobile species that complicate the solution via integration schemes designed for the transport operator alone. Recent approaches use iterative solutions between the algebraic and differential equations (Yeh and Tripathi, 1989) or a decoupling of the system (Rubin, 1990). In the other limiting case where all the reactions have a finite rate, then a sink term appears for each species in the evolution equation and the solution can proceed conventionally through operator-splitting (e.g., Chiang et al., 1990; Valocchi and Malmstead, 1992) or explicit integration schemes (e.g., Widdowson et al., 1988; Celia et al., 1989). Finally, a mixture of both nonequilibrium and equilibrium transformations results in a “differential-algebraic equation” (DAE) system that may be solved, under particular conditions regarding the “index” of the system (Gear and Petzold, 1984; Gear, 1988), through specialized iterative or noniterative decoupling approaches (e.g., Steefel and MacQuarrie, 1996).

We now turn our attention to a demonstration of some of the available approaches for dealing with reactions in the presence of physical and chemical heterogeneities in natural media. The goal of the

scaling approaches is to derive maximum information about the solute fate and transport when data on the spatial variability of the natural media is limited. We describe two such approaches, an Eulerian stochastic-analytic method that is useful for describing expected plume behavior under linear reactions, given statistical (correlation) properties of the spatially variable aquifer flow and reaction characteristics; and a Lagrangian-streamtube approach useful for predicting solute arrivals at observation locations (as opposed to plume evolution) when the reactions involve nonlinear transformations. Beyond the current scope is a third approach that is quite useful when available aquifer data includes conditioning data on the flow and reaction property values on significant scales, such as identification and incorporation of hydrofacies geometry (Bierkens and Weerts, 1994; Scheibe and Freyberg, 1995).

In the following sections, we introduce concepts of heterogeneity defined as spatial variability in physical, or chemical, or microbiological properties, within the context of scales of observation, to set the stage for upscaling or averaging methods, that is the topic of this chapter. We then turn to some basic demonstrations of how one can upscale various occurrences of heterogeneity. This is first described in the Eulerian setting for some elementary single-component reactions. Next, the Lagrangian perspective is described and demonstrated for a multicomponent reaction network with nonlinear kinetics. Finally, we describe a connection between the Eulerian and Lagrangian settings.

This review is not comprehensive either with respect to reactions encountered in the subsurface or with respect to most up-to-date contributions. Our target is rather to demonstrate some of the basic results in upscaling solute fate and transport in heterogeneous formations, for some simple cases of reactions. Thus, our goal is to provide an introductory review of the concepts and some approaches used in dealing with subsurface heterogeneity and its effects on solute fate and transport.

19.2 Physical, Chemical, and Microbiological Heterogeneity

It is clear that transport models for multiple species of reactive and degradable chemicals in a homogeneous environment must be very complicated. Couple this already complicated transport phenomena with heterogeneity, and the resultant problem increases by an order of magnitude in complexity. To date, only the simplest of such problems have been solved analytically, and even Monte Carlo methods have been applied to only moderately complex systems.

Aquifer properties are commonly found to vary on multiple spatial scales in a variety of settings, including alluvial (e.g., Weissman and Fogg, 1999; Schulmeister et al., 2003) and marine depositional (Tompson et al., 1996; DeFlaun et al., 2001) environments. Such heterogeneity arises in part from aquifer depositional structures that exhibit variations from the centimeter (e.g., cross-bedding) to the meter (lithofacies) scales. In the previous chapter, physical heterogeneity was discussed in the context of a random hydraulic conductivity field. There are two other commonly encountered forms of physical heterogeneity: a spatially variable porosity field and a spatially variable “local” dispersivity field. Both are commonly considered of secondary importance to flow and transport, though recent Monte Carlo simulations (Hassan et al., 1998) suggest that the correlation between conductivity and porosity may play a significant role in transport. As in Chapter 18, however, we will equate physical heterogeneity with ignorance about the values of particular parameters and how they vary over space. This will usually be approached by conceptualizing, for example, the hydraulic conductivity as a “random” conductivity field. In such a field the value of a parameter such as conductivity is viewed as a random deviate from a distribution that can be characterized by its moments and the like (mean, standard deviation, etc.). When the moments of the distribution themselves do not depend on space, the random field is termed “stationary” or, equivalently, “statistically homogeneous.” While stationary field representations are among the simplest (and most commonly relied upon), they are not to be confused with “homogeneous” fields where conductivity value is the same everywhere. Another critically important aspect of a “random” field is “ergodicity.” In loose terms, a field is ergodic if the probability distribution of values that can occur at a point in space is equivalent to the frequency distribution of values over space for the given subsurface in question. Ergodicity is critically important because if it holds for a given aquifer, then the statistics for

describing the random field can be estimated to some degree by compiling data from boreholes and other (e.g., geophysical) sources from different locations within the given subsurface in question. Unfortunately, proof of ergodicity would require not only exhaustive characterization of an entire subsurface domain, but also definition and identification of a set of subsurface domains that could arguably serve as a probability space with samples from each one at a consistent point in space — clearly an impossibility. Therefore ergodicity can not be proved.

Superimposed on the random conductivity field is a random geochemical field. Historically and spatially varying distributions of mineral deposition, formation, diagenesis, and weathering and organic matter occurrence and deposition give rise to this natural heterogeneity. Superimposed over both the physically and chemically heterogeneous fields is a dynamic random distribution of microbial populations. The major difference between physical/chemical heterogeneity and microbial heterogeneity is the dynamic nature of the latter. Microbes are transported and adsorbed as much as are chemicals and other colloidal particles, though the adsorptive mechanisms are completely different, with physiologically dependent residence times on surfaces leading in some cases to biofilm formation and eventual decay.

19.3 Scale of Observation and Constitutive Concepts

Engineered porous media are generally designed to be homogeneous when viewed over the scale on which they are to be used in a given technology. Natural porous formations, on the other hand, are almost always inhomogeneous on any relevant scale of observation. As mentioned in the previous section these natural formations are not only inhomogeneous physically, but chemically and biologically as well. However, the degree and type of these heterogeneities depends to a large extent on the scale of observation. It is for this reason that “effective” parameters such as so-called macro dispersivity have little meaning in natural systems.

A primary goal of modelers of transport in heterogeneous media is to develop functional forms for constitutive variables and to know the regions of space-time that are important to their definition. If the constitutive theory depends upon what is happening at a point (here a point has meaning only with respect to the window and scale of resolution associated with the measurement process) in space-time, or a very small neighborhood of the point, then the theory is said to be local. On the other hand, if information is needed to define the constitutive theory from regions of space-time distinct from a neighborhood of the space-time point where evaluation of the theory is to be made, then the model is said to be nonlocal. These ideas will be made more concrete in subsequent sections.

19.4 Upscaling

The concept of “upscale” is one of the most important in modern hydrological modeling circles. Upscaling simply means taking information on scales smaller than those of interest, homogenizing it, and studying the resultant “larger” scale system. The rationale behind the need to upscale is that we simply cannot make enough detailed measurements at the smaller scales to be of much use on the larger scales wherein relevant transport processes are taking place.

There are numerous methods that may be used to upscale physics in hydrology, including matched asymptotic expansions, Taylor–Aris–Brenner method of moments, renormalization group theory, transform methods, volume averaging, particle tracking, and so on. Many of these are illustrated in Cushman et al. (2002) and the interested reader is referred there. The next subsection reviews works regarding mineralogical or otherwise reactivity-associated properties and the effects of heterogeneity therein.

19.4.1 Methods for Chemical Heterogeneity

Here we review some concepts and results in the use of the Lagrangian approach when not only physical but also chemical and microbiological properties are heterogeneous. First, we term any reactive property

of an aquifer “aquifer reactivity.” This includes any immobile reactive component, for example, sorbed organic matter, organic mineral assemblages, (co-)precipitated minerals, mineral-oxide surface coatings, or biotic phases such as microbial colonies or biofilms.

Glassley et al. (2001) characterized a small sample of fractured rock from Yucca Mountain, Nevada, and compared both high-resolution (modeled aquifer porosity and reactivity) and effective-homogeneous models and found that the effect of mineralogical heterogeneity on reactive solute transport depends on how solute flux is distributed over contact times with minerals. At larger scales this dependency is often viewed in the context of how reactive transport depends on the way in which the patterns of geochemical and physical properties are overlain in space. Investigations have involved both deterministic patterns of geochemical and physical properties in experimental (e.g., Knapp et al., 1998) and simulated (e.g., Burr et al., 1994) settings.

In sandy aquifer materials metal (iron, aluminum, manganese) oxides are among the most important solid species that impart the positively charged sites on an otherwise negatively charged silicate surface, thus controlling the sorption of negatively charged solute particles. The occurrence of these microscale heterogeneities in reactivity and the impact on transport of colloidal and viral tracers in particular has been studied by Elimelech and coworkers (e.g., Ryan et al., 1999; Elimelech et al., 2000; Chen et al., 2001; Bhattacharjee et al., 2002). There, negatively charged colloidal depositions are demonstrated to be influenced by the positively charged surface (e.g., Chen et al., 2001). The presence of organic matter absorbed on the mineral surface decreases the electrostatic force between the positively charged surface and negatively charged colloids (Ryan et al., 1999). The overall impacts of such “patchwise” heterogeneity of positively charged metal oxyhydroxide grain coatings, on the transport of virus tracers is described for instance in Bhattacharjee et al. (2002). Larger scales of variation were examined, in that both layered and randomly (log-normally) distributed physical and geochemical heterogeneities were considered.

The most direct and simple approach to upscaling microscale (e.g., patchwise) heterogeneity of reactivity, in both Eulerian and Lagrangian settings, is by replacing the heterogeneous medium with an effective homogeneous medium (HEM) with the same total reactive surfaces available (but distributed now, homogeneously). This usually involves the assumption that the solute has equal access to reactive surfaces occurring within any representative elementary volume (REV) that is complete mixing. In one study without this assumption, Lichtner and Tartakovsky (2003) show that even with a single sorption rate but surface area variable within the REV due to a log-normal grain size distribution, the effective rate is not a constant but varies with time. When complete mixing is assumed and the sorption reaction is fast but involves multiple sorbents (e.g., mineralogies) that are spatially variable at the grain scale, equilibrium models written for homogeneous reactivity can give wrong simulations especially when calibrated in batch but applied in the presence of transport (e.g., Bosma et al., 1993; Wise, 1993; Chen and Wagenet, 1995; Szecsody et al., 1998 for a multicomponent reaction system). When kinetically controlled sorption reactions are involved, this situation can give rise to multi-rate or multi-site reaction kinetics at the continuum and larger scales.

Sposito (1993) points out how multiple reactive mineral phases involved in a surface reaction with a single reversibly sorbing solute give rise to the multi-rate reaction model (e.g., Haggerty and Gorelick, 1995) and under Gamma-distribution of reactivity, a power-law distribution of sorption rates and thus residence times (e.g., Metzler and Klafter, 2002). If some of the reactions in a multi-rate model are fast on the timescale of transport and the remaining reactions, then they can in general be combined to a single equilibrium reaction that operates in parallel with the remaining kinetically controlled reactions. Various combinations of serial and parallel reactions models can be constructed from particular conceptual models of grain surfaces, sometimes involving intra-granular porosity and diffusion limitations approximated as kinetically controlled reactions (e.g., Villermaux, 1981; Valocchi, 1990; Cunningham et al., 1997).

Approaches to dealing with heterogeneity in reactive properties on larger (e.g., meter) scales are often based on hypothesized correlations between macroscopic physical properties and reactive properties. Kabala and Sposito (1991) derive an effective convective-dispersive transport equation for the expected value of a solute undergoing transport and equilibrium sorption in random velocity and distribution coefficient fields, in the absence of microscopic dispersion. Their results are extended to higher moments

and higher order statistics in Kabala and Sposito (1994). Burr et al. (1994) present a numerical study of three-dimensional reactive transport with a geostatistical description of hydraulic conductivity and sorption equilibrium distribution coefficient designed to resemble Borden aquifer characteristics. Using spatial moment analyses on a small Monte Carlo ensemble, the authors showed that a negative correlation between the physical and reactive properties led to an increased spreading of solute plumes, as reflected in the second spatial moment within individual plumes. Similar results are reflected in Dagan (1989), Chrysikopoulos et al. (1990), and Bellin et al. (1993). Tompson et al. (1996) provide nine separate case simulations of multicomponent reactive transport of a uranyl–citric acid mixture in multi-dimensional physically and chemically heterogeneous media, with speciation and surface-complexation (equilibrium) reactions involving hydrous-ferric oxide (goethite) coatings on aquifer materials (sand). The authors highlight the importance of the third space dimension in mixing, the effect of multicomponent interactions, and showed that fate is controlled by the abundance of the oxide as a function of the specific sand surface area as well as by larger scale patterns of oxide occurrence. Burr et al. (1994) also showed that the joint variability induces a pseudokinetic behavior, in that the ensemble mean bulk retardation factor can increase with time and plume displacement distance, even though the sorption reaction is equilibrium. The same result is obtained by Espinoza and Valocchi (1997) in a perturbation analysis and Monte Carlo study of one-dimensional solute transport with kinetically controlled sorption, and by Mishra et al. (1999) for reversible kinetic sorption.

The usefulness of HEM models noted above for upscaling grain-scale reactivity heterogeneity has also been examined at larger scales. For instance, Chen et al. (2001) and Knapp et al. (1998) investigated transports of colloids and bacteria, respectively, in experimental columns with porous media that were physically homogenous but chemically heterogeneous. They compared results to reactivity-homogenized cases where aquifer reactivity patterns were erased and replaced with homogeneous material containing equivalent total reactivity. They concluded that the actual spatial distribution of iron-coated surfaces did not affect the column-scale attenuation and transport of the solutes as long as the overall fractions of reactive surface area in the column were the same. However, these results are restricted to one-dimensional transport with chemical heterogeneity occurring in the same dimension (e.g., varying along the flow path). If reactivity were arranged in a second space dimension, for example, in bands aligned parallel to the direction of the flow, the colloidal and bacterial breakthrough curves would indeed depend on pattern, because some of the solute flux would experience zero reactivity. This dependence is partly reflected in two of the nine experiments of Tompson et al. (1996) that involved banded goethite patterns, although the only pattern tested was banding diagonal to flow that spanned the no flow boundaries (e.g., parallel-to-flow banding was not examined). The authors state that the impact of correlation between reactive surface area and hydraulic conductivity seemed less significant than the overall abundance and distribution of the reactive area, such as the kind of banded goethite patterns observed in a coastal sand body. Finally, studies with unidirectional flow cannot reflect the coupled effects of physical and reactivity heterogeneity that occur via nonuniform flow in multiple dimensions (including dilution, e.g., Tompson et al., 1996). Thus, conclusions for unidirectional flow cases are not generally applicable in naturally heterogeneous porous media because in natural systems solute flux may experience a variety of “frequency of reaction sites” exposures. This is the same notion constructed by Glassley et al. (2001).

19.5 Linearity and Nonequilibrium Eulerian Stochastic Models

19.5.1 Deterministic Nonequilibrium Sorption

It is assumed the chemical is transported by convection and local dispersion under steady, saturated, incompressible groundwater flow in a nondeformable porous medium, of constant porosity. The chemical undergoes first-order linear nonequilibrium kinetics with its surroundings. We assume a natural scale exists, the “Darcy” scale (as in Chapter 18), and we wish to upscale without knowledge of a scale separation between the Darcy and the “Reservoir” scale. Darcy-scale transport is governed by Equation 19.1. It is assumed the mean-flow is in the x_1 -direction, that K_r and K_d are deterministic constants, and as a result

of randomness in $\ln K$, V_i , S , and C are random. Solving for the mean and fluctuation equations for Equation 19.1 and using a transform analysis gives (Cushman et al., 1996), assuming triplet correlation fluxes can be neglected,

$$\begin{aligned} \frac{\partial \bar{C}}{\partial t} + \bar{V}_1 \frac{\partial \bar{C}}{\partial x_1} - \frac{\partial}{\partial x_j} \int_0^t \int_{\mathbb{R}^3} D_{ij}(\mathbf{x} - \mathbf{y}, t - t') \frac{\partial \bar{C}(\mathbf{y}, t')}{\partial y_i} d\mathbf{y} dt' \\ = -K_r K_d \int_0^t \left[\delta(t - t') - e^{-K_r(t-t')} \right] \bar{C}(\mathbf{x}, t') dt' + K_r e^{-K_r t} S_0 \end{aligned} \quad (19.3)$$

where

$$D_{ij}(\mathbf{x} - \mathbf{y}, t - t') = d_{ij} \delta(\mathbf{x} - \mathbf{y}, t - t') + \bar{v}_i \bar{v}_j(\mathbf{x} - \mathbf{y}) G(\mathbf{x} - \mathbf{y}, t - t') \quad (19.4)$$

and

$$G^\circ(\mathbf{k}, \omega) = \left[\omega \left(1 + \frac{K_r K_d}{\omega + K_r} \right) + d_i k_i^2 + i k_1 \bar{V}_1 \right]^{-1} \quad (19.5)$$

where $\bar{v}_i \bar{v}_j$ is the assumed stationary fluctuating-velocity covariance, S_0 and C_0 are initial concentrations, the \circ represents Fourier–Laplace transform where t is dual to ω and \mathbf{x} is dual to \mathbf{k} .

The form of the mean Equation 19.3 is quite different from the original local scale Equation 19.1. This dispersive flux is nonlocal (third term of the left-hand side of Equation 19.3), as is the source (left-hand side of Equation 19.3) term. The upscaling process is the source of nonlocality and it results from a lack of information on the small-scale details of the transport process.

19.5.2 Random Nonequilibrium Sorption

Here the local-scale model is exactly as the previous case except it is assumed K_d is a random spatial field that may be correlated with $\ln K$. In this case, the resultant mean equation is more complicated and given by Hu et al. 1995:

$$\begin{aligned} \frac{\partial \bar{C}}{\partial t} + \bar{V}_1 \frac{\partial \bar{C}}{\partial x_1} - \frac{\partial}{\partial x_j} \int_0^t \int_{\mathbb{R}^3} D_{ij}(\mathbf{x} - \mathbf{y}, t - t') \frac{\partial}{\partial y_i} \bar{C}(\mathbf{y}, t') d\mathbf{y} dt' \\ - \frac{\partial}{\partial x_j} \int_0^t \int_{\mathbb{R}^3} G(\mathbf{x} - \mathbf{y}, t - t') \bar{k}_d \bar{v}_j(\mathbf{x} - \mathbf{y}) \bar{C}(\mathbf{y}, t') d\mathbf{y} dt' \\ = -K_r \left\{ \bar{K}_d \bar{C} - e^{-K_r t} S_0 - K_r \bar{K}_d \int_0^t e^{-K_r(t-t')} \bar{C}(\mathbf{x}, t') dt' - \int_0^t \left[\delta(t - t') - K_r e^{-K_r(t-t')} \right] \right. \\ \times \left[\int_0^{t'} \int_{\mathbb{R}^3} G(\mathbf{x} - \mathbf{y}, t' - t'') \bar{k}_d \bar{k}_d(\mathbf{x} - \mathbf{y}) \bar{C}(\mathbf{y}, t'') d\mathbf{y} dt'' \right. \\ \left. + \int_0^{t'} \int_{\mathbb{R}^3} B(\mathbf{x} - \mathbf{y}, t' - t'') \bar{k}_d \bar{v}_i(\mathbf{x} - \mathbf{y}) \frac{\partial}{\partial y_i} \bar{C}(\mathbf{y}, t'') d\mathbf{y} dt'' \right] dt' \left. \right\} \end{aligned} \quad (19.6)$$

with

$$D_{ij}(\mathbf{x} - \mathbf{y}, t - t') = d_{ij}\delta(\mathbf{x} - \mathbf{y}, t - t') + \overline{v_i v_j}(\mathbf{x} - \mathbf{y})B(\mathbf{x} - \mathbf{y}, t - t') \quad (19.7)$$

$$\overset{\circ}{B}^{-1}(\mathbf{k}, \omega) = \omega \left(1 + \frac{K_r \overline{K_d}}{\omega + K_r} \right) + d_i k_i^2 + i k_1 \overline{V} \quad (19.8)$$

and

$$\overset{\circ}{G}(\mathbf{k}, \omega) = \overset{\circ}{B}(\mathbf{k}, \omega) \frac{\omega K_r}{\omega + K_r} \quad (19.9)$$

The addition of randomness in K_d again increases the complexity of the mean equation. In this case, in addition to nonlocal sources/sinks and dispersion, there is also nonlocality in the convective flux (fourth term on the left-hand side of Equation 19.6).

19.5.3 Random Nonequilibrium Sorption and Equilibrium Microbial Decay

It is possible to generalize the previous results to the case where there is linear first order microbial degradation. In this case, the local scale Equation 19.1 is replaced by

$$\frac{\partial C}{\partial t} + \frac{\partial S}{\partial t} = \frac{\partial}{\partial x_j} \left(d_{ij} \frac{\partial C}{\partial x_i} \right) - \frac{\partial (V_i C)}{\partial x_i} - K_c C - K_s S \quad (19.10)$$

which is coupled with Equation 19.2. Here K_c and K_s are the solute degradation rates in the solution and sorbed phases. The mean equation can be obtained for this problem, but it is too complicated to present here (see Hu et al. [1997] for its presentation and development). In the next subsection we will, however, present some numerical solutions to it.

19.5.4 Numerical Results for Random K_d and No Degradation

To apply Equation 19.6, one needs to know three covariance functions $\overline{v_i v_j}$, $\overline{k_d k_d}$, and $\overline{k_d v_i}$ and where v_i and k_d are the fluctuating velocity in the i th-direction and the fluctuating partition coefficient, respectively. As it is difficult, if not impossible, to measure v_i , it is common to functionally relate v_i to f (the fluctuating log-conductivity) and then to measure the covariance of f and its cross-covariance with k_d . First-order expressions relating v_i to f and v_i to k_d are (Gelhar and Axness, 1983; Hu et al., 1995)

$$\widehat{\overline{v_i v_j}}(\mathbf{k}) = \left(\frac{Q_g}{n} \right)^2 \left[\delta_{i1} - \frac{k_1 k_i}{|k|^2} \right]^2 \widehat{\overline{f f}}(\mathbf{k}) \quad (19.11)$$

and

$$\widehat{\overline{v_i k_d}}(\mathbf{k}) = \frac{Q_g}{n} \left[\delta_{i1} - \frac{k_1 k_i}{|k|^2} \right] \widehat{\overline{f k_d}}(\mathbf{k}) \quad (19.12)$$

where Q_g is the product of the geometric mean conductivity with the mean gradient in the x_1 -direction, n is the porosity, and caret indicates Fourier space transform. Both $\overline{f f}$ and $\overline{f k_d}$ must be obtained from field experiments. For illustrative purposes, we assume various correlation structures for these covariances. For perfect positive correlation (Model A) or perfect negative correlation (Model B) to first order, we have

$$\widehat{\overline{k_d k_d}}(\mathbf{k}) = (K_d^G)^2 e^{\sigma_f^2} \widehat{\overline{f f}}(\mathbf{k}) \quad (19.13)$$

and

$$\widehat{\overline{k_d f}}(\mathbf{k}) = \pm K_d^G e^{\sigma_f^2} \widehat{f\bar{f}}(\mathbf{k}) \quad (19.14)$$

with plus for Model A and minus for Model B, where K_d^G is the geometric mean of $K_d(x)$. For uncorrelated K_d (Model C), to first order,

$$\widehat{\overline{k_d k_d}}(\mathbf{k}) = (K_d^G)^2 e^{\sigma_w^2} \widehat{\overline{WW}}(\mathbf{k}) \quad (19.15)$$

and

$$\widehat{\overline{f k_d}}(\mathbf{k}) = 0 \quad (19.16)$$

where W is a normally distributed random space function with zero mean, variance σ_W^2 , and covariance

$$\widehat{\overline{WW}}(z) = \sigma_W^2 e^{-zl/l_w} \quad (19.17)$$

where l_w is the integral scale of W and l is the correlation length of $\ln K$.

We take the initial concentration C_0 to be constant, C_m over a rectangular prism $2a_1 \times 2a_2$ and assume $S_0 \equiv 0$. The log-conductivity is chosen to be exponential

$$\overline{f}(z) = \sigma_f^2 \exp[-|z|^2/l^2] \quad (19.18)$$

Model D will correspond to deterministic and constant K_d .

In Figure 19.1 (from Hu et al., 1995, Figure 3), we plot the various spatial moments as a function of nondimensional time $t' = t\bar{V}/\bar{R}l$ where $\bar{R} = 1 + \overline{K_d}$. It is seen that the first moment is little affected by the models irrespective of K_r . Relative to the uncorrelated model, negative correlation between the chemical heterogeneity and physical heterogeneity increases the second longitudinal moment and positive correlation decreases it. The difference between the three models becomes larger as K_r increases. Deterministic K_d can produce significantly different results from random K_d . Figure 19.1c shows the difference in the second transverse moments between the four models is small, but distinct at large time.

19.5.5 Numerical Results with Random Chemical and Microbiological Conditions

The mean concentration for this case can be obtained given a much larger number of covariance functions, $\overline{v_i v_j}$, $\overline{v_i k_d}$, $\overline{v_i k_c}$, $\overline{v_i k_s}$, $\overline{k_d k_d}$, $\overline{k_d k_s}$, $\overline{k_d k_c}$, $\overline{k_c k_c}$, $\overline{k_s k_s}$, and $\overline{k_s k_c}$. We call the case without degradation model D. In this case, $K_s \equiv K_c \equiv 0$. We again must relate v_i , k_d , k_c , and k_s to f and in general, it follows to first order (with $a = d, c, s$).

$$\widehat{\overline{k_a k_a}}(\mathbf{k}) = \frac{Q_g}{n} \left[\delta_{il} - \frac{k_l k_i}{|k|^2} \right] \widehat{f k_a}(\mathbf{k}) \quad (19.19)$$

For illustrative purposes, we assume k_c and k_s are Gaussian and

$$\overline{k_c k_s} = \sigma_c \sigma_s \exp[-r^2/l_c^2] \quad (19.20)$$

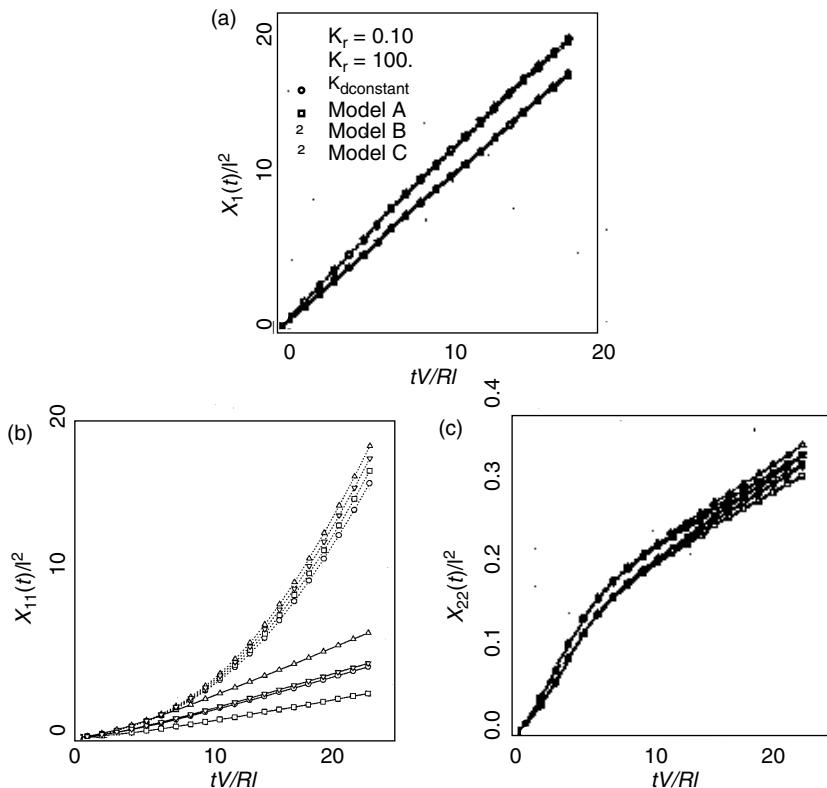


FIGURE 19.1 Spatial moments for various models with $K_d^G = 0.2$: (a) first moment, (b) second longitudinal moment, c) second transverse moment. (Adapted from Figure 3 of Hu et al., 1995, *Water Resour. Res.* 31(9):2219–2237.)

We assume positive (Model A), negative (Model B), and uncorrelated (Model C) covariance structures in the form

$$\begin{aligned}\overline{k_cf} &= \pm \sigma_c \sigma_f \exp[-r^2/l_{cf}^2] \\ \overline{k_sf} &= \pm \sigma_s \sigma_f \exp[-r^2/l_f^2]\end{aligned}$$

where + is for Model A, and - for Model B, and k_d is linearly negatively correlated with f . Use the same initial data as the previous case with $l_c = l_{cf} = l_f = 1.0$ m, $\sigma_f^2 = 0.2$, $V_1 = 1.0$ m/day, $d_1 = 0.05$ m²/day, $d_2 = 0.005$ m²/day, $c_m = 1$, $\bar{K}_d = 1$ /day, $\bar{K}_c = 0.01$ /day, $\bar{K}_s = 0.003$ /day, $\sigma_c^2 = 0.01$, $\sigma_s^2 = 0.003$.

Figure 19.2 plots mean concentration contours for the various models at 80 days. Models A, B, and C produce similar mean distributions. The shape of Model D's contours are similar to the others, but they are larger in area and have higher mean concentrations, owing to a lack of degradation.

19.6 Upscaling Microbial Dynamics

Even when microbes do not stick to the porous medium, they should be considered as reactive. This is because their self-motility allows them to move in a nonpassive way. They can move against the mean flow, transverse to it, or move rapidly in the direction of the mean flow and effectively can behave as retarded or enhanced particles. Initially this behavior was modeled as enhanced diffusion, however, in recent years researchers have begun to model microbial transport as nonlocal via Lévy motions (see Figure 1, Park et al., 2005a, 2005b). One of the biggest challenges faced when modeling microbial dynamics is how to

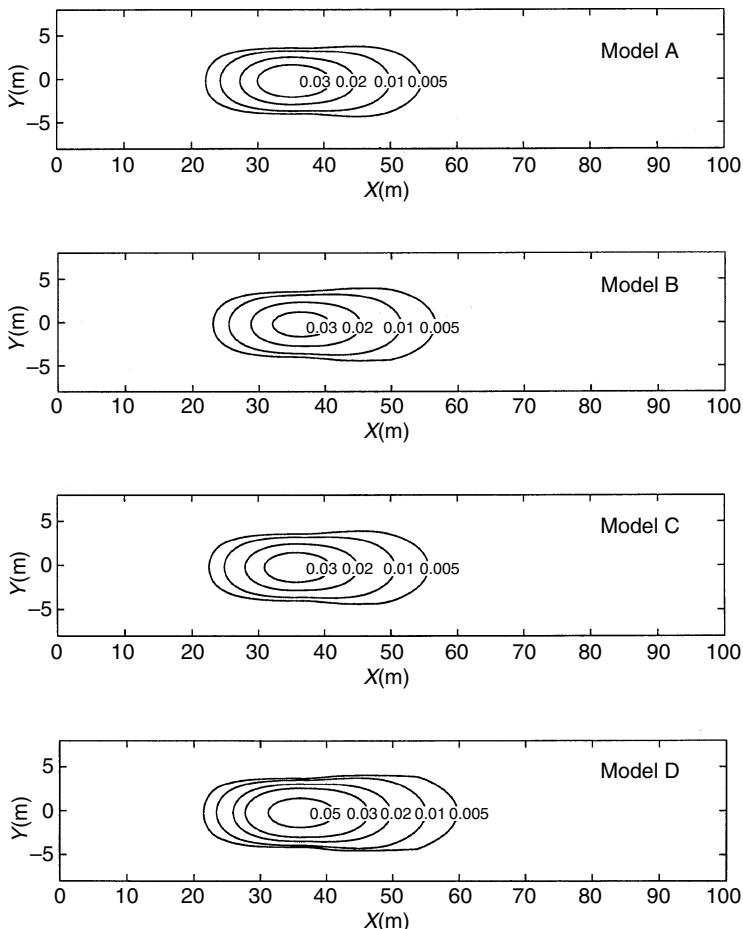


FIGURE 19.2 Concentration contours for various models at $t = 80$ days for $K_r = \bar{K}_d = 1/\text{day}$, $\bar{K}_c = 0.01/\text{day}$, and $\bar{K}_s = 0.003/\text{day}$. (Adapted from Figure 1 of Hu et al., 1997, *Adv. Water Resour.* 20(5–6):293–308.)

upscale from the pore to the Darcy to the field scale, especially when the field scale may be fractal. One method of attack, proposed in Park et al. (2005a, 2005b) is via the use of generalized central limit theorems (CLT). Essentially one writes the stochastic differential equation that the microbe obeys on the smallest scale and then applies the CLT to upscale. We present only a rudimentary summary of the results here.

On the pore scale, the stochastic ordinary differential equation (SODE) the motile particle is assumed to satisfy has the form

$$X^{(0)}(t) = X^{(0)}(0) + \int_0^t v^{(0)}(X^{(0)}(r))dr + \rho^{(0)}L^{(0)}(t) \quad (19.21)$$

where the drift velocity $v^{(0)}(X^{(0)}(r))$ is assumed stationary, ergodic, and Markov with mean \bar{v} and the fluctuating position $L^{(0)}(t)$ is α_1 -stable Lévy and accounts for the self-motility. A classical CLT (Bhattacharya, 1982; Bhattacharya and Gupta, 1990) coupled with the fact that Lévy motions dominate Brownian shows $X^{(0)}(nt) \rightarrow L^{(1)}(t)$, where $L^{(1)}(t)$ is a Darcy-scale Lévy motion which is a renormalized pore trajectory.

Thus the Darcy scale SODE is

$$X^{(1)}(t) = X^{(1)}(0) + \int_0^t v^{(1)}(X^{(1)}(r))dr + \rho^{(1)}L^{(1)}(t) \quad (19.22)$$

where $v^{(1)}(X^{(1)})$ is the Lagrangian velocity, $V^{(1)}$. In form, the previous two equations are identical, but mathematically $v^{(1)}(X^{(1)}(r))$ is rarely stationary and so the classical CLT cannot be applied to obtain the field scale trajectory $X^{(2)}$. However, generalized CLTSs can be proved (Park, 2005a) and applied if it is assumed the Lagrangian drift velocity at the Darcy scale can be modeled as an α_ν -stable Levy process. With this assumption and a few technical constraints, it can be shown that at the field scale the ADE for the transition density (concentration) in 1-D is

$$\frac{\partial f}{\partial t} = -v^{(2)} \frac{\partial f}{\partial x} + D(t) \frac{\partial^{\alpha_\nu} f}{\partial x^{\alpha_\nu}} \quad (19.23)$$

where

$$v^{(2)} = [v^{(1)}] + \mu_l^{(1)} \quad (19.24a)$$

$$[v^{(1)}] = E \left[\frac{1}{|A| |B|} \int_A \int_B v^{(1)}(t, x_0) dt dx_0 \right] \quad (19.24b)$$

$$A = A(x_0(\omega)) = \{x_0(\omega) : x_0(\omega) \in [0, 1]\} \quad (19.24c)$$

$$B = B(t, \omega) = \left\{ t \geq 0; \int_0^t X^{(1)}(\tau, \omega) d\tau \in [0, 1] \right\} \quad (19.24d)$$

where ω is an elementary event. The α_ν derivative $\partial^{\alpha_\nu} f / \partial x^{\alpha_\nu}$ is the well-known fractional derivative, and $D(t) = -(t\sigma_\nu)^{\alpha_\nu} / \cos(\pi\alpha_\nu/2)$ with σ_ν a parameter in the definition of the α_ν stable Lévy motion; $S_{\alpha_\nu}(t^{1/\alpha_\nu} \sigma_\nu, 1, v^{(1)})$ is the distribution for the Lévy motion.

19.7 Lagrangian Stochastic Models for Nonlinear Reactions

Mass transfer equations representing natural biological and chemical transformations in saturated subsurface systems typically involve strongly nonlinear transformations among multiple species. Eulerian methods that result in ensemble-averaged expected plume behavior have been successful in upscaling the effects of idealized heterogeneities on the fate and transport of linearly sorbing solutes at both equilibrium (e.g., Berglund and Cvetkovic, 1996) and nonequilibrium rates (e.g., above, cf. also Dagan and Cvetkovic, 1993; Quinodoz and Valocchi, 1993; Andricevic, 1995), and solutes undergoing linear decay (e.g., Hu et al., 1995; Miralles-Wilhelm and Gelhar, 1996). These methods are however, generally inappropriate for the application to reactive transport where nonlinearities exist in nonequilibrium transformations or in multicomponent (coupled) dependencies. Sudicky et al. (1990) point out that the ergodic requirement (that, simply, the solute domain is large relative to the dominant scale of heterogeneity) is invalid for a nonconservative (e.g., sorbing or degraded) solute whose domain shrinks with time. Also, the use of a macrodispersion parameter effectively represents as mixing what may actually be plume deformation due to nonuniform convections, as described in the previous chapter. In the case of dually-limiting reactants (interior and exterior to the plume, such that reactions only occur at the mixing fringe), this artificial mixing will significantly overestimate the reaction sink. The scaling procedure, as an equation averaging operation, is valid when the averaging operator commutes with the (linear) kinetic reaction operator, so

that the average of quantities *within* the reaction term may be obtained. When nonlinearities exist in the kinetics of transformation, the averaging and reaction operations do not generally commute, and so the scaling approach fails to generate useful equations for average plume behavior.

An approach to scaling complex reactive transport in heterogeneous materials may be obtained when one first applies the reaction operation to solute parcels as they follow nonuniform convective paths, and then averages the resulting reacted solute convections over the appropriate convective paths contributing to an observation. In this approach, known as the solute flux-based (Cvetkovic and Dagan, 1996; Cvetkovic et al., 1998) or stochastic-convective reaction (SCR) method (Simmons et al., 1995), the reactions are cast in terms of the solute residence-time during convection. The method is based in the stochastic-convective representation of dispersive transport of a passive tracer (Simmons, 1982). The travel-time of the solute is defined as the time of residence of solute moving through the aquifer, from the point of spill or contamination entry to the system, to the observation location or regulatory boundary. Solute particles that enter the system and arrive at the observation location follow particular paths, or streamtubes, and the set of streamtubes delivering solutes to an observation sampler (point), line (well), or plane (regulatory boundary) makes up an ensemble of travel paths, and so the travel-time describing the ensemble is a distributed quantity, characterized by a distribution function. When the reactive transport model along any given streamtube can be expressed in terms of travel-time, the solution to the model is in terms of travel-time, and the solute arrival at the observation is recovered by averaging the solution of the model against the travel-time distribution function.

The advantages of this approach are (1) the reliance upon travel-time, a global measure of the effects of locally-varying Eulerian velocities, as opposed to estimates of the full nonuniform velocity field, and (2) the solution for each streamtube equation system can be obtained before averaging, thus eliminating the operator commutation problem with scaling nonlinearly reactive transport. This latter feature, it should be noted, results from the assumption that lateral mixing between streamtubes is negligible, that is, the flow in which the transport occurs is “segregated.” Recent work by Viswanathan and Valocchi (2004) and by Robinson and Viswanathan (2003) provides some innovative effective methods to deal with the occurrence of lateral mixing. This scaling approach is usually applied to provide predictions of observed solute arrivals at observation samples, or arrivals at an imaginary control plane. It is a direct although cumbersome matter to use the SCR approach to reconstruct estimates of plume evolution, which may be done if the travel-time distribution function is given or estimated at many points in the aquifer. Similar streamtube principles have been used to gain computational speedup in two- and three-dimensional reactive transport simulations in a deterministic setting (Thiele et al., 1995a, 1995b).

The method proceeds in three basic steps. First, the existence of an ensemble of streamtubes that is suitable for averaging must be ensured by checking the validity of the SCR assumptions for the particular reaction system, initial, and boundary conditions. Second, the reaction system must be solved for the streamtube ensemble (or canonical streamtube, defined below, when it exists). Finally, the travel-time distribution function must be determined for the particular flow regime and initial and boundary conditions. The following two subsections borrow heavily from Simmons et al. (1995).

19.7.1 Streamtube Formulation

The concept of the streamtube as a path taken by solute parcels from a source to an observation is made under the requirement that flow in it (a subset of the total flow domain) is constant for both space and time. The assumptions under which such an ensemble of streamtubes may exist are as follows: The members of the ensemble are physically independent, with no mass transfer among streamtubes. This allows treatment of the solute paths as integral, averagable quantities. This restricts diffusive/dispersive mass transports lateral to streamtubes; for example, this representation is incapable of incorporating local transverse dispersion. For simplicity we ignore also longitudinal mixing processes, although this is not a requisite assumption as a longitudinal mixing may be incorporated within the streamtube (Ginn, 2001). Second, the streamtube orientation in Eulerian space is arbitrary save that it must of course connect the source to the measurement location, and that the projection of its axis, s , onto the x -axis connecting the

source and observation be one-to-one (such that the $x(s)$ is monotone increasing). This assumption allows us to associate with each streamtube a 1D velocity function $v(x)\epsilon\{v(x)\}$ that is positive, and so allows the definition of an associated travel-time function. Further, each streamtube obeys a known reaction function r . A streamtube in three-space (which may have variable areal cross section) is defined such that its average cross-sectional velocity $v_s(x')$ times the corresponding cross-sectional water content $n_s(x')$ is a constant, q . To represent a streamtube in three-dimensions with an equivalent 1D streamtube with the same flow q we must honor the velocity to preserve the transport characteristics. Velocity is defined by $v_s(x') = q/n_s(x')$ in the streamtube and by $v(x') = q/n(x')$ in the equivalent 1D system. Mass balance implies $v_s(x')n_s(x')A_s(x') = v(x')n(x')A_0$ where A_0 is a (constant) reference area, and so we may define an effective porosity $n(x') = n_s(x')A_s(x')/A_0$ to give $v(x') = v_s(x')$ along the equivalent 1D system. Under this construction the convective-reactive transport along a given streamtube is represented in the 1D form

$$\frac{\partial(nC)}{\partial t} + \frac{\partial(nvC)}{\partial x} = nr \quad (19.25)$$

By construction $n(x)v(x) = q$ is a constant of x , and n is assumed constant of t , so Equation 19.25 simplifies to

$$\frac{\partial C}{\partial t} + v(x)\frac{\partial C}{\partial x} = r \quad (19.26)$$

A simple change of variables in Equation 19.26 allows expression of velocity in terms of solute travel-time. The travel-time corresponding to the streamtube velocity $v(x)$ is the time elapsed as a tracking particle is transported from x_0 to x ,

$$T(x; x_0) = \int_{x_0}^x dx'/v(x') = \int_0^x dx'/v(x') - \int_0^{x_0} dx'/v(x') = T(x) - T(x_0) \quad (19.27)$$

where $T(x)$ means $T(x', 0)$. By positivity of velocity, $T(x)$ is a monotone increasing function in x and the inverse function, $\Xi(T'(x)) = T^1(\tau) = x$, exists and is monotone increasing in variable τ . Note that mass loss over the ensemble corresponds to the case where ensemble maximum travel-time is infinity. In this case, a finite T^* would be used that represents the largest travel-time of a tracer particle captured at a measurement location, and the mass lost would be accounted in the specification of the streamtube frequency distribution. Now invoking the chain rule with $dT(x) = dx/v(x)$ and incorporating the inverse function $\Xi(T)$ we may write for Equation 19.26:

$$\frac{\partial C}{\partial t} + \frac{\partial C}{\partial T} = r(\Xi(T), t, C) \quad (19.28)$$

This last form expresses the result that C can be obtained explicitly as a function of global travel-time. Equation 19.7 is the basic streamtube model. Note that if r is streamtube invariant (requiring that r' contain effective streamtube porosity as factor) the solution to Equation 19.26 depends only on v . Any set of streamtubes for which the velocity statistics are identical to those of $\{v\}$ (determined from the observation of conservative tracer breakthrough or from conductivity statistics), and for which r accurately represents the reactions, suffices. A working hypothesis of this approach is that the streamtube model more closely represents the actual conditions for the definition and calibration of the kinetic model r than does the conventional averaging approach. Closed-form solutions to Equation 19.26 may be found under general conditions on r , including nonlinear and dynamic forms (e.g., r depends on a parameter such as microbial mass that evolves according to its own differential equation). We will return to a few examples after introducing the travel-time distribution function.

Ensemble averaging of the solution to Equation 19.28, described next, is greatly simplified when the solution to Equation 19.28 may be written in terms of dimensionless travel-time, and rescaled to fit any particular travel-time. When the solution to the streamtube equation is rescalable over travel-time, the solution is termed “canonical.”

19.7.2 Stochastic–Convective Averaging

Here we examine averaging of the foregoing solutions according to the travel-time pdf, a measure of streamtube frequency. Once a canonical solution of Equation 19.7 is obtained as $C(T, t)$, the streamtube ensemble average solution, $\langle C \rangle(x, t)$, is found by averaging $C(T, t)$ over the travel-time pdf $p(\tau; x)$ for the streamtube travel-times to location x . This pdf is defined over the streamtube ensemble in the frequency sense as (Simmons, 1982; Shapiro and Cvetkovic, 1988; Dagan and Nguyen, 1989)

$$p(\tau; x) = \langle \delta(\tau - T(x)) \rangle \quad (19.29)$$

where δ is the Dirac distribution function and the angle brackets indicate average over the streamtube ensemble. The expectation can be expressed as

$$\langle C \rangle(x, t) = \int_0^{\infty} C(\tau, t) p(\tau; x) d\tau \quad (19.30)$$

In Equation 19.30, $p(\tau; x)$ is the pdf for random travel-time τ to a plane normal to the major flow axis at x , measured along that axis (Simmons, 1982; Dagan and Nguyen, 1989). The averaging in Equation 19.30 is over the realizations of (random) travel-time at a certain point in x space, and not over the deterministic measure of travel-time $T(x)$ itself. That is, the average in the right-hand side of Equation 19.30 is exactly the streamtube ensemble average, as is demonstrated on substitution of Equation 19.29:

$$\int_0^{\infty} C(\tau; t) p(\tau; x) d\tau \equiv \int_0^{\infty} C(\tau; t) \langle \delta(\tau - T(x)) \rangle d\tau = \left\langle \int_0^{\infty} C(\tau; t) \delta(\tau - T(x)) d\tau \right\rangle = \langle C(T(x), t) \rangle \quad (19.31)$$

The numerical calculation Equation 19.31 via discrete integration requires first the enumeration of the surfaces $C_i(T, t)$ for each realization of travel-time function $T_i(x)$. When a canonical solution is available this is accomplished by simply rescaling the travel-time axis of the canonical solution $C(T, t)$ to the indicated $T_i(x)$. This approach is used by Wise and Carbeneau (1994), whose averaging is associated with the measured cumulative passive breakthrough (which they term the “fractional breakthrough,” that is the integral part of our travel-time pdf).

Finally, one may also introduce heterogeneity into the reaction term through either the space-time dependency or the solute component dependency in the right-hand side of Equation 19.30 (for instance, when $r(x, t, C) = g(x, t)h(C)$, and g indicates concentration of reactive iron oxides or attached biomass). Further, correlations between the flow properties and the reaction properties may be cast as a conditioning of the travel-time pdf upon the value of the reaction properties. When such heterogeneity is represented through pdfs of the random properties and when these pdfs are known, the expectation may be computed in many cases by simple sequential averaging of $C(T, t)$ according to each random property (with pdfs appropriately conditioned to reflect correlations). Suppose, for instance, randomness in the case of a factorable hyperbolic kinetic decay (e.g., biodegradation following Michaelis–Menten kinetics), is such

that we write Equation 19.30 as

$$r(x, t, C) = \{g\}(x, t)h(C) \quad (19.32)$$

where g represents time evolution of a random space function selected from the ensemble $\{g\}$. That is, Equation 19.32 is the reactive term written for one realization (streamtube) of the random pathways. In this case (Simmons et al., 1995) the streamtube solution is cast in terms of not only travel-time (cumulative reciprocal velocity) but also cumulative “reaction property” G , defined as the integral of g along the streamtube coordinate. Then the travel-time averaging in Equation 19.30 is followed by averaging over the ensemble distribution of G , appropriately conditioned;

$$\langle C(x, t) \rangle_G = \int_0^\infty C(\tau, t) p(\tau; x|G) d\tau \quad (19.33a)$$

$$\langle C(x, t) \rangle = \int_{-\infty}^\infty \langle C(\tau, t) \rangle_G p(G; x) dG \quad (19.33b)$$

19.7.3 Travel-Time Distribution Function

The travel-time $T(x)$ determines the mapping of the random velocity along a streamtube onto a random velocity field for an equivalent one-dimensional system with distance coordinate x . Thus the travel-time pdf is experimentally observed as the conservative tracer breakthrough curve at the control plane at x in response to a unit Dirac- δ input function. Alternatively, the pdf may be deconvolved from the observed response to other known input functions (e.g., Ginn et al., 1995). In either case, it should be noted that pore-scale mixing processes (such as diffusion and pore-scale dispersion) serve to smooth the observed breakthrough curve in an irreversible way. One may also hypothesize forms for $p(\tau; x)$, a priori, such as lognormal or inverse-Gaussian (Simmons, 1982).

Frequently, tracer test results are not available for characterizing transport between a source and an observation location. It is possible to estimate the expected travel-time distribution function that corresponds to the average arrival time distribution over an ensemble of flow fields. To derive the expected travel-time distribution for particles arriving at an observation point at distance x_1 from source centroid, $g(t; x_1)$, at first order, we follow Dagan (1989, chap. 5.8; see also Appendix A in Cvetkovic and Dagan 1996). For simplicity it is assumed that all particles start within a finite plane source that is oriented normal to the average flow direction. A particle starts at coordinates $(0, a_2, a_3)$, at initial velocity parallel to x -axis $v_0 = V_1(0, a_2, a_3)$. The probability that a solute particle has crossed the regulatory control plane (assumed normal to the flow direction \bar{x} , located at x_1) at time $t = \tau$ is the same as the probability that the particle’s displacement along the flow axis is larger than the distance to the control plane, that is, $X_1 > x_1$. Hence, with $\phi(X_1, v_0; t)$ as the joint pdf of X_1, v_0 , at time t , one has for the cumulative arrival distribution

$$G(\tau; v_0, x_1) = \int_{x_1}^\infty \phi(X_1; v_0, \tau) dX_1 \quad (19.34)$$

It is important to note that the concentration probability function ϕ is a statistical construction that is not equivalent to any particular physical realization unless the transport is ergodic, for example, that the lateral domain of the source is large compared to the heterogeneity scale. Various models for ϕ have been derived based on different constructs of the statistical description of the controlling (conductivity) field and boundary conditions (Dagan, 1991; Dagan et al., 1992). The expected travel-time pdf is obtained from the cdf G by averaging Equation 19.34 over possible initial velocities v_0 (where v_0 is assumed Gaussian

with mean U) and differentiating with respect to τ .

$$g(\tau; x_1) = \frac{1}{U} \int v_0 g(\tau, v_0; x_1) dv_0 = \frac{1}{U} \frac{\partial}{\partial \tau} \int_0^\infty \int_{x_1}^\infty v_0 \phi(X_1, v_0; \tau) dX_1 dv_0 \quad (19.35)$$

For instance, for asymptotic conditions where the impact of v_0 and τ correlations vanish, g is given by

$$g(t; x_1) = -\frac{1}{\sqrt{2\pi}} \exp\left(\frac{-(x_1 - Ut)^2}{2X_{11}}\right) \frac{\partial}{\partial \tau} \left(\frac{x_1 - Ut}{\sqrt{X_{11}}} \right) \quad (19.36)$$

where X_{11} is the longitudinal particle covariance, a measure of the ensemble-averaged longitudinal second spatial moment, and U is the source-averaged initial velocity.

19.7.4 Some Reaction System Examples

Here we summarize some recent examples of reaction systems that have been constructed for particular biogeochemical systems and used in streamtube-ensemble reactive transport models. The examples include a generalized linear sorption model, and a model of aerobic biodegradation with microbial transport.

Lassey (1988) provided a generalized formulation of a model combining equilibrium, kinetic-reversible, and kinetic-irreversible sorption processes and its solution for particular boundary conditions. The solution to the streamtube Equation 19.30 with r given by the kinetic reversible sorption model of Equation 19.1a and Equation 19.1b (e.g., $r = -(\rho_b/n)\partial S/\partial t$), in the case of an instantaneous injection at the upstream boundary, is adopted from Lassey (1988; cf. Also Cvetkovic and Dagan, 1996; Cushman et al., 1996)

$$C(\tau, t) = e^{-K_r K_d t} \delta(t - \tau) + K_r^2 K_d \tau e^{[-K_r K_d t - K_r(t-\tau)]} \bar{I}_1[K_r^2 K_d \tau(t - \tau) H(t - \tau)] \quad (19.37)$$

where $\bar{I}_1(z) = z^{-1/2} I_1(2z^{1/2})$, and I_1 is the modified Bessel function of the first kind of order one. This solution is appropriate for averaging as indicated in Equation 19.30. Note that as pointed out in Simmons et al. (1995) and Theile (1995a), the so-called initial value problems, for example, those involving the initial distribution of a reactant (such as attached biomass) in space, impose certain restrictive assumptions on the use of the streamtube approach. In the most restrictive case, uniformity of the distributed property over the streamtube ensemble is required. (This is only slightly less restrictive than the requirement of spatial biochemical homogeneity overall.) In the event of spatially nonuniform initial distribution of biomass, full averaging of the streamtube-component evolution equations is feasible, but requires the joint probability of biomass occurrence and streamtube travel-time (as in Equation 19.33).

Biogeochemical and biodegradation systems involving attached phase biomass or biofilms are typically formulated as initial value problems, and thus may be complicated by heterogeneity in not only initial microbial distributions, but also by the dynamic microbial heterogeneity associated with microbial growth. This is the case in the next example, involving nonlinear degradation of an organic solute by a microorganism that is initially uniformly attached to the solid phase but partitions to the aqueous phase according to a nonequilibrium and nonlinear rate equation. This example describes meter-scale laboratory experiments in aerobic biodegradation in physically heterogeneous media, and is greatly condensed from Murphy et al. (1997) and Ginn et al. (2001). Simplified evolution equations for the organic substrate benzoate, oxygen, mobile microbes, immobile microbes, and passive tracer bromide, are shown here in one-dimensional (streamtube projection) format. Because we have retained one-dimensional longitudinal dispersivity d in the transport operation in Equation 19.38, to preserve the mixing between oxygenated and oxygen-depleted waters that is important to degradation, the equations are cast in terms of their streamtube velocity projected onto the x -axis, as opposed to travel-time. This aspect, and the possibility

of nonuniformities in the initial biomass concentrations, violates the requirements for a travel-time (i.e., constant-velocity) representation of streamtube velocities. Nevertheless for simplicity in calculating the SCR predictions, a constant velocity ensemble, corresponding to an equivalent travel-time ensemble, is used as an approximation of the transport.

The reactions involve degradation of benzoate (C_b , Equation 19.38a) as the electron-donor, with oxygen consumed as the terminal electron-acceptor (C_o , Equation 19.38b). Degradation induces an increase in biomass, the rates of which are the same for aqueous (C_{mm} , Equation 19.38c) and attached (C_{im} , Equation 19.38d) biomass. The mass transformations associated with the degradation are governed by the bracketed dual-Monod terms in (Equations 19.38a, b, c) that effectively shut down the transformation rate under limiting conditions, for example, when C_0 or C_b values drop to the order of K_0 or K_b , respectively. Biomass partitions kinetically and reversibly, but with the added complexity that the equilibrium partition coefficient is dependent on the local ionic strength, changes in which are due to the presence of an inert bromide tracer (C_T , Equation 19.38e) that effectively compresses the electrostatic repulsive layer of the quartzitic solid phase. For simplicity, attached biomass is treated as a stationary volumeless dissolved species. These transformations incur a strong nonlinear coupling among the four mobile components (organic solute, oxygen, aqueous biomass, and bromide) and the one immobile component (attached biomass). More complex forms, involving microbial metabolic lag and dynamic endogenous respiration, are described in Murphy et al. (1997). The transport takes place at average pore-water velocity of 50 cm/d through a 1 m flow cell that is 0.1 m × 0.2 m in cross-section, in inoculated porous media consisting of high hydraulic conductivity sand containing a set of randomly located low conductivity inclusions. Details on the heterogeneity pattern and pre-experimental modeling are found in Ginn et al. (1995).

$$\frac{\partial C_b}{\partial t} + \frac{\partial v_x C_b}{\partial x} - d \frac{\partial^2 C_b}{\partial x^2} = -\frac{\mu}{Y} (C_{mm} + C_{im}) \cdot \left\{ \frac{C_0}{(K_0 + C_0)} \frac{C_b}{(K_b + C_b)} \right\} \quad (19.38a)$$

$$\frac{\partial C_o}{\partial t} + \frac{\partial v_x C_o}{\partial x} - d \frac{\partial^2 C_o}{\partial x^2} = -F \frac{\mu}{Y} (C_{mm} + C_{im}) \cdot \left\{ \frac{C_0}{(K_0 + C_0)} \frac{C_b}{(K_b + C_b)} \right\} \quad (19.38b)$$

$$\frac{\partial C_{mm}}{\partial t} + \frac{\partial v_x C_{mm}}{\partial x} - d \frac{\partial^2 C_{mm}}{\partial x^2} = -\mu C_{mm} \left\{ \frac{C_0}{(K_0 + C_0)} \frac{C_c}{(K_c + C_c)} \right\} - K_r K_d (C_T) C_{mm} + K_r C_{im} \quad (19.38c)$$

$$\frac{\partial C_{im}}{\partial t} = -\mu C_{im} \cdot \left\{ \frac{C_0}{(K_0 + C_0)} \frac{C_c}{(K_c + C_c)} \right\} + K_r K_d (C_T) C_{mm} - K_r C_{im} \quad (19.38d)$$

$$\frac{\partial C_T}{\partial t} + \frac{\partial v_x C_T}{\partial x} - d \frac{\partial^2 C_T}{\partial x^2} = 0 \quad (19.38e)$$

The initial conditions in the flow cell are fully oxygenated water ($C_o = 8$ mg/l) with equilibrium distributed biomass at 5×10^6 cells/ml in the aqueous phase. The distribution coefficient is roughly 2/3 in the absence of bromide and doubles at 300 mg/l bromide. The boundary conditions are continuous injection of oxygenated (8 mg/l) water with a 1d pulse of benzoate (300 mg/l) and bromide (392 mg/l) in solution injected to evaluate the resulting biodegradation. Data collected include breakthrough curves at the outlet of the flow cell for benzoate, oxygen, and aqueous microorganisms. The first step in obtaining the SCR solution is the determination of a travel-time distribution function, which was taken from the breakthrough curve of the bromide tracer, shown by the "+" marks in the top panel of Figure 19.3. Using the principle of superposition in reverse allows deconvolution of the breakthrough curve to determine the breakthrough curve expected to result from an instantaneous injection (Dirac pulse) at the boundary, that is, the travel-time distribution function. This is shown in the bottom panel of Figure 19.3, and the deconvolved simulation of the bromide breakthrough is shown as the solid line in the top panel of Figure 19.3. (The deconvolution is not a trivial exercise because it involves solving a Volterra integral equation and

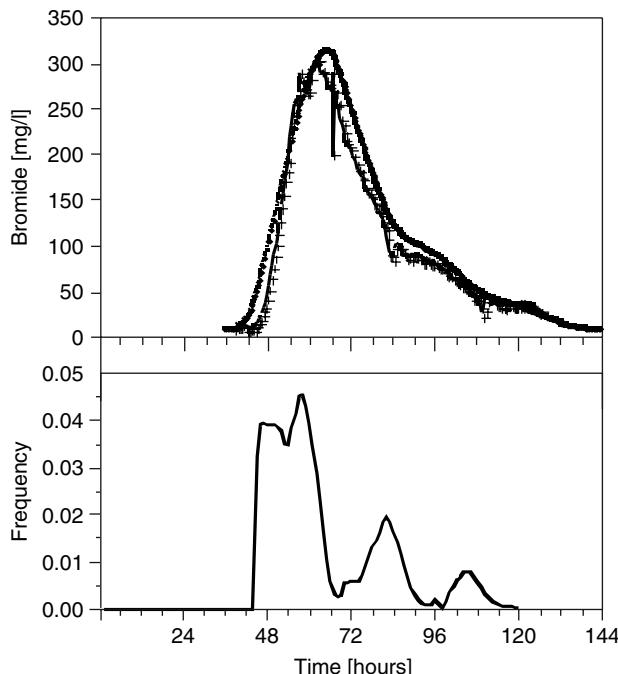


FIGURE 19.3 *Top panel:* Observed breakthrough of bromide (+) and reconstituted breakthrough curve (solid line) using convolution of the system travel-time distribution function with the 1-day square impulse function. *Bottom panel:* Travel-time distribution function.

is usually unstable numerically, although a unique solution exists.) In the second step, the SCR solution ensemble was generated by solving Equation 19.38 for 100 different but constant velocities corresponding to the travel-time domain shown in Figure 19.3. That is, effective velocities were selected according to $v = x/T$, where T is discretized at 40, 120 h. These solutions were obtained using RAFT, a deterministic flow and reactive transport simulator capable of general reaction systems (Chilakapati, 1995).

The SCR predictions are obtained in the third step by averaging the ensemble of breakthrough curves per solute against the travel-time distribution function shown in Figure 19.3. This average was carried out by solving a discrete form of Equation 19.30 using an ensemble of 100. The resulting predictions are shown with observed data for benzoate, oxygen, and aqueous microbes in Figure 19.4.

Numerous reaction systems have been addressed with the streamtube approach, in part because of the facility in developing solutions in one transport dimension. In addition to the foregoing, generalized solution surfaces $C(\tau, t)$ have been derived for multicomponent nonlinear equilibrium sorption (Charbeneau, 1988; Dagan and Cvetkovic, 1996; Ginn, 2001), immiscible two-phase displacement (the Buckley–Leverett problem, Thiele et al., 1995a), and simplified nonlinear decay, precipitation, or dissolution rate equations (Simmons et al., 1995; Dagan and Cvetkovic, 1996; Cirpka and Kitanidis, 2000), to name a few.

19.7.5 Toward Handling Heterogeneity in Reactive Properties

As was pointed out in Section 19.4.1, a number of works (Tompson, 1993; Szecsody et al., 1998; Glassley et al., 2001; Bhattacharjee et al., 2002) state explicitly that upscaling reactive transport in natural porous media requires not necessarily the aquifer reactivity volume fraction or otherwise Eulerian information, but the distribution of flux over cumulative reactivity.

Simmons et al. (1995) describe such an approach involving the distribution of solute flux over both travel-time τ and cumulative reactivity (defined as the integral along a streamtube path of the spatially

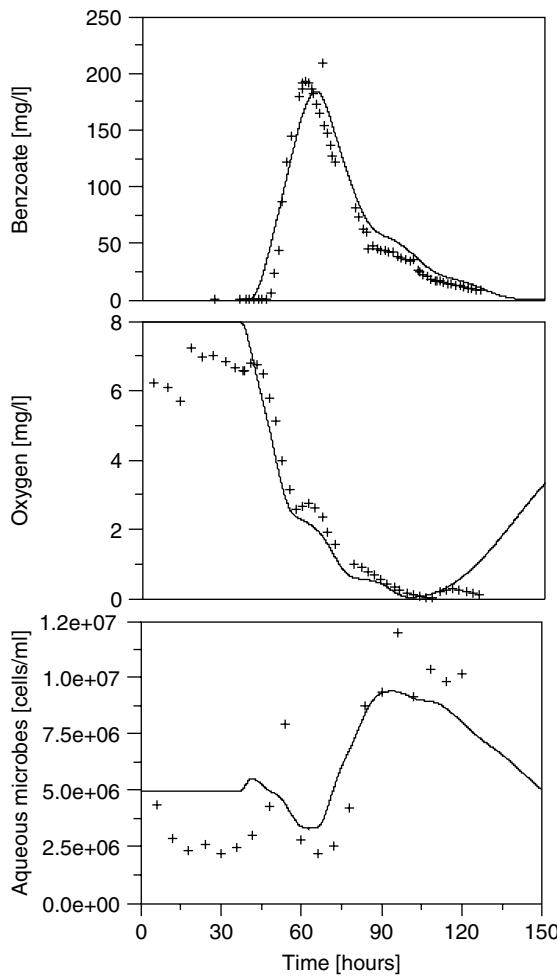


FIGURE 19.4 Top panel: Observed (+) and simulated (solid line) breakthrough of benzoate. Middle panel: Observed (+) and simulated (solid line) breakthrough of oxygen. Bottom panel: Observed (+) and simulated (solid line) breakthrough of aqueous microorganisms.

variable reactivity), G . The field scale solute flux is reconstructed as an integral over the ensemble of streamtubes with joint frequency distribution $p(\tau, G)$ of flux (Simmons et al., 1995, Equation 36). The method is outlined for general linear or nonlinear kinetically controlled irreversible reactions and is shown admissible for any kinetic where the reaction rate factors into (explicitly) spatially dependent term and concentration-dependent terms. Cvetkovic et al. (1998) provide a stochastic interpretation of the same basic approach and define a “reaction flow path” μ as the cumulative reactivity G above, and describe $p(\tau, \mu)$ in terms of the stochastic Lagrangian transport approach of Dagan (1989). They show first-order analytical and Monte Carlo methods to obtain statistics of $p(\tau, \mu)$. The conceptual model also differs in that transport is via advective streamlines whereas Simmons et al. (1995) conceptualize streamtubes. Cvetkovic et al. (1998) provide detailed temporal and spatial moment analyses for several equilibrium and kinetically controlled reactions in statistically homogeneous, stationary, and isotropic media. This approach is subsequently focused to the case of reactive transport in fractures in Cvetkovic et al. (1999) and used to simulate sorbing tracer transport at Yucca Mountain in Painter et al. (2001) and in a network model based on the Äspö (Sweden) site in Cvetkovic et al. (2004). The authors point out that their results stress the need for further studies on the upscaling of travel-time and cumulative reactivity distributions.

19.8 Connection between Approaches

The Eulerian methods described in Section 19.6 and the Lagrangian approach described above can be examined in light of one another through the use of a formal coordinate transformation that shows the particle (Lagrangian) paths in terms of the spatial (Eulerian) velocities (Dagan et al., 1992; Dagan and Cvetkovic, 1993; Cvetkovic and Dagan, 1994, 1996; Cushman et al., 1996). In this section, we present this coordinate transformation for reference. We begin conceptually again with Equation 19.1, taking $d_{ij} = 0$ and steady divergence-free velocity field $\mathbf{v}(\mathbf{x})$, the mean of which is in the x_1 direction. Further, the magnitude of $v_1(\mathbf{x})$ is restricted to be positive. Now to trace the path of a solute parcel or particle in such a system, define a particle coordinate tracking function $\mathbf{X}(t; t^0, \mathbf{x}^0)$ as a solution to the Lagrangian equation $d\mathbf{X}/dt = \mathbf{V}(\mathbf{X})$, with the initial condition $\mathbf{X}(t^0; t^0, \mathbf{x}^0) = \mathbf{x}^0$. Under steady velocity, we have also $\mathbf{X}(t; t^0, \mathbf{x}^0) = \mathbf{X}(t - t^0; \mathbf{x}^0)$. We restrict ourselves to the tracking of the particle path until it strikes an infinite plane positioned normal to the axis of mean flow, at x_1 , and introduce the following coordinates to put transport in the x_1 -direction in terms of travel-time and transport in the remaining two directions in terms of deviations from the starting point \mathbf{x}^0 .

$$T(x; x_0) = t - t_0 \quad (19.39)$$

$$\eta(x; x_0) = X_2(\tau; x_0) \quad (19.40)$$

$$\xi(x; x_0) = X_3(\tau; x_0) \quad (19.41)$$

The travel-time function T is defined as the inverse of the particle position function \mathbf{X} . The new coordinate system satisfies by construction $\tau = 0$, $\eta = x_2$, and $\xi = x_3$ for $x_1 = x_1^0$, and

$$\frac{d\tau}{dx_1} = \frac{1}{v_1(x_1, \eta, \xi)} \quad (19.42a)$$

$$\frac{d\eta}{dx_1} = \frac{d\eta}{d\tau} \frac{d\tau}{dx_1} = \frac{v_2(x_1, \eta, \xi)}{v_1(x_1, \eta, \xi)} \quad (19.42b)$$

$$\frac{d\xi}{dx_1} = \frac{d\xi}{d\tau} \frac{d\tau}{dx_1} = \frac{v_3(x_1, \eta, \xi)}{v_1(x_1, \eta, \xi)} \quad (19.42c)$$

Thus $x_2 = \eta(x_1, \mathbf{x}_0)$ and $x_3 = \xi(x_1, \mathbf{x}_0)$ are equations of streamlines (streamtube centerlines) for the particle path. In this coordinate system, along the streamlines, Equation 19.1a takes the form of Equation 19.30, with r given by Equation 19.1b [e.g., $r = -(\rho_b/n)\partial S/\partial t$]. Thus the transformation Equation 19.39 to Equation 19.42 formally translates the governing equations into each other. Cushman et al. (1996) have used this transformation to compare the spatial moments of a plume evolving according to Equation 19.1 obtained by both the Eulerian and Lagrangian (according to Dagan and Cvetkovic [1993], using Lassey's solution, [Equation 19.37]) approaches. In that case the zero and first spatial moments are equivalent between the two approaches, but the second moments of the Eulerian solution contain a linear dependence on time with coefficient in d_{ij} . In particular, the long-time transverse second moment is linear in time and otherwise constant, so the Eulerian approach predicts the plume width will grow linearly in time while the Lagrangian approach predicts the plume width will reach a constant.

Note that this discrepancy is not inherent to any fundamental difference between the approaches themselves; rather, it arises from the assumption of zero dispersion d_{ij} in the Lagrangian approach. In

the previous chapter, the importance of pore-scale mixing (e.g., nonzero d_{ij}) to average plume dilution is illustrated and in the previous section; the effect of mixing on scaling multicomponent nonlinear reactions is noted. This highlights a subtle aspect of mathematical scaling of processes in general. The total mass moved by pore-scale mixing is in general relatively small as the process is small scale in space and time (it is slow) — however it can play a significant role in the eventual large scale and longtime behavior of the system.

Definition of Symbols

Mathematical Adornments

- represents Fourier–Laplace transform where t is dual to ω and \mathbf{x} is dual to \mathbf{k}
- ~ represents Fourier space transform where \mathbf{x} is dual to \mathbf{k}
- represents expected value

Principal Actors

A_o	constant reference area for effective streamtube [L^2]
C	aqueous concentration [M/L^3]
C_0	initial condition for aqueous concentration [M/L^3]
$C_o, C_b, C_{mm}, C_{im}, C_T$	aqueous concentrations of oxygen, benzoate, mobile microbes immobile microbes, and bromide tracer, respectively
$D_{ij}(\mathbf{x} - \mathbf{y}, t - t')$	nonlocal dispersion tensor defined in Equation 19.4
d_{ij}	symmetric microscale dispersion tensor [L^2/T]
$\delta(t)$	Dirac-delta function
F	mass ratio of oxygen consumed per benzoate consumed by microbial degradation [1]
f	fluctuation in the natural log of $K = \overline{\ln K} - \ln K$
$\phi(X_1, v_0; t)$	joint probability distribution of X_1, v_0 , at time t
G	cumulative reactivity, integral of g along a streamtube, a distributed quantity
$g(x)$	reaction rate coefficient along a given streamtube, a random field
$H(t)$	Heaviside function, integral of $\delta(t)$
I_1	modified Bessel function of the first kind of order one
\mathbf{K}	hydraulic conductivity tensor [L/T]
K_r	desorption rate [$1/T$]
K_d	equilibrium distribution coefficient [L^3/M]
K_d^G	geometric mean of $K_d(x)$
k_d	fluctuation in K_d from its expected value $= \overline{K_d} - K_d$
K_f	$= K_r K_d$ is the forward sorption rate [$1/T$]
K_r	desorption rate [M/TL^3]
K_c	solute degradation rate in the solution [$1/T$]
K_s	solute degradation rate in the sorbed phase [$1/T$]
K_o	half-saturation constant for oxygen biodegradation [M/L^3]
K_b	half-saturation constant for benzoate biodegradation [M/L^3]
μ	specific growth rate for microbial degradation of benzoate [$1/T$]
$p(\tau; x)$	distribution of solute flux over travel-times
R	$= 1 + \rho_b K_d/n$ is the retardation factor [1]
ρ_b	bulk density of the aquifer material [M/L^3]
n	porosity
$r(C)$	algebraic relation specifying S as a function of C at equilibrium
r	reaction rate in streamtube [$M/L^3 T$]
S	sorbed concentration per unit mass of solid [M/M]

S_0	initial condition for sorbed concentration [M/M]
t'	nondimensional time = $t \bar{V} / \bar{R}l$ where $\bar{R} = 1 + \bar{K}_d$
τ	travel-time [T] as a distributed variable
$T(x; x_0)$	deterministic travel-time along streamtube from x_0 to x
\mathbf{V}	porewater velocity vector [L/T]
\mathbf{v}	fluctuation in \mathbf{V} from its expected value = $\bar{\mathbf{V}} - \mathbf{V}$
$v(x)$	Lagrangian velocity along a streamtube [L/T]
$\Xi(T)$	inverse function of travel-time (delivering x)
Y	yield coefficient for microbial degradation of benzoate [M/M]
W	normally distributed random space function

Variances, Symbolically

σ_a^2 the variance of the fluctuation a of the quantity A

Covariances, Symbolically

\overline{ab}	the expected value of the product of the fluctuations a and b (of A and B , respectively), that is, the covariance. Some examples:
\overline{ff}	fluctuating log-conductivity covariance
$\overline{fk_d}$	fluctuating log-conductivity-partition coefficient covariance
$\overline{v_i v_j}$	fluctuating velocity covariance
$\overline{k_d v_j}$	fluctuating velocity-partition coefficient covariance
$\overline{k_d k_d}$	fluctuating distribution coefficient covariance
\overline{WW}	covariance of W

Integral Scales (a.k.a., Correlation Lengths)

l_w is the integral scale of W , [L]

l is the correlation length of $\ln K$ [L]

Glossary

Cumulative reactivity: integral of *reactivity* along a streamtube.

Equilibrium reaction: a chemical or biological transformation that occurs at a rate much faster than the other fate and transport processes governing a particular species.

Eulerian: referring to a coordinate system that is independent of the motion of groundwater, solutes, and other items of interest, that is, fixed.

First-order kinetics: a means of expressing the rate of a kinetically controlled reaction as proportional to the difference between the current state of the reactants and their equilibrium point.

Fluctuation: departure of the value of a property from some defined mean or expected value of that property.

Half-saturation constant: parameter of the Monod kinetic rate formula for biodegradation.

Heterogeneous (property): a property (e.g., porosity) whose value varies with spatial location.

Heterogeneous reaction: a chemical or biological transformation that involves a change of phase (typically between solid and aqueous phases, such as a sorption reaction).

Homogeneous (property): a property (e.g., porosity) whose value is not dependent on spatial location.

Homogeneous reaction: a chemical or biological transformation that involves a single phase.

Inhomogeneous: heterogeneous.

Isotherm: a means of expressing the equilibrium relationship among reactants and products of an equilibrium reaction (defined at a given temperature, thus the name), typically as an algebraic relation.

Kinetically controlled reaction: a chemical or biological transformation that occurs at a rate on the same order as that of the other fate and transport processes governing a particular species.

kinetics: the expressions of the rates of reactions for *kinetically controlled* reactions.

Lagrangian: referring to a coordinate system that is associated with a point moving with the groundwater flow.

Monte Carlo: a computational strategy for dealing with uncertainty, where unknown parameters are characterized by some frequency distribution and repeated simulations are made using values chosen from such distribution(s).

Operator-splitting: a computational scheme wherein distinct fate and transport processes are treated within a given time step by overlaying their effects sequentially for that time step; this is an approximation.

Pseudokinetic: rates of transformation that appear kinetically controlled but are associated with (multiple) equilibrium reactions in heterogeneous media.

Random: uncertain.

Reactivity: reactive site density (sites per bulk volume) associated with surfaces within aquifer materials.

Streamtube: REV-scale (e.g., super-pore scale) path associated with streamlines of transport that is steady state in direction.

Travel-time distribution: distribution of fractional flux over travel-time from start location to observation location.

Acknowledgments

This work was supported by the Environmental Management Science Program, Office of Health and Environmental Research and Office of Science and Technology, United States Department of Energy. The second author was supported by the National Science Foundation through the project grant no. 0420374 entitled “Biogeochemical cycling of heavy metals in lake Coeur d’Alene sediments: The role of indigenous microbial communities.”

References

- Andricevic, R. 1995. Comment on “Stochastic analysis of the transport of kinetically sorbing solutes in aquifers with randomly heterogeneous hydraulic conductivity,” by H.A.M. Quinodoz and A.J. Valocchi. *Water Resour. Res.* 31(1):237–243.
- Bellin, A., Rinaldo, A., Bosma, W.J.P., Van Der Zee, S.E.A.T.M., and Rubin, Y. 1993. Linear equilibrium adsorbing solute transport in physically and chemically heterogeneous porous formations: 1. Analytical solutions. *Water Resour. Res.* 29(12):4019–4030.
- Berglund, S. and Cvetkovic, V. 1996. Contaminant displacement in aquifers: coupled effects of flow heterogeneity and nonlinear sorption. *Water Resour. Res.* 32(1):23–32.
- Bhattacharya, R.N. 1982. On the functional central limit theorem and the law of integrated logarithms for Markov processes. *Z. Wahrsch. Verw. Gebiete* 60:185.
- Bhattacharya, R.N. and Gupta, V.K. 1990. Application of central limit theorems to solute dispersion in saturated porous media: from kinetic to field scales. In *Dynamics of Fluids in Hierarchical Porous Media*, J.H. Cushman (ed.), Chapter IV, Academic, San Diego, CA.
- Bhattacharjee, S., Ryan, J.N., and Elimelech, M. 2002. Virus transport in physically and geochemically heterogeneous subsurface porous media. *J. Contam. Hydrol.* 57:161–187.

- Bierkens, M.F.P. and Weerts, H.J.T. 1994. Block conductivity of cross-bedded fluvial sediments. *Water Resour. Res.* 30:2665–2678.
- Burr, D.T., Sudicky, E.A., and Naff, R.L. 1994. Non-reactive and reactive solute transport in three-dimensional heterogeneous porous media: mean displacement, plume spreading, and uncertainty. *Water Resour. Res.* 30:791–815.
- Chen, W. and Wagenet, R.J. 1995. Solute transport in porous media with sorption-site heterogeneity. *Environ. Sci. Technol.* 29:2725–2734.
- Chen, J.Y., Ko, C.H., Bhattacharjee, S., and Elimelech, M. 2001. Role of spatial distribution of porous medium surface charge heterogeneity in colloid transport. *Colloid Surf. A — Physicochem. Eng. Asp.* 191:3–15.
- Chiang, C.Y., Dawson, C.N., and Wheeler, M.F. 1990. Modeling of *in situ* bioremediation of organic compounds in groundwater. *Tech. Rep. TR90-31*, Rice University.
- Chilakapati, A. 1995. A simulator for reactive flow and transport of groundwater contaminants. PNNL-10636, Pacific Northwest National Laboratory.
- Chrysikopoulos, V.D., Kitanidis, P.K., and Roberts, P.V. 1990. Analysis of one-dimensional solute transport through media with spatially variable retardation factor. *Water Resour. Res.* 26(3):437–446.
- Cirpka, O.A. and Kitanidis, P.K. 2000. An advective-dispersive stream tube approach for the transfer of conservative-tracer data to reactive transport. *Water Resour. Res.* 36:1209–1220.
- Cunningham, J.A., Werth, C.J., Reinhard, M., and Roberts, P.V. 1997. Effects of grain-scale mass transfer on the transport of volatile organics through sediments 1. Model development. *Water Resour. Res.* 33(12):2713–2726.
- Cushman, J.H. (Ed.). 1990. *Dynamics of Fluids in Hierarchical Porous Media*. Academic Press, London.
- Cushman, J.H., Bennethum, L.S., and Hu, B.X. 2002. A primer on upscaling tools for porous media. *Adv. Water Resour.* 25:1043–1067.
- Cushman, J.H., Hu, B.X., and Deng, F.-W. 1996. Comparison of Eulerian to Lagrangian expected spatial moments for transport in a heterogeneous porous medium with deterministic linear nonequilibrium sorption. *Chem. Eng. Commun.* 148–150:5–21.
- Cvetkovic, V. and Dagan, G.. 1996. Reactive transport and immiscible flow in geological media. II. Applications. *Proc. R. Soc. Lond. A* 452:303–328.
- Cvetkovic, V., Dagan, G., and Cheng, H. 1998. Contaminant transport in aquifers with spatially variable hydraulic and sorption properties. *Proc. R. Soc. Lond. A — Math. Phys. Eng. Sci.* 454:2173–2207.
- Cvetkovic, V., Painter, S., Outters, N., and Selroos, J.O. 2004. Stochastic simulation of radionuclide migration in discretely fractured rock near the Äspö Hard Rock Laboratory. *Water Resour. Res.* 40(2), Art. No. W02404.
- Cvetkovic, V., Selroos, J.O., and Cheng, H. 1999. Transport of reactive tracers in rock fractures. *J. Fluid Mech.* 378:335–356.
- Dagan, G. 1989. *Flow and Transport in Porous Media*. Springer-Verlag, New York.
- Dagan, G. 1991. Dispersion of a passive solute in non-ergodic transport by steady velocity fields in heterogeneous formations. *J. Fluid Mech.* 233:197–210.
- Dagan, G. and Cvetkovic, V. 1993. Spatial moments of a kinetically sorbing solute plume in a heterogeneous aquifer. *Water Resour. Res.* 29(12):4053–4061.
- Dagan, G. and Cvetkovic, V. 1996. Reactive transport and immiscible flow in geological media. II. General theory. *Proc. R. Soc. Lond. A* 452:285–301.
- Dagan, G., Cvetkovic, V., and Shapiro, A. 1992. A solute-flux approach to transport in heterogeneous formations. 1. The general framework. *Water Resour. Res.* 28:1369–1376.
- Dagan, G. and Nguyen, V.. 1989. A comparison of travel time and concentration approaches to modeling transport by groundwater. *J. Contam. Hydrol.* 4:79–91.
- DeFlaun, M., Balkwill, D., Dobbs, F., Fredrickson, J., Fuller, M., Green, M., Ginn, T., Griffin, T., Holben, W., Hubbard, S., Johnson, W., Long, P., Majer, E., McInerney, M., Onstott, T., Phelps, T., Scheibe, T., Swift, D., White, D., and Wobber, F. 2001. Breakthroughs in field-scale bacterial transport. *EOS Transac. AGU* 82(38):417,423–425.

- Dzombak, D.A. and Morel, F.M.M. 1990. *Surface Complex Modeling; Hydrous Ferric Oxide*. Wiley-Interscience, New York.
- Elimelech, M., Nagai, M., Ko, C.H., and Ryan, J.N. 2000. Relative insignificance of mineral grain zeta potential to colloid transport in geochemically heterogeneous porous media. *Environ. Sci. Technol.* 34:2143–2148.
- Fetter, C.W. 1993. *Contaminant Hydrology*. Macmillian Publishing Co., New York.
- Gear, C.W. 1988. Differential-algebraic equation index transformations. *SIAM J. Sci. Comput.* 9(1):39–47.
- Gear, C.W. and Petzold, L.R. 1984. ODE methods for the solution of differential/algebraic systems. *SIAM J. Numer. Anal.* 21:716–728.
- Gelhar, L.W. and Axness, C.L. 1983. Three-dimensional stochastic analysis of macrodispersion in aquifers. *Water Resour. Res.* 19:161–190.
- Ginn, T.R. 2001. Stochastic-convective transport with nonlinear reactions and mixing: finite streamtube ensemble formulation for multicomponent reaction systems with intra-streamtube dispersion. *J. Contam. Hydrol.* 47:1–28.
- Ginn, T.R., Murphy, E.M., Chilakapati, A., and Seboonruang, U. 2001. Stochastic-convective transport with nonlinear reactions and mixing: application to intermediate-scale experiments in aerobic biodegradation in saturated porous media. *J. Contam. Hydrol.* 48:121–149.
- Ginn, T.R., Simmons, C.S., and Wood, B.D. 1995. Stochastic-convective transport with nonlinear reaction: biodegradation and microbial growth. *Water Resour. Res.* 31:2689–2701.
- Haggerty, R. and Gorelick, S.M. 1995. Multiple-rate mass transfer for modeling diffusion and surface reactions in heterogeneous media. *Water Resour. Res.* 31(10):2383–2400.
- Hassan, A.E., Cushman, J.H., and Delleur, J.W. 1998. The significance of porosity variability to flow and transport in random heterogeneous porous media. *Water Resour. Res.* 34(9):2249–2259.
- Hu, B.X., Cushman, J.H., and Deng, F.W. 1997. Nonlocal theory of reactive transport with physical, chemical, and biological heterogeneity. *Adv. Water Resour.* 20(5–6):293–308.
- Hu, B.X., Deng, F.W., and Cushman, J.H. 1995. A nonlocal theory of reactive transport with physical and chemical heterogeneity: linear nonequilibrium sorption with random K_d . *Water Resour. Res.* 31(9):2219–2237.
- Kabala, Z. and Sposito, G. 1991. A stochastic model of reactive solute transport with time-varying velocity in a heterogeneous aquifer. *Water Resour. Res.* 27(3):341–350.
- Kabala, Z. and Sposito, G. 1994. Statistical moments of reactive solute concentration in a heterogeneous aquifer. *Water Resour. Res.* 30(3):759–768.
- Karickhoff, S. 1983. Organic pollutant sorption in aquatic systems. *J. Hydraul. Eng.* 110:707–733.
- Knapp, E.P., Herman, J.S., Hornberger, G.M., and Mills, A.L. 1998. The effect of distribution of iron-oxyhydroxide grain coatings on the transport of bacterial cells in porous media. *Environ. Geol.* 33:243–248.
- Lichtner, P.C., Steefel, C.I., and Oelkers, E.H. (eds.) 1996. *Reactive Transport in Porous Media; Reviews in Mineralogy, Volume 34*. P.H. Ribbe, series editor. Mineralogical Society of America, Washington, D.C.
- Lichtner, P.C. and Tartakovsky, D.M. 2003. Upscaled effective rate constant for heterogeneous reactions, *Stochast. Environ. Res. Risk Assessment (SERRA)* 17(6):419–429.
- Metzler, R. and Klafter, J. 2000. The random walk's guide to anomalous diffusion: a fractional dynamics approach. *Phys. Rep.* 339:1–77.
- Miralles-Wilhelm, F. and Gelhar, L.W. 1996. Stochastic analysis of transport and decay of a solute in heterogeneous aquifers. *Water Resour. Res.* 32(12):3451–3459.
- Murphy, E.M., Ginn, T.R., Chilakapati, A., Resch, C.T., Phillips, J.L., and Wietsma, T.W. 1997. The influence of physical heterogeneity on microbial degradation and distribution in porous media. *Water Resour. Res.* 33:1087–1104.
- Novak, C.F. and Sevougian, S.D. 1993. Propagation of dissolution/precipitation waves in porous media. In *Migration and Fate of Pollutants in Soils and Subsoils*, D. Petruzzelli, F. G. Helffereich (eds.) NATO ASI Series, Series GL Ecological Sciences, vol. 32. Springer-Verlag, Berlin.

- Painter S., Cvetkovic, V., and Turner, D.R. 2001. Effect of heterogeneity on radionuclide retardation in the alluvial aquifer near Yucca Mountain, Nevada. *Ground Water* 39:326–338.
- Park, M., Kleinfelter, N., and Cushman, J.H. 2005a. Scaling laws and Fokker–Planck equations for 3-dimensional porous media with fractal mesoscale. *SIAM Multiscale Model. Simul.* 4(4): 1233–1244.
- Park, M., Kleinfelter, N., and Cushman, J.H. 2005b. Renormalizing chaotic dynamics in fractal porous media with application to microbe motility. *Geophysical Res. Ltr.* 33(1), Art. No. L01401.
- Quinodoz, H.A.M. and Valocchi, A.J. 1993. Stochastic analysis of the transport of kinetically sorbing solutes in aquifers with randomly heterogeneous hydraulic conductivity. *Water Resour. Res.* 29(9):3227–3240.
- Robinson B.A. and Viswanathan, H.S. 2003. Application of the theory of micromixing to groundwater reactive transport models, *Water Resour. Res.* 39(11).
- Rubin, J. 1983. Transport of reacting solutes in porous media: relationship between mathematical nature of problem formation and chemical nature of reactions. *Water Resour. Res.* 19(5):1231–1252.
- Rubin, J. 1990. Solute transport with multisegment, equilibrium-controlled reactions: a feed forward simulation method. *Water Resour. Res.* 26(9):2029–2055.
- Ryan, J.N., Elimelech, M., Ard, R.A., Harvey, R.W., and Johnson, P.R. 1999. Bacteriophage PRD1 and silica colloid transport and recovery in an iron oxide-coated sand aquifer. *Environ. Sci. Technol.* 33:63–73.
- Scheibe, T.D. and Freyberg, D.L. 1995. Use of sedimentological information for geometric simulation of natural porous media structure. *Water Resour. Res.* 31:3259–3270.
- Schulmeister, M.K., Healey, J.M., McCall, G.W., Birk, S.M and Butler, Jr, J.J. 2003. High-resolution chemical characterization of an alluvial aquifer. In *Modelcare 2002: Calibration and Reliability in Groundwater Modeling: A Few Steps Closer to Reality*, K. Kovar and Z. Hzkal (eds.), IAHS Publication 277, International Association of Hydrological Sciences, Wallingford, UK, pp. 419–423.
- Simmons, C.S. 1982. A stochastic–convective transport representation of dispersion in one-dimensional porous media systems. *Water Resour. Res.* 18:1193–1214.
- Simmons, C.S., Ginn, T.R., and Wood, B.D. 1995. Stochastic–convective transport with nonlinear reaction: mathematical framework. *Water Resour. Res.* 31:2674–2688.
- Sposito, G. 1993. *Chemical Equilibria and Kinetics in Soils*. Oxford University Press, Inc., New York.
- Steefel, C.I. and MacQuarrie, K.T.B. 1996. Approaches to modeling of reactive transport in porous media. In *Reactive Transport in Porous Media; Reviews in Mineralogy, Volume 34*. Lichtner et al. (eds.), P.H. Ribbe, series editor. Mineralogical Society of America, Washington, D.C.
- Sudicky, E.A., Schellenberg, S.L., and MacQuarrie, K.T.B. 1990. Assessment of the behavior of conservative and biodegradable solutes in heterogeneous porous media. In *Dynamics of Fluids in Hierarchical Porous Media*. J.H. Cushman (ed.), pp. 429–462. Academic, New York.
- Szecsody, J.E., Chilakapati, A., Zachara, J.M., and Garvin, A.L. 1998. Influence of iron oxide inclusion shape on $\text{Co}^{II/III}$ EDTA reactive transport through spatially heterogeneous sediment. *Water Resour. Res.* 34(10):2501–2514.
- Thiele, M.R., Blunt, M.J., and Orr, Jr, F.M. 1995a. Modeling flow in heterogeneous media using streamtubes I. Miscible and immiscible displacements. *In Situ* 19:299–339.
- Thiele, M.R., Blunt, M.J., and Orr, Jr, F.M. 1995b. Modeling flow in heterogeneous media using streamtubes II. Compositional displacements. *In Situ* 19:367–391.
- Tompson, A.F.B., Schafer-Perini, A.L., and Smith, R.W. 1996. Impacts of physical and chemical heterogeneity on co-contaminant transport in a sandy porous medium. *Water Resour. Res.* 32(4):801–818.
- Valocchi, A.J. 1990. Use of temporal moment analysis to study reactive solute transport in aggregated porous media. *Geoderma* 46:233–247.
- Valocchi, A.J. and Malmstead, M. 1992. Accuracy of operator splitting for advection–convection–reaction problems. *Water Resour. Res.* 28(5):1471–1476.

- Villermaux, J. 1981. Theory of linear chromatography, In *Percolation Processes, Theory and Applications*, A.E. Rodrigques and D. Tondeur (eds.), NATO ASI Series, Serial E, Vol. 33, pp. 83–140.
- Viswanathan, H.S. and Valocchi, A.J. 2004. Comparison os streamtube and three-dimensional models of reactive transport in heterogeneous media. *J. Hydraul. Res.* 42:141–145.
- Widdowson, M.A., Molz, F.J., and Benefield, L.D. 1988. A numerical transport model for oxygen- and nitrate-based respiration linked to substrate and nutrient availability in porous media. *Water Resour. Res.* 24(9):1553–1565.
- Wise, W.R. 1993. Effects of laboratory-scale variability upon batch and column determinations of nonlinearly sorptive behavior in porous media. *Water Resour. Res.* 29:2983–2992.
- Wood, W.W., Kramer, T.P., and Hearn, Jr, P.P. 1990. Intergranular diffusion: an important mechanism influencing solute transport in clastic aquifers? *Science* 247:1569–1572.
- Yeh, G.T. and Tripathi, V.S. 1989. A critical evaluation of recent developments in hydrological transport models of reactive multichemical components. *Water Resour. Res.* 25(1):93–108.

Further Reading

Regarding Reactive Transport Modeling in Hydrogeology

- Schulz, H. D. and Teutsch, G. (Ed.). 2002. *Geochemical Processes: Conceptual Models for Reactive Transport in Soil and Groundwater. Research Report*, Wiley Publishing Company.
- Lichtner, P. C., Steefel, C. I., and Oelkers, E. H. (Ed.). 1997. *Reactive Transport in Porous Media*, Volume 34, Reviews in Mineralogy, American Mineralogical Society.

Regarding Eulerian Upscaling Approaches to Reactive Transport

- Cushman, J. H. (Ed.). 1990. *Dynamics of Fluids in Hierarchical Porous Media*. Academic Press, London.
- Cushman, J.H., Bennethum, L.S., and Hu, B.X. 2002. A primer on upscaling tools for porous media. *Adv. Water Resour.* 25, 1043–1067.

Regarding Lagrangian/Streamtube Approaches to Reactive Transport

- Cirpka, O.A. and Kitanidis, P.K. 2000. An advective-dispersive stream tube approach for the transfer of conservative-tracer data to reactive transport. *Water Resour. Res.* 36:1209–1220.
- Cvetkovic, V. and G. Dagan 1996. Reactive transport and immiscible flow in geological media. II. Applications. *Proc. R. Soc. Lond. A* 452:303–328.
- Ginn, T. R. 2001. Stochastic-convective transport with nonlinear reactions and mixing: Finite streamtube ensemble formulation for multicomponent reaction systems with intra-streamtube dispersion, *J. Contam. Hydrol.* 47:1–28.

20

Groundwater Flow and Solute Transport in Fractured Media

20.1	Introduction.....	20-1
20.2	Structural Geology of Fractured Rock	20-2
20.3	Fundamentals of Groundwater Flow and Solute Transport in a Single Fracture	20-4
	Groundwater Flow in a Single Fracture • Solute Transport in a Single Fracture	
20.4	Field Techniques	20-9
	Fracture Mapping • Core Analysis • Borehole Geophysical Methods • Hydraulic Testing Methods • Single-Well Hydraulic Tests • Multiwell Hydraulic Tests • Point Dilution Method • Tracer Experiments • Borehole Instrumentation	
20.5	Conceptual Models	20-26
	Conceptual Models for a Single Fracture • Conceptual Models for a Fracture Network	
20.6	Modeling Flow and Transport in Fractures and Fracture Networks	20-31
	Flow in Discrete Fracture Networks • Solute Transport in a Discrete Fracture Network	
	Glossary.....	20-37
	References	20-38
	Further Information	20-43

K.S. Novakowski
Queen's University

P.A. Lapcevic
*British Columbia
Ministry of the Environment*

E.A. Sudicky
University of Waterloo

20.1 Introduction

Many of the world's largest aquifers reside in fractured consolidated media. Studies have shown that even media traditionally considered to be of low permeability, such as shale and granite, are fractured to the extent that significant groundwater flow may occur. In North America, there are many locales at which bedrock of either sedimentary or crystalline origin is used for either domestic or commercial water supply. In many geological settings, the lack of a protective layer of low-permeability clays or clay tills makes these water supplies particularly vulnerable to groundwater contamination.

Understanding groundwater flow in fractured consolidated media has long been important when undertaking engineering tasks such as dam construction, mine development, the abstraction of petroleum,

slope stabilization, and the construction of foundations. To study groundwater flow in support of these tasks, often only relatively simple measurements of permeability conducted *in situ* are required. More recently, the discovery of groundwater contamination in fractured bedrock materials has led to the advent of many new investigative techniques that are quite different from those previously used in the geotechnical and petroleum industries. In particular, the focus of most hydrogeological investigations has been on the characterization of the transport properties of the higher permeability fractures in the rock mass.

On many occasions, groundwater investigations benefit from the application of a numerical model that is used to simulate the flow system. However, because groundwater flow is primarily governed by discrete fractures at the local scale, traditional mathematical methods, in which it is assumed that the hydrogeological properties vary in a smooth and continuous fashion (a single continuum), provide a poor approximation. Only at large, regional scales of groundwater flow or where fracturing is very dense might continuum methods be applied in fractured media. Further, the often complex and heterogeneous arrangement of discrete fractures makes for difficulties in the representation of these features on an individual basis. There are several modeling approaches that can be used to circumvent this problem, including the use of stochastic or mixed deterministic-stochastic techniques (e.g., Smith and Schwartz 1984) and the use of scale-dependent multi-continua (e.g., Bear, 1993). However, numerical models that can represent two-dimensional groundwater flow and solute transport in each discrete fracture within a three-dimensional randomly fractured domain of practical size have yet to be widely applied. As it is extremely expensive to characterize a field site to the degree necessary to support this type of model, a compromise such as the alternate approaches suggested above will almost always be required.

In this chapter, we will try to provide a basic theory and practical advice for guiding investigations of the hydrogeology of fractured rock sites. Our aim has been to focus on the hydrogeological representation of fractured rock systems at scales below which an equivalent porous media (EPM) approach is adequate. To this end, the modeling and field sections will outline techniques to characterize and simulate discrete fractures and fracture networks in the subsurface. A basic introduction to structural geology will be provided to assist in determining how the features that govern groundwater flow in fractured media are formed and how to go about finding them. A brief outline of the principal processes involved in the flow of groundwater and transport of contaminants in individual fractures is presented. In addition, some methods commonly used in the field characterization of fractured rock will be described. These elements become the building blocks for the development of a conceptual model of groundwater flow. A discussion of conceptual models for both single fractures and fracture networks is included. Finally, an overview of analytical and numerical modeling techniques for fractured rock will be presented.

20.2 Structural Geology of Fractured Rock

In this chapter, the term fracture is used to refer to any discontinuity in a solid material. In geological terms, a fracture is any planar or curviplanar discontinuity that has formed as a result of a process of brittle deformation in the earth's crust (Engelder et al., 1993).

Planes of weakness in rock respond to changing stresses in the earth's crust by fracturing in one or more ways depending on the direction of the maximum stress and the rock type. Three broad types of fractures have been identified based on the motion involved in the formation of the fracture. Opening mode (Mode I) cracks form by the separation of the crack walls, under the action of tensile stresses normal to the crack plane (Figure 20.1a). Mode I fractures are thus extension fractures in which the fracture walls have displaced normal to the crack plane and no appreciable displacement parallel to the fracture plane has occurred. Sliding mode (Mode II) cracks propagate by mutual shearing of the crack walls with the shearing oriented in the direction normal to the crack front (Figure 20.1b). Tearing mode (Mode III) cracks advance when the crack walls are subject to a shearing aligned parallel to the crack front (Figure 20.1c). Mode II and III features are defined as shear fractures.

Joints (Mode I fractures) are the most common type of fracture structure and are found in all types of rocks as well as in partly consolidated to unconsolidated sediments. Joint sets are characterized as

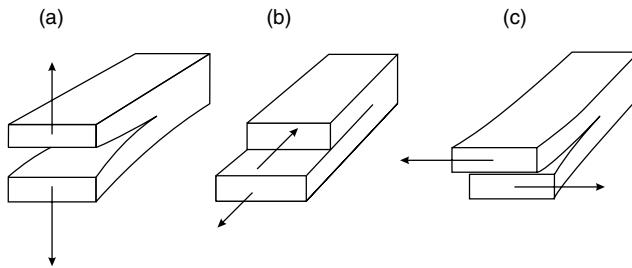


FIGURE 20.1 Schematic diagram showing mechanism of fracture formation: (a) Mode I (opening), (b) Mode II (sliding), (c) Mode III (tearing). (From Atkinson, B.K., 1987. *Fracture Mechanics of Rock*, Academic Press, Orlando, FL, 534 pp.)

systematic or nonsystematic. Joints belonging to a set maintaining consistent characteristics such as orientation, morphology, mineralization, and distribution between outcrops in a local and regional scale are defined as systematic (Engelder et al., 1993). Nonsystematic joints have highly curved and irregular fracture surfaces and are not easily related to a recognizable stress field.

Groundwater movement in fractured rock depends on discontinuities at a variety of scales ranging from microcracks (micrometers, μm , in length and width) to faults (kilometers in length and meters in width). Fracturing in a large-scale rock mass can be loosely classified on the basis of three scales of discontinuity (1) microscopic, (2) mesoscopic, and (3) megascopic.

At the microscopic scale, isolated or continuous disk-shaped microcracks ranging in length from 100 to 1000 μm and 1 to 2 μm in width are common to both crystalline and sedimentary rocks. Microcracks are normally Mode I fractures, and often form in a disk-shape. In some cases, particularly adjacent to shear fractures, microcracks may form under Mode II conditions.

In crystalline rocks, the total porosity of the rock matrix (i.e., the rock bounded by macroscale fractures) is due to microcracks and may range from 0.01 to 1.0%. In sedimentary rock, in addition to microcracks, significant porosity in the matrix may also occur due to incomplete cementation, the weathering of vugs, or imperfect lattice development. Matrix porosity in sedimentary rock may range from less than 1% to as much as 50%.

At the mesoscopic scale, fractures (or joints) ranging between <0.5 m and >1000 m in length with widths ranging from <10 μm to >10 mm predominate the groundwater flow system. These fractures are often grouped in sets having similar orientation that are related to the mode of fracture genesis. Fracture zones at this scale are defined as sets of closely spaced fractures contained within a narrow band. Shear fractures are also observed at this scale. A common and extremely important class of fractures at this scale is known as sheeting structure (Holzhauser, 1989). Sheet structures form due to the vertical expansion of the rock resulting from erosional unloading. Sheet structures are present in all types of rock, often found parallel to the earth's surface, and are common conduits for groundwater flow at the regional scale. In plutonic and metamorphic rocks they are often independent of the original macrostructure and fabric of the rock and are found at decreasing frequency with increasing depth (Trainer, 1988).

At the megascopic scale, faults and shear zones dominate the fracture morphology. A fault is defined as a planar discontinuity between two blocks of uniform rock where displacement has taken place parallel to the discontinuity. These are large-scale fractures with many associated features and may be either of high or low permeability depending on infilling and local stress fields. Faults may be of any orientation, although it is the sub-vertical features that are most easily identified through air-photo or map interpretation. Shear zones are similar to faults (i.e., a large-scale fracture zone that has undergone shearing) and are typically found in metamorphic terrain. These features are often of high permeability. In many fractured rock environments, the frequency and spacing of faults and shear zones is sparse. For example, Figure 20.2 illustrates in plan view the possible surface expression of a hypothetical distribution of vertical faults in the southern Ontario region of central Canada. These faults are postulated to vertically penetrate the entire thickness of the sedimentary strata that ranges from ~200 m to ~2 km, in this area (Sanford et al., 1985).

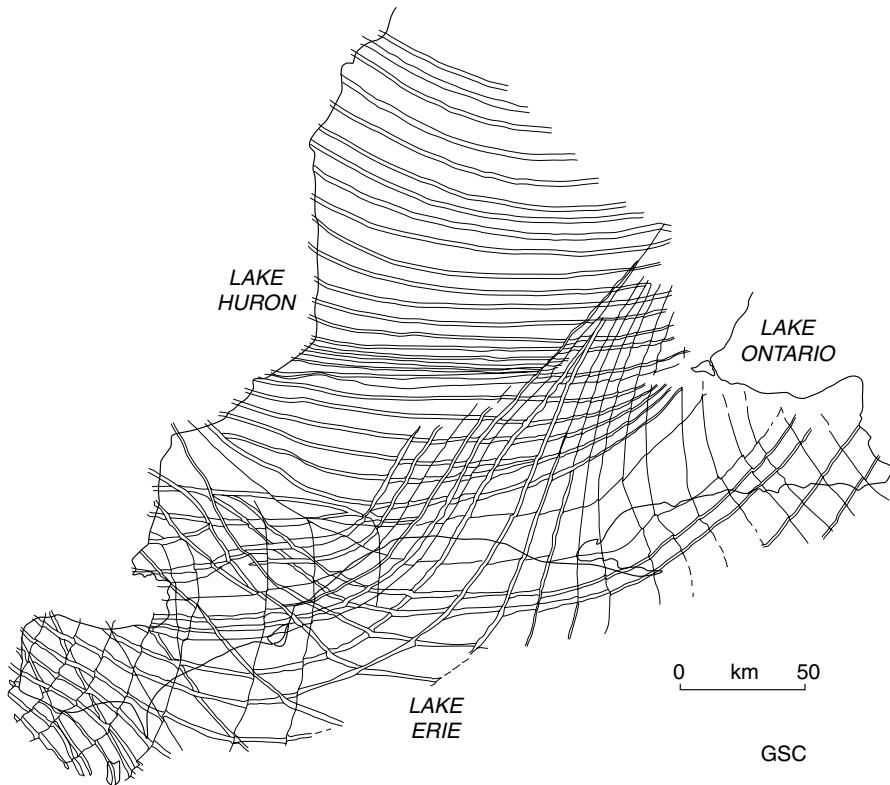


FIGURE 20.2 Distribution of regional faults in Southern Ontario. (From Sanford, B.V., Thompson, F.J., and McFall, G.H., 1985. Phanerozoic and recent tectonics of the Canadian Craton, in *The Geoscience Program — Proceedings of the Seventeenth Information Meeting of the Nuclear Fuel Waste Management Program*, Atomic Energy of Canada Limited, TR-299, pp. 334–352.)

Note that the spacing between faults is several to tens of kilometers. Thus, at the local to regional scale, any one of these faults may dominate the groundwater flow system.

Karst aquifer systems develop due to the extensive dissolution of fractured carbonate rock by flowing groundwater. The hydrogeology of this type of rock terrain is discussed in the following chapter.

20.3 Fundamentals of Groundwater Flow and Solute Transport in a Single Fracture

20.3.1 Groundwater Flow in a Single Fracture

Groundwater flow in a single fracture is often interpreted by assuming the fracture walls are analogous to parallel plates separated by a constant aperture. Using this analogy, solution of the Navier–Stokes equations for laminar flow of a viscous, incompressible fluid bounded by two smooth plates leads to an expression referred to as the cubic law. Figure 20.3 shows the parabolic distribution of velocity predicted by the solution to the Navier–Stokes equations in a cross section of a fracture having parallel walls. A rigorous mathematical development of the cubic law can be found in Bear (1972). Expressed as volumetric flow rate, Q , per unit drop in hydraulic head, Δh , the cubic law can be written as

$$\frac{Q}{\Delta h} = C(2b)^3 \quad (20.1)$$

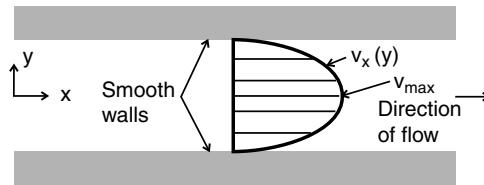


FIGURE 20.3 Schematic diagram showing the parabolic distribution of velocity in a fracture having parallel walls.

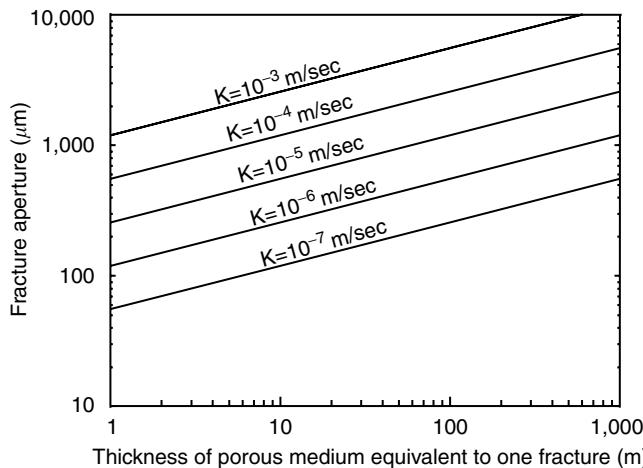


FIGURE 20.4 Comparison of aperture of a single fracture to equivalent thickness of porous media. (From de Marsily, G., 1986. *Quantitative Hydrogeology*, Academic Press, Orlando, FL.)

where $2b$ is fracture aperture and C is a constant related to the properties of the fluid and the geometry of the flow domain. Thus, the flow rate is proportional to the cube of the aperture, hence the term cubic law. For straight uniform flow, the constant of proportionality is

$$C = \frac{\rho g}{12\mu} \frac{W}{L} \quad (20.2)$$

where ρ is fluid density, g is acceleration of gravity, μ is kinematic viscosity, W is the width of the fracture, and L is the length of the fracture. For flow in radial domain, C is given as

$$C = \frac{\rho g}{12\mu} \frac{2\pi}{\ln(r_e/r_w)} \quad (20.3)$$

where r_w is well radius and r_e is radius of influence. By substitution of the cubic law into Darcy's law, the transmissivity (T_{fr}) and hydraulic conductivity (K_{fr}) of a fracture can be defined as

$$T_{fr} = \frac{\rho g}{12\mu} (2b)^3 = K_{fr} 2b \quad (20.4)$$

Thus, using Equation 20.4 we can relate terms typically used to describe the permeability of a porous medium to that for discrete fractures.

Figure 20.4 illustrates for a range of hydraulic conductivities, the thickness of porous medium that would be equivalent to a single fracture of a given aperture. For example, using Figure 20.4, we see that the volume of groundwater passing through a 10-m section of porous media having $K=10^{-4}$ m/sec is

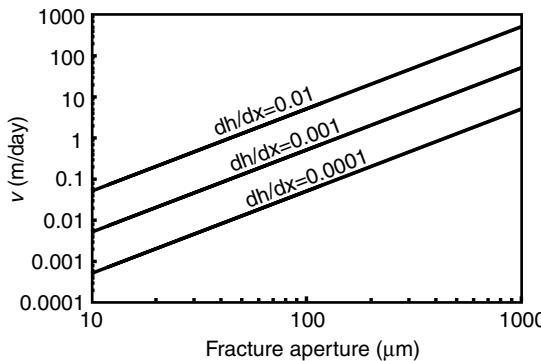


FIGURE 20.5 Groundwater velocity in a single fracture as predicted by the cubic law.

the same as can pass through a fracture less than 1 mm in width under the same driving force. Thus, even in sparsely fractured rock, considerable volumes of groundwater are moved about under typical hydraulic gradients.

The average groundwater velocity in a fracture can be found by combining Equation 20.4 and Darcy's Law:

$$\bar{v} = -\frac{\rho g}{12\mu} (2b)^2 \frac{dh}{dx} \quad (20.5)$$

Groundwater velocity predicted using the cubic law over a range of hydraulic gradients is presented in Figure 20.5. Note that even for very small hydraulic gradients, groundwater velocity in discrete fractures is very high in comparison to velocities typical of porous media.

As in open-channel or closed-pipe regimes, turbulent flow is possible where very large hydraulic gradients are present. To estimate the onset of turbulent flow in a discrete planar fracture, the transition from laminar to turbulent flow is defined using the Reynolds number, R_e :

$$R_e = \frac{\bar{v} D_h \rho}{\mu} \quad (20.6)$$

where \bar{v} is mean velocity, and D_h is the hydraulic diameter. For planar fractures, D_h is defined as:

$$D_h = 2 \cdot 2b \quad (20.7)$$

In addition, the relative roughness of a fracture, R_r , can be defined as

$$R_r = \frac{\varepsilon}{D_h} \quad (20.8)$$

where ε is the mean height of the asperities (the peaks in the topographic surface of the fracture walls). The transition from smooth to rough flow occurs at a R_r of 0.033 and the transition from laminar to turbulent flow at a R_e of 2300 (de Marsily, 1986). Using the expression for velocity (Equation 20.5) and the definition of R_e , it is found that extremely high hydraulic gradients are required to generate turbulent conditions in smooth fractures. For example, using an aperture of 1000 μm (on the higher end of the range typically found in nature) and a hydraulic gradient of 0.01 (very high end of the range observed in nature), a R_e of only 8.5 is calculated.

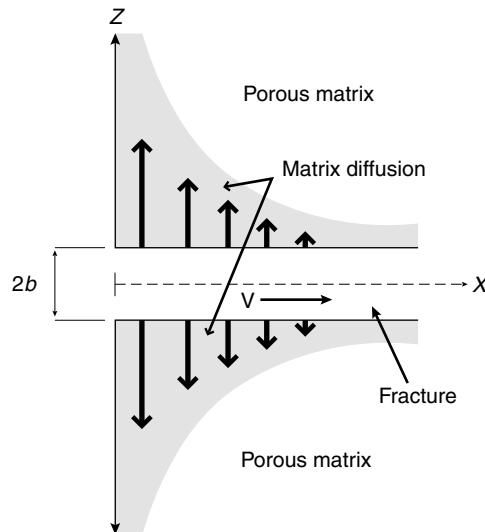


FIGURE 20.6 Schematic diagram of a single fracture in a porous rock.

20.3.2 Solute Transport in a Single Fracture

The transport of both conservative and reactive aqueous constituents in discrete fractures is governed by (1) hydrodynamic dispersion, (2) advection, (3) radioactive or biological decay, and (4) diffusion into the rock matrix. In addition, the transport of reactive compounds is influenced by geochemical processes such as acid–base reactions, oxidation–reduction reactions, precipitation–dissolution, or adsorption–desorption.

The governing equation for solute transport in a single fracture (Figure 20.6) which incorporates advective transport, longitudinal dispersion, molecular diffusion from the fracture into the matrix, adsorption onto the face of the matrix, adsorption within the matrix, and radioactive decay is written as (Tang et al., 1981):

$$2b \left[R_a \frac{\partial c}{\partial t} + \bar{v} \frac{\partial c}{\partial x} - D_L \frac{\partial^2 c}{\partial x^2} + \lambda R_a c \right] + 2q = 0 \quad 0 \leq x \leq \infty \quad (20.9)$$

where the \$x\$-coordinate is in the direction of the fracture axis, \$\lambda\$ is a decay constant, and \$q\$ is the diffusive flux perpendicular to the fracture axis. The retardation factor, \$R_a\$, is defined by

$$R_a = 1 + K_{df}/b \quad (20.10)$$

where \$K_{df}\$ is the distribution coefficient for the fracture walls defined as mass of solute adsorbed per unit area of surface divided by the concentration of solute in solution. To relate this to the distribution coefficient for the rock matrix, \$K_d\$, which is based on the bulk density of the material, an estimate of internal specific surface area of the medium, \$\gamma\$, is required. The specific surface area is defined as the area of the pore walls exposed to the fluid and has units of area per unit mass. Thus, the relation is given as

$$K_d = \gamma K_{df} \quad (20.11)$$

The hydrodynamic dispersion coefficient, D_L , is defined following the standard formulation for porous media:

$$D_L = \alpha_L \bar{v} + D' \quad (20.12)$$

where α_L is the longitudinal dispersivity in the direction of the fracture axis and D' is the effective diffusion coefficient which incorporates the geometric effect of the pathways in the pore space. The governing equation for the matrix as derived by Tang et al. (1981) is

$$\frac{\partial c'}{\partial t} - \frac{D'}{R'} \frac{\partial^2 c'}{\partial z^2} + \lambda c' = 0 \quad b \leq z \leq \infty \quad (20.13)$$

assuming a single fracture penetrating an infinite medium where c' is concentration in the matrix, and the matrix retardation coefficient R' is defined as

$$R' = 1 + \frac{\rho_b}{\theta_m} K_d \quad (20.14)$$

where ρ_b is the bulk density of the matrix, and θ_m is matrix porosity. The diffusive loss term q in Equation 20.9 represents diffusive mass flux crossing the fracture–matrix interface. Using Fick's first law this can be expressed as

$$q = -\theta_m D' \left. \frac{\partial c'}{\partial z} \right|_{z=b} \quad (20.15)$$

Thus, the final coupled equation for the fracture becomes

$$2b \left[R_a \frac{\partial c}{\partial t} + \bar{v} \frac{\partial c}{\partial x} - D_L \frac{\partial^2 c}{\partial x^2} + \lambda R_a c \right] - 2\theta_m D' \left. \frac{\partial c'}{\partial z} \right|_{z=b} = 0 \quad 0 \leq x \leq \infty \quad (20.16)$$

In porous and fractured rock, the loss of solute from the fractures into the rock matrix through the process of diffusion can be significant. To illustrate the effect of matrix diffusion, Figure 20.7 uses the solution

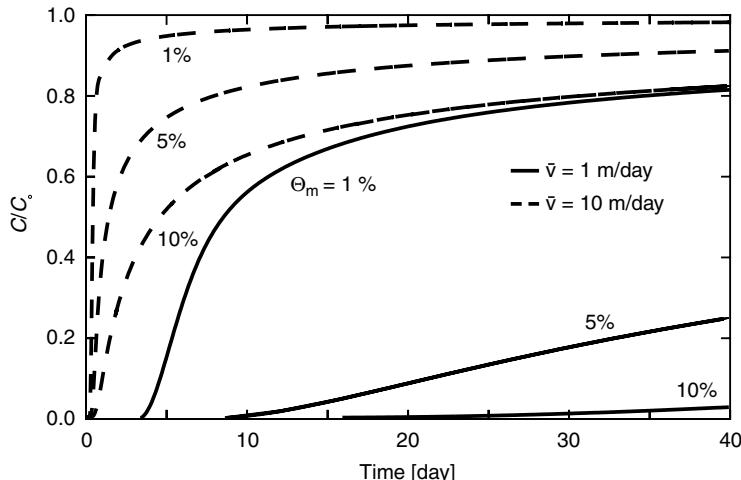


FIGURE 20.7 Dimensionless concentration profiles in a 500- μm fracture, 50 m from a continuous source, illustrating the influence of matrix porosity on solute transport in a single fracture.

to Equation 20.16 to show the relative concentration of solute in a single $500\text{ }\mu\text{m}$ fracture pervading a medium having a porosity of 1 to 10%. The velocity in the fracture ranges from 1 to 10 m/day and dispersivity is varied from 0.05 to 0.5 m. Concentration input is continuous and the point of measurement is 50 m from the source. It is immediately evident that the effect of solute loss to the matrix is profound. The effect is so significant that the influence of hydrodynamic dispersion in the longitudinal spreading of solute is completely overwhelmed (i.e., different α_L results in little change in the shape of the curve).

The loss of solute to the matrix results in an additional problem when undertaking the clean-up of a fractured porous aquifer. Removal of the source of aqueous contamination to a fracture network may be conducted via a number of means (e.g., pump-and-treat), however, the back diffusion process will continue to supply contamination from the matrix for many years, depending on the porosity of the matrix rock. This is shown in Figure 20.8 where Equation 20.16 is used to simulate the exposure of a single fracture-matrix system to a source of aqueous contamination (5000 ppb) for a period of 10 years. The fracture aperture is $100\text{ }\mu\text{m}$, the groundwater velocity is 1 m/day, the point of observation is 45 m down-gradient from the source, and the results for three matrix porosities (0.5%, 2%, and 10%) are shown. Although the influence of matrix diffusion at the lowest porosity does not appear significant, note that for a contaminant having a low drinking water limit (e.g., $5\text{--}10\text{ }\mu\text{g/L}$), the presence of contaminants above that limit will persist for many years following the cutoff of the source. At higher matrix porosity this problem is significantly exacerbated, with contamination persisting for many decades. It is the process of back diffusion that presently confounds efforts to remediate contaminated bedrock aquifers.

20.4 Field Techniques

To develop a defensible conceptual model for a given flow system in fractured rock, some characterization of the medium is necessary. The degree of characterization is dependent on a variety of factors including the inherent complexity of the fracture system (i.e., sedimentary versus crystalline rock), the type of conceptual model desired, and the amount of funding available. Because of the inherent heterogeneity of fractured rock masses (e.g., presence of sheeting structures), some characterization of the subsurface is essential in supporting conceptual models based only on surface measurements (i.e., fracture maps).

In the following, methods for the measurement of the surface and subsurface expression of fractures will be discussed. In addition, techniques for the measurement of the hydraulic parameters of both the fractures and unfractured matrix are discussed and an overview of tracer methods that are used

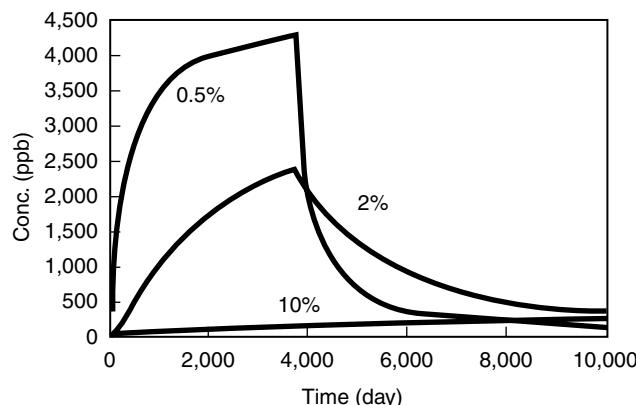


FIGURE 20.8 Concentration profiles in a $100\text{-}\mu\text{m}$ fracture, 45 m from a source having a concentration of 5000 ppb, lasting 10 years. Groundwater velocity is 1 m/day.

to measure solute transport parameters in fractured rock, is presented. Finally, permanent multilevel borehole instrumentation specifically designed for fractured rock investigations is described. An essential component of comprehensive investigations of groundwater flow and solute transport in fractured media is the geochemistry of the groundwater. A discussion of groundwater geochemistry in fractured media, however, is beyond the scope of this chapter. An example of using inorganic and isotopic geochemistry to conceptualise groundwater flow pathways can be found in Zanini et al. (2000).

20.4.1 Fracture Mapping

To determine the distribution of systematic fractures, outcrops, quarry walls, pavements, or mine tunnels are mapped to estimate the strike, dip, and spacing of individual fractures. In practise, reference lines are established on the rock surface using a measuring tape oriented from the north. The orientation, spacing, and trace length of all fractures intersecting these scanlines are measured. The bias introduced by the intersection of the fractures with an oriented scanline in this type of sampling must be considered during post-processing (LaPointe and Hudson, 1985). Other useful observations of the fractures on traces include the nature of the termination, mineralization, staining, and surface roughness. Qualitative observations of water or fluid seepage along fractures in road cuts or gorge faces can also be useful indicators of groundwater flow. It is important to note that in many cases surface fracture measurements will be reflective of near surface stress conditions or may be influenced by the removal of the overlying rock in quarrying operations. Consequently, physical estimation of fracture aperture cannot reliably be conducted from surface measurements. In addition, due to the presence of sheeting structure, the fracture measurements obtained from surface exposures may not agree with those obtained *in situ*.

Measurements of fracture orientation are usually presented in one of two ways (1) rose diagrams or (2) stereographic plots. Rose diagrams are histograms in which the frequency (usually as a percentage of total measurements) of joint orientations within 5–10° intervals is plotted by azimuth. Rose diagrams are commonly used in flat-lying sedimentary rock where most non-horizontal fractures are vertical in orientation. Figure 20.9 shows a rose diagram developed from core collected in a Silurian-aged dolostone. The diagram indicates the presence of a dominant fracture set oriented at approximately 100° and two secondary sets oriented at approximately 10° and 140°. In crystalline rock or deformed sedimentary rock with fracture sets of more random orientation, data are often presented as the projection of the poles of the planes (fractures) plotted on equal-area stereonets (Ragan, 1973). An example of this method is given in Figure 20.10 using fracture data collected in monzonitic gneiss. There are two dominant fracture sets and one minor set shown by the concentration of the poles.

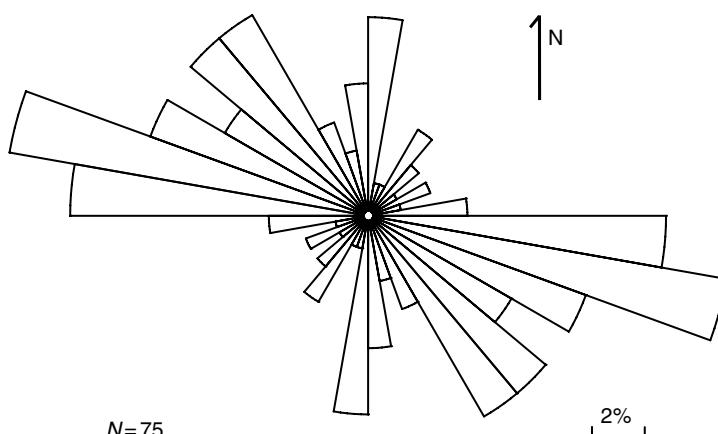


FIGURE 20.9 Rose diagram of orientations of vertical fractures intersecting seven boreholes to depths of 60 m in a flat-lying Silurian dolostone.

20.4.2 Core Analysis

Boreholes installed for groundwater investigations in fractured rock, using either percussion, tri-cone, or diamond drilling techniques, generally range in diameter from 48 to 152 mm. In North America, 76 or 96 mm diameter boreholes are generally drilled as test or monitoring wells while 152 mm or greater diameter boreholes are used for water supply. Rock core obtained from diamond-drilled boreholes can be used to determine the subsurface pattern of fractures. While it is common to install only vertical boreholes, for most hydrogeological site investigations, installing boreholes in the inclined orientation will maximize the number of non-horizontal and vertical fractures intersected. There is only a nominal cost differential between inclined and vertical boreholes and working in inclined boreholes is no more difficult than vertical boreholes. In flat-lying sedimentary rock, inclined boreholes may be of better quality and less prone to blockages due to spalling rock.

The physical examination of rock core can be used to identify potentially open fractures through which groundwater flows. Natural fractures can be distinguished from mechanical breaks (breaks resulting from drill vibration etc.) using the following criteria (1) roughness of the fracture surface (wavy irregular fracture traces in natural fractures contrast to sharp, clean mechanically induced breaks), (2) presence

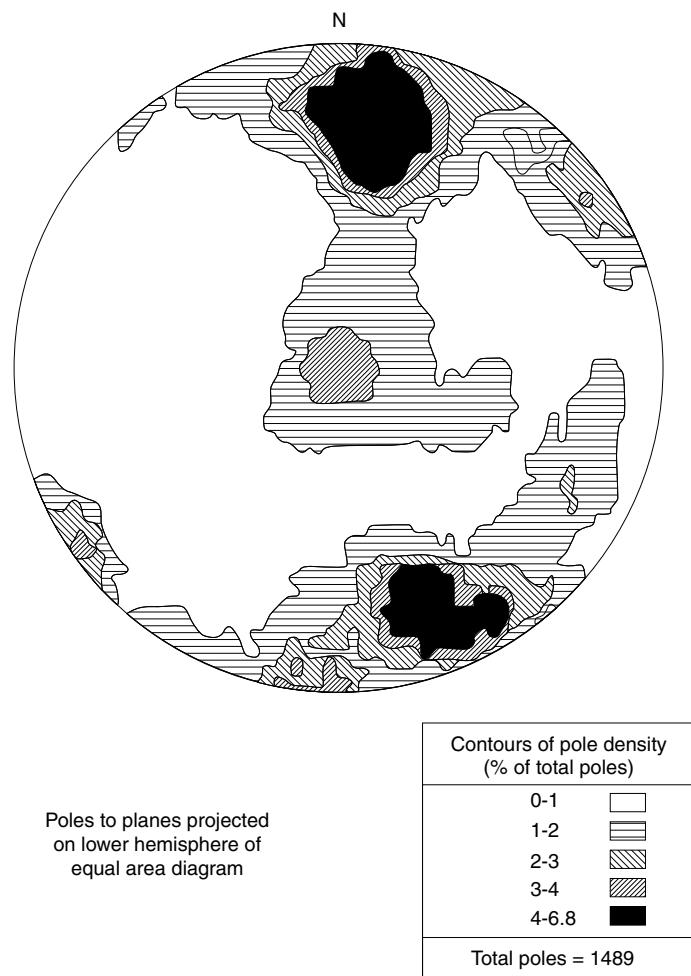


FIGURE 20.10 Equal area stereonet of fracture orientations plotted as poles to planes. (From Raven, K.G., 1986. Hydraulic characterization of a small ground-water flow system in fractured monzonitic gneiss, *Nat. Hyd. Res. Inst. Scientific Series No. 149*, No. 30, 133pp.)

of infilling (normally, calcite or sulfate mineralization), and (3) in crystalline or deformed sedimentary rock, evidence of shear or slickenside structure. Evidence of weathering or iron staining is indicative of groundwater flow. Examination of the core in detail for lithological changes is also important because discontinuities at geological contacts are frequently water-bearing fractures. In a detailed comparison between core analysis and subsequently-obtained hydraulic testing results (West et al., 2005), a ranking system ranging from 1 to 10 based on the likelihood of a specific core break being open to flow was shown to have a good correlation to the hydraulic results.

It is important to note that the aperture or width of naturally open fractures cannot be determined from physical measurements on the core. This is because removing core from the subsurface disrupts the natural stress conditions such that aperture values become arbitrary.

If the orientation of bedding is known or the core is oriented using an orientation device, it is a relatively simple procedure to obtain the orientation of non-horizontal fractures by measuring the arc length of the fracture ellipse in the core (Lau, 1983). Orientation data may be displayed using the same methods as used for surface fractures. Fracture spacing from the core can be determined as the spacing between naturally open fractures (eliminating all other fractures). Further discussion of the analysis of core, particularly in sedimentary rock is provided by Kulander et al. (1990).

20.4.3 Borehole Geophysical Methods

There are a broad number of borehole geophysical tools available for the investigation of hydrogeological features in fractured rock aquifers. Many such as the gamma, neutron, and electric logs are most helpful with lithological interpretation and of limited application in identifying fracture features relevant to groundwater flow. Identification of fractures and other hydraulic features is best determined using calliper, temperature, conductivity, borehole flowmeter, and borehole image logs.

Borehole acoustic televiewer, is one of the most useful borehole geophysical tools for determining the location and orientation of fractures in an open borehole (Zemanek et al., 1969). As the televiewer probe travels the borehole length, a rotating electronic transducer emits a pulsed ultrasonic beam. The reflection of this acoustic energy from the borehole wall produces an image that can be used to identify the strike, dip, and relative size of the fractures intersecting the borehole. The system characterizes fractures only in the immediate vicinity of the borehole. Additionally, it is often difficult to distinguish open fractures from the effects of drilling damage. Recent developments in acoustic probe hardware, optical televiewer probes , and the software for the interpretation of the acoustic or optical signal (Deltombe and Schepers, 2004) have made for more automated and accurate assessment of the presence of fractures in a variety of borehole conditions. Many geophysical logging companies in North America and abroad have the capability to conduct televiewer logs.

A borehole video camera (BVC) survey is a cost-effective alternative to acoustic or optical televiewers, particularly in shallow boreholes (<100 m depth). BVC's film the borehole walls and produce either a color or black and white image of the borehole. Many of the cameras designed to inspect sewers or pipelines are easy to adapt to a borehole application. Automated depth counters allow for the location of specific features such as fractures, vugs, precipitates, mineralization, changes in lithology, or the presence of flowing fractures. Mirrors can be attached to permit a side view of the borehole walls. A BVC survey conducted prior to hydraulic testing or the installation of permanent packer strings will reduce the chances of packer damage or zone leakage due to the inflation of packers over an enlarged portion of the borehole.

Recently, borehole temperature probes and borehole flowmeters have been used to detect the presence of large, open fractures using either ambient flow in the borehole (Paillet and Kapucu, 1989), or pumping-induced vertical flow (Hess and Paillet, 1990; Paillet, 2004). In both cases, these methods are sensitive only to the largest one or two fractures, although quantitative information about the hydraulic properties of these features can be obtained with manipulation and correction of the data (Paillet, 2004). Other geophysical logging tools suitable for fracture detection in open boreholes include (1) multi-arm caliper, (2) formation microscanner, and (3) azimuthal laterlog. Conventional and imaging logs are compared for their ability to detect fractures in Jouanna (1993).

20.4.4 Hydraulic Testing Methods

The characterization of groundwater flow systems in fractured rock is generally dependent on measurements of the permeability of the rock (hydraulic testing). The principal aim of hydraulic testing is to determine the permeability of the fractures and of the rock mass to assist in the estimate of groundwater velocity through the fractures and to appraise groundwater resources. The key to successful field characterization of fractured rock systems lies in the design of hydraulic testing programs. This design should be developed from a preliminary conceptual model of the fracture framework and groundwater flow system. The choice of hydraulic test will depend on several factors including the expected permeability, fracture frequency, matrix properties, and type of rock, depth of the borehole, inter-borehole spacing, and purpose of investigation. In this section, constant-head injection tests, and pulse interference testing will be described. In addition, some of the concepts important in the interpretation of pumping tests (i.e., wellbore storage and skin effects) will be introduced. A general description of various testing methods applicable to hydrogeological investigations in fractured rock can be found in Doe et al. (1987), and a comprehensive case study is provided by Novakowski et al. (1999).

20.4.4.1 Hydraulic Testing Equipment

Essential equipment for testing the hydraulic properties of open boreholes in rock, are packers of some type and devices to measure pressure (i.e., pressure transducers). Packers are used to temporarily seal portions of an open borehole to allow for testing in specific zones. In some instances, hydraulic testing is carried out after the installation of permanent borehole instrumentation (discussed later in this section). Typically, for studies conducted at shallow depth (i.e., <100 m) in open boreholes, test zones may range from 2 to 10 m in length with packer seal lengths ranging from 1 to 2 m. In shallow systems, packers are normally inflated with compressed air or nitrogen. In deeper systems or in low-permeability rock, water can be used to provide a stiffer seal. Commercially available packers generally consist of a gland of rubber, steel-reinforced rubber or neoprene over steel or PVC pipe with either a fixed-head or sliding-head design. Fixed-head packers rely on the stretch of the gland material during inflation to provide the seal. Conversely, with sliding-head packers, o-rings in the bottom head of the packer permit the gland to slide upward during inflation. This latter design allows for the use of stiffer (i.e., reinforced) materials to inflate against the borehole wall. The fixed-head design is adequate for shallow rock systems with competent borehole walls, while sliding-head packers are more suitable for use in deeper flow systems.

A two-packer system having a single isolated zone is usually adequate for hydraulic testing programs conducted in moderate to sparsely fractured rock. For conditions where fracture frequency is high and packer leakage or short-circuiting effects around the main packers are anticipated, additional packers may be installed above and below the test section to monitor these effects.

For multiwell testing, multiple-packer strings are normally used in the observation boreholes. In some cases, it may be convenient and cost-effective to install permanent packer strings in the observation boreholes (described later in this section).

The pressure response in the isolated source zone or in the observation zones to perturbations generated by injecting or withdrawing water from the source zone is usually measured using pressure transducers or transmitters. The most commonly used types of pressure transducers are based on either strain gauge or vibrating wire membranes. Descriptions of transducer electronics and other geotechnical instrumentation useful to hydrogeological studies can be found in Dunncliff (1988).

20.4.5 Single-Well Hydraulic Tests

20.4.5.1 Constant-Head Injection Tests

As open fractures may intersect a borehole at any depth, a continuous measurement of permeability with depth is necessary to locate these features. The most widely used method for testing individual boreholes is the constant-head injection technique. The principle behind a constant-head injection test (CHIT) is to inject or withdraw water at a constant hydraulic head into an isolated portion of the borehole and

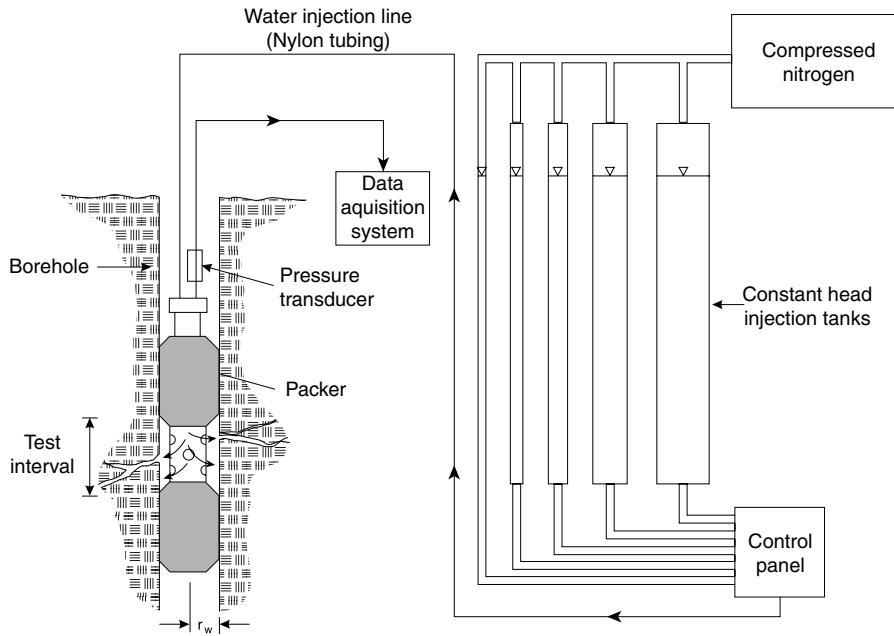


FIGURE 20.11 Schematic diagram of a typical constant-head injection testing system. A series of five tanks is shown on the right while the left shows the downhole instrumentation consisting of a two-packer system and pressure transducer.

measure the flow rate at steady-state conditions (Ziegler, 1976). CHITs may also be referred to as “Lugeon tests,” or “constant pressure tests.” While a pump and water supply could be used to inject water into the test zone, a series of different diameter tanks fitted with manometers is more commonly used and extends the limits of the testing apparatus to generate a wide range of injection flow rates (Figure 20.11). A range of 10^{-10} to $10^{-3} \text{ m}^2/\text{sec}$ in transmissivity, T , can typically be determined using a series of three or more tanks ranging in diameter from 0.01 to 0.3 m. Flow rate is measured by timing the change in water level as viewed through the tank manometers. Compressed nitrogen gas added above the water in each tank ensures a constant pressure as the tank empties. Flow meters can be used to measure the flow rate; however, most flow meters have a very limited measurement range and thus a complex series of flow meters connected by valves would be required to achieve a range in measurement of T similar to the tank system. Thus, the great advantage to this testing method is the broad range of measurement.

Using measurements of flow rate, Q , and the change in pressure expressed as a change in hydraulic head, ΔH , at steady conditions, the T of the tested zone is calculated using the Theim equation:

$$T = \frac{Q}{\Delta H 2\pi} \ln \left(\frac{r_e}{r_w} \right) \quad (20.17)$$

The radius of influence, r_e , or outer flow boundary, can generally be assumed to be between 10 and 15 m for moderate values of T (Bliss and Rushton 1984). To improve the estimate of r_e , the following simple expression can be used by assuming a value for storativity, S :

$$r_e = 2\sqrt{T/S \cdot t} \quad (20.18)$$

where t is the duration of the injection period for the test. The value of S can be estimated or determined from a slug or pumping test. The initial value of r_e is calculated using an estimate of T and then revised using the result from Equation 20.17. Following a few iterations, r_e and T stabilize and a more accurate

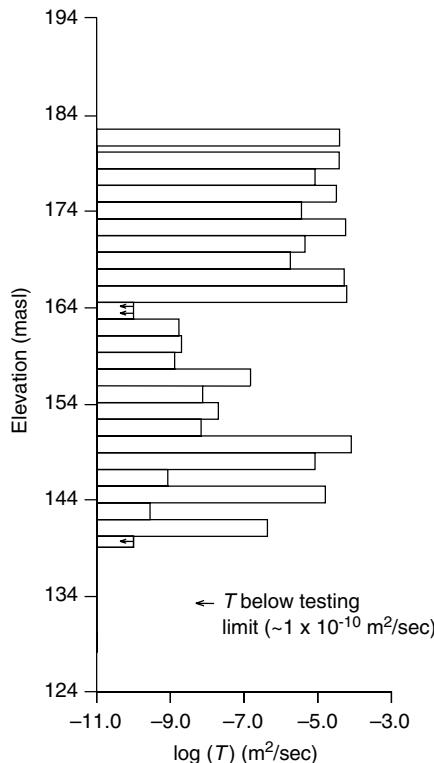


FIGURE 20.12 Typical profile of T with respect to depth obtained from a borehole drilled in Silurian dolostone and tested with constant-head injection tests using a 2-m packer test interval.

estimate of T is obtained than when using Equation 20.17 alone (Novakowski, 2005). Further discussion of the r_a issue can be found in Doe and Remer (1980).

For tests in which it is assumed that only one fracture intersects the test section, an equivalent single fracture aperture, $2b_{eq}$, can be determined from the test results by using the cubic law:

$$2b_{eq} = \left[T \cdot \frac{12\mu}{\rho g} \right]^{1/3} \quad (20.19)$$

The vertical distribution of T or $2b_{eq}$ is determined by systematically testing the length of the borehole in sections, using a two-packer system. If possible, two to three increasing increments of pressure should be used during each test. In most studies, injection pressures leading to $\Delta H \leq 10$ m are recommended. Care must be taken not to use pressures that generate fracture dilation or hydro-fracturing. Figure 20.12 shows a typical profile of T with respect to depth obtained from a borehole drilled in Silurian dolostone. T varies by over seven orders of magnitude over 50 m of elevation illustrating the heterogeneity commonly observed in fractured rock.

It is important to note that a single permeable fracture within a test zone will predominate the permeability measurement. Thus, the choice of test zone length will depend on the expected fracture spacing with depth. A test interval too short in length will lead to an unnecessarily large number of tests while a test interval too long in length will not capture any of the variability in T and lead to general overestimates of the permeability of the rock mass. To properly characterize a borehole in detail, an initial survey using a large packer separation length will identify the more permeable zones while a subsequent second survey using a shorter packer separation length can be used to characterize the properties of individual permeable features and fracture zones (Figure 20.13).

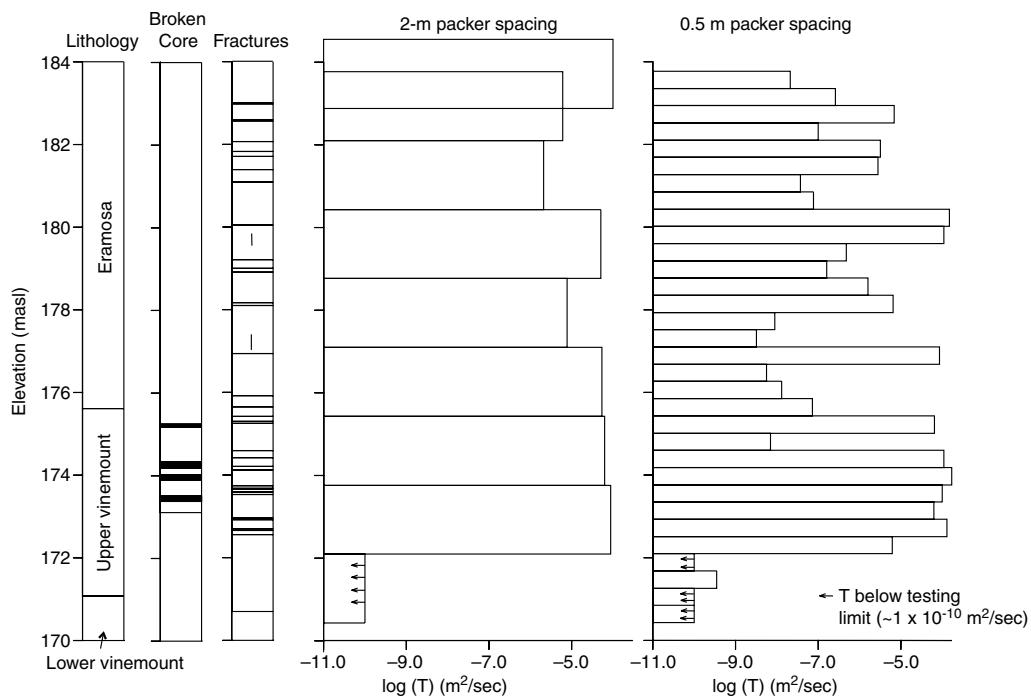


FIGURE 20.13 Results of two testing surveys (2-m and 0.5 m packer intervals) compared to fractures and broken core zones identified in core examination.

20.4.5.2 Slug Tests

The slug (pulse or bail) test is a transient single-well method that is also frequently used in hydrogeological studies of fractured rock to obtain estimates of the hydraulic properties of a given length of borehole. Type curves based on radial flow in a confined porous aquifer are usually employed to obtain estimates of T and S (Cooper et al., 1967). Although, in principle, this method should provide estimates of T that are more accurate than CHITs, slug tests conducted in fractured rock, frequently exhibit results that deviate from the ideal (Cooper et al. 1967) response. This has led to the development of numerous conceptual models that consider the effects of alternate boundary geometries, composite zones, and the other physical conditions more typically found in fractured rock systems. For example, in the presence of a damaged zone around the well, known as wellbore skin, T and S determined from a slug test may reflect only the skin zone (e.g., Faust and Mercer, 1984; Sageev, 1986). Wellbore skin can refer to a zone of enhanced permeability (negative skin) or of reduced permeability (positive skin). Double porosity effects (e.g., Doughtery and Babu, 1984), where both the fractures and matrix contribute to the flow system can also be incorporated. Composite models are used to account for separate regions surrounding the well, which have properties that are individually different from the formation properties as a whole (e.g., Moench and Hsieh, 1985; Karasaki et al., 1988). Open-hole slug tests are generally limited to zones of $T \geq 10^{-7} \text{ m}^2/\text{sec}$ due to the time required to achieve sufficient recovery for interpretation. In lower permeability material ($T < 10^{-10} \text{ m}^2/\text{sec}$) slug tests are conducted under shut-in (no free-water surface) conditions, reducing the effects of storage in the borehole (Bredehoeft and Papadopoulos, 1980; Neuzil, 1982).

20.4.6 Multiwell Hydraulic Tests

20.4.6.1 Interference Tests

In the following, the term “interference testing” is used to refer to all types of multiwell hydraulic tests. Interference testing between specific zones in two or more boreholes serves two main objectives

(1) establish the presence of hydraulic connection between the zones and (2) obtain estimates of interwell T and S . In the case of isolated discrete fractures, interwell aperture can also be calculated. In hydrogeological investigations, the choice of interference test method will be influenced by the diameter of the borehole, ability or desire to pump, or inject large quantities of water into the formation, and the presence of a water table. Pumping tests conducted in fractured rock are done in a similar manner to those conducted in porous media. The difference lies in the factors influencing the response to pumping. Similar to the case for slug tests, pumping interference tests are influenced by (1) wellbore storage, (2) skin effects, (3) double porosity, (4) boundary effects (e.g., a linear feature), and (5) vertical or near-vertical fractures. Interpretation of interference tests is accomplished by matching the pressure or hydraulic head response in both the source and observation zones to theoretically derived response curves. This is accomplished by either manually matching field data to generate theoretical type curves (e.g., Earlougher, 1977) or using automated algorithms (e.g., Piggott et al., 1996).

20.4.6.2 Wellbore Storage

Fracture apertures usually range from a few microns to a few millimeters. Consequently, the volume of water stored in an open borehole will be several orders of magnitude greater than the volume of water found in the fractures intersecting the borehole. This will influence the pressure response measured in the well during pulse interference and pumping tests. The wellbore storage factor, C_s , for an open well, is defined as πr_c^2 where r_c is the radius of the casing. The dimensionless wellbore storage coefficient, C_D , is defined as

$$C_D = \frac{C_s}{2\pi r_w S} \quad (20.20)$$

In low storativity media (most fractured rock systems), wellbore storage will significantly influence the early time response to a pressure disturbance in either a source or observation well. To accurately measure the hydraulic properties of low storativity media, wellbore storage must either be eliminated (by eliminating the free-water surface in the testing equipment) or accounted for in the interpretation of transient multiwell tests.

20.4.6.3 Pulse Interference Tests

Single-pulse interference tests are an easy yet powerful extension to single-well slug tests. In this case, the pressure response in one or more observation zones due to the introduction of an instantaneous slug of water in a nearby source well is measured (Figure 20.14). A pulse interference test can be analyzed quickly using a graphical method (Novakowski, 1989). To do so, the magnitude of peak response and time lag in the observation zone must be measured. Peak response is defined as the ratio of the maximum rise in hydraulic head at the observation well (Δh), divided by the initial rise in hydraulic head in the source well (H_0). Time lag (t_L) is the elapsed time between the start of the slug test in the source well and the maximum peak response in the observation well. The remaining required dimensionless parameters (r_D , t_D , and dimensionless wellbore storage coefficients in both the observation and source well (C_{DO} , C_{DS}) are defined as follows:

$$r_D = \frac{r}{r_w} \quad C_{DS} = \frac{C_s}{2\pi r_w^2 S} \quad C_{DO} = \frac{C_o}{2\pi r_w^2 S} \quad t_D = \frac{Tt_L}{r_w^2 S} \quad \Delta h_{DO} = \frac{\Delta h}{H_0} \quad (20.21)$$

where r is the distance between the source and observation boreholes and C_s , and C_o are the storage coefficients in the source and observation boreholes, respectively. One of two sets of graphs is used to determine T and S via the dimensionless variables. Set 1 (Figure 20.15) is used when there is no observation well storage ($C_{DO} = 0$) whereas Set 2 (Figure 20.16) is used when the storage capacity in the source well is equal to the storage capacity in the observation well ($C_{DO} = C_{DS}$). First, using h_{DO} and r_D , a value for C_{DS} is determined from the graphs in part (a) and S is calculated from the definition

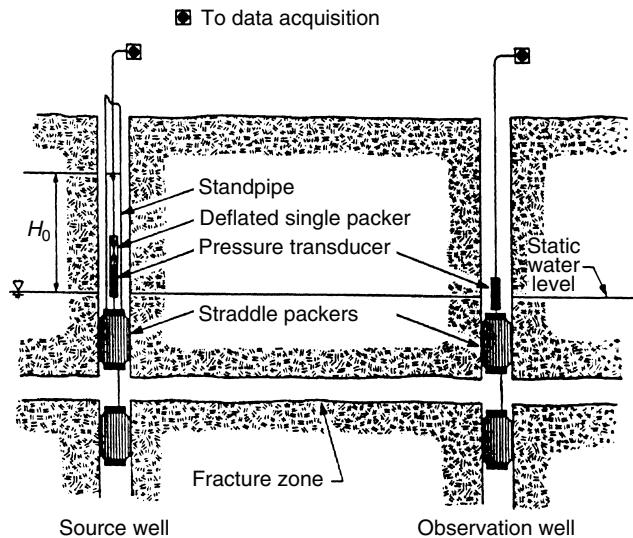


FIGURE 20.14 Schematic diagram showing field setup for pulse interference tests. (From Novakowski, K.S., 1989. *Water Resour. Res.*, 25(11): 2377–2387.)

of C_{DS} in Equation 20.21. Second, using h_{DO} and r_D again, a value for t_D is determined from the graphs in part (b) and T is calculated from the definition of t_D in Equation 20.21. Figure 20.17 shows an example of both source and observation zone responses to a typical pulse interference test where $r = 22$ m, $T = 5 \times 10^{-4}$ m²/sec and $S = 4 \times 10^{-5}$. This method was recently employed in the development of a detailed conceptual model for groundwater flow pathways in a structurally complex folded limestone (Stephenson et al., 2006)

20.4.7 Point Dilution Method

Measuring groundwater velocity in discrete fractures is often difficult due to the large uncertainty associated with estimates of low hydraulic gradient (i.e., hydraulic head varies by only centimeters over distances of many meters). The point dilution method may be employed to determine direct measurements of groundwater velocity. This method is based on the decay in concentration with time of a mixed tracer in a single well due to dilution caused by the natural groundwater flow through the borehole (Drost et al., 1968; Grisak et al., 1977). To conduct a point dilution experiment, a small section of the borehole having one or more fractures is isolated using a set of two pneumatic packers. To minimize the duration of the experiment, the mixing volume must be reduced by using a short spacing between the two packers. The experiment is initiated by instantaneous injection of a small volume of conservative tracer into the test zone and mixing is continued throughout the duration of the experiment. Figure 20.18 shows schematically a typical arrangement for a point dilution experiment. Suitable conservative tracers include bromide or chloride, some fluorescent dyes (e.g., Lissamine FF), radioactive isotopes (e.g., Tritium), or stable isotopes (e.g., Deuterium).

The decay in concentration is interpreted using (Drost et al., 1968):

$$\frac{dc}{dt} = -Av_a \frac{c}{V} \quad (20.22)$$

where A is the cross-sectional area available to flow, v_a is the apparent velocity of groundwater flowing through the wellbore and V is the volume of the sealed off portion of the borehole in which the dilution occurs. The solution to Equation 20.22 for the case of instantaneous injection and relating the apparent

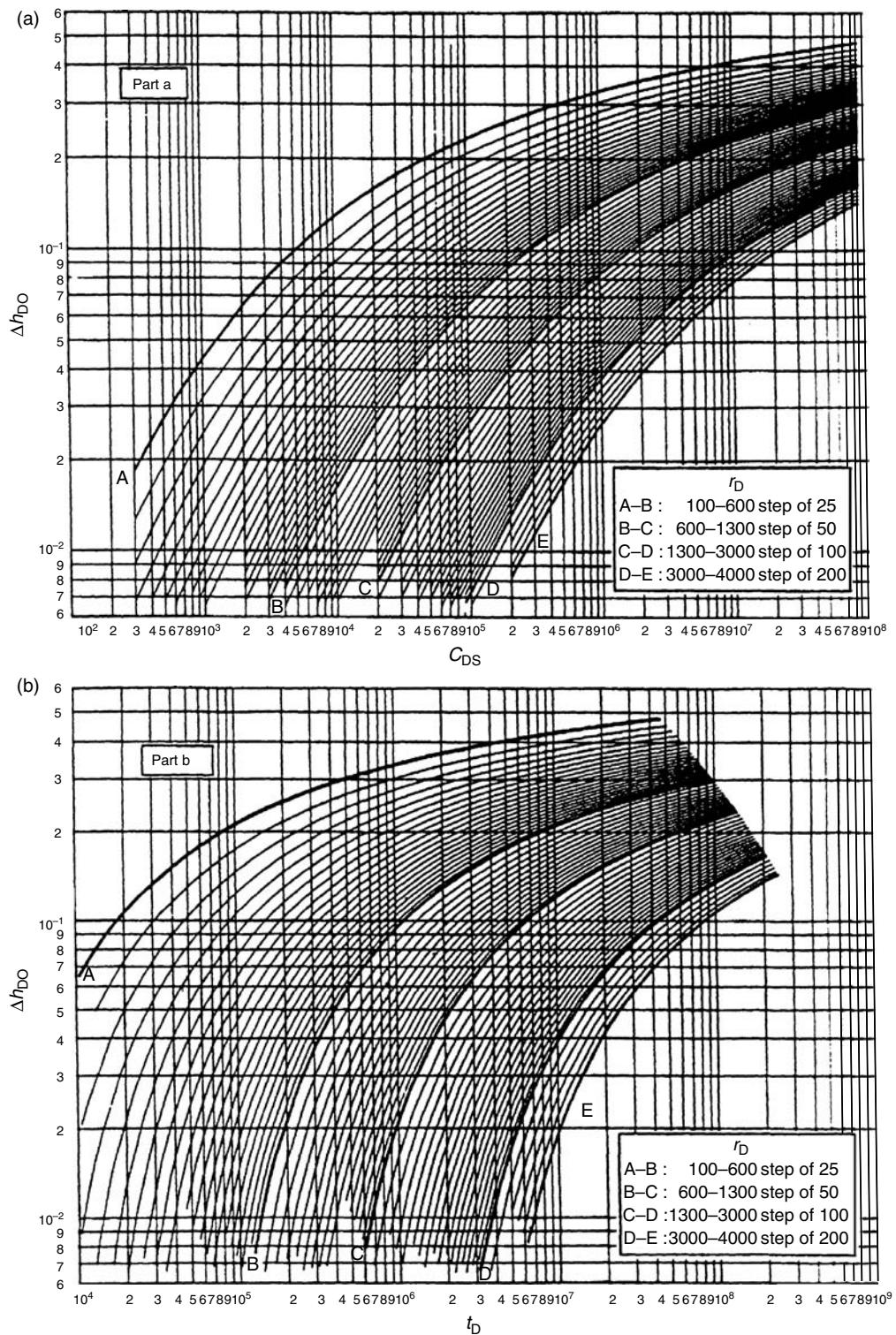


FIGURE 20.15 Type curves used to interpret pulse interference tests by graphical method when there is no observation well storage (a) curves used to calculate S and (b) curves used to calculate T . (From Novakowski, K.S., 1989. *Water Resour. Res.*, 25(11): 2377–2387.)

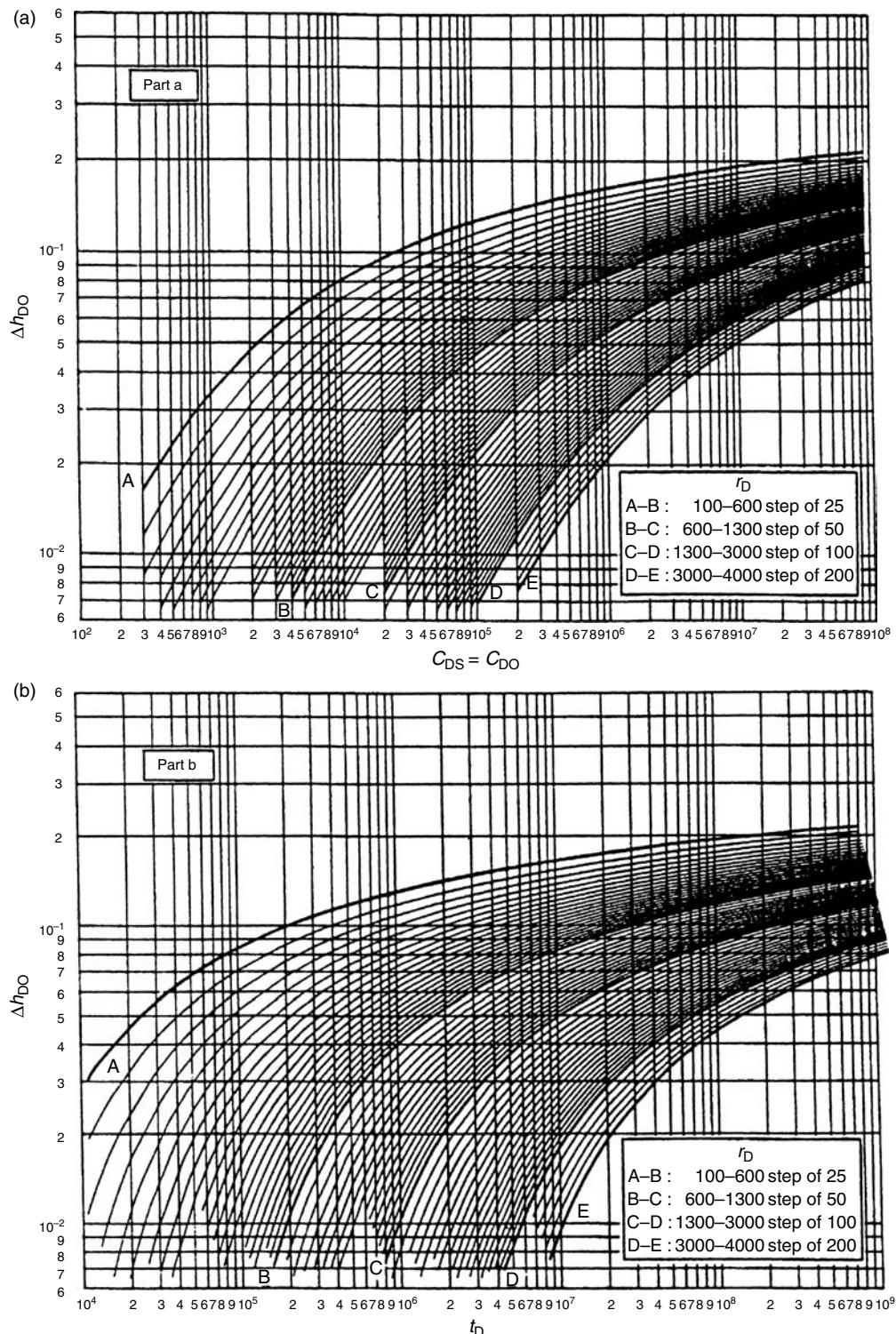


FIGURE 20.16 Type curves used to interpret pulse interference tests by graphical method when observation well storage = source well storage (a) curves used to calculate S and (b) curves used to calculate T . (From Novakowski, K.S., 1989. *Water Resour. Res.*, 25(11): 2377–2387.)

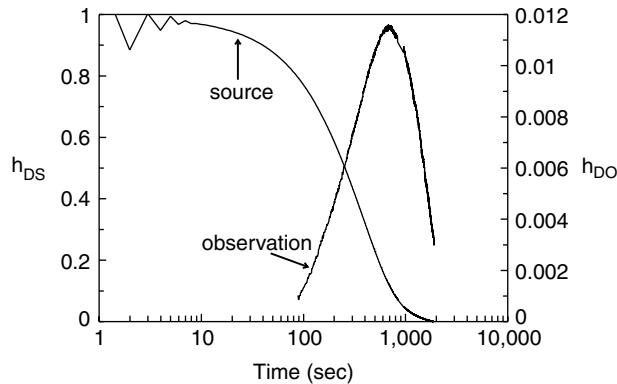


FIGURE 20.17 Example of source well pulse and observation zone response. Test was conducted in a 500- μm fracture where the observation zone was approximately 22 m from the source zone.

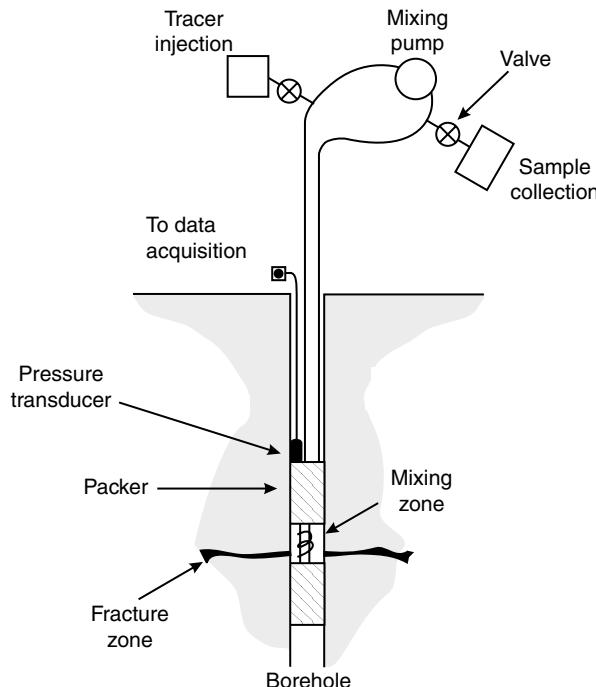


FIGURE 20.18 Schematic diagram of a point dilution experimental setup.

velocity in the test section to the true formation velocity, v_f , yields (Drost et al., 1968):

$$v_f = \frac{V}{\xi At} \ln \frac{c}{c_0} \quad (20.23)$$

where c_0 is the initial concentration at $t=0$, c is the concentration at time t after the tracer was injected, and ξ is a dimensionless correction factor accounting for additional flow captured by the open well due to the convergence of flowlines in the neighborhood of the wellbore. Groundwater velocity in a single fracture intersecting the test zone (v_f) is determined from the measurements of dilution of tracer in the sealed off portion of the borehole by plotting the results in the form of $\ln c$ vs. t and fitting a linear regression

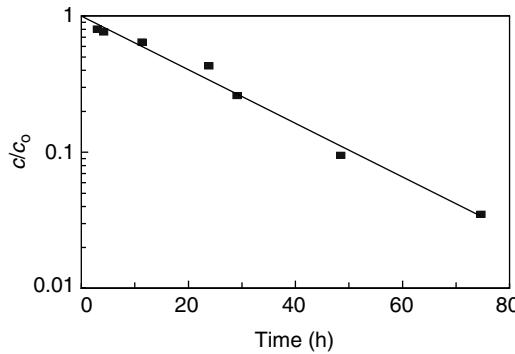


FIGURE 20.19 Example point dilution response curve. In this test a velocity of 4 m/day was measured in a 220 μm fracture.

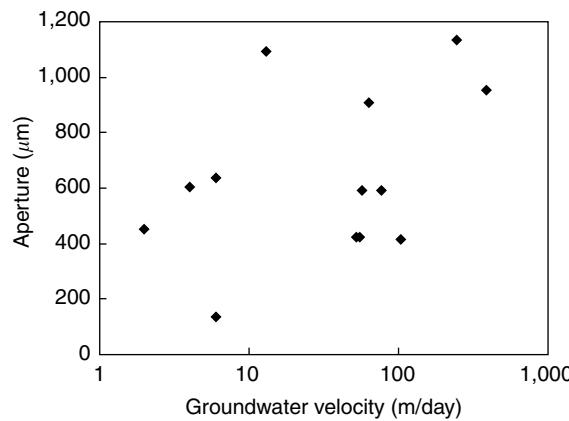


FIGURE 20.20 Average groundwater velocity vs. fracture aperture for point dilution experiments conducted in an orthogonal fracture network.

line to the data (Figure 20.19). Using the time (t_m) at $c/c_0=0.5$ from the regression line, Equation 20.23 reduces to:

$$v_f = \frac{0.693V}{\xi At_m} \quad (20.24)$$

In the case of a zone containing a single fracture, A equals the diameter of the borehole times the hydraulic aperture and $\xi=2$. Where more than one fracture intersects the borehole, the interpretation is more complicated. To accurately estimate velocity, the aperture of each individual fracture must be known, otherwise the result will represent an average.

In a detailed study of groundwater velocity conducted in a fracture network that pervades a flat-lying Silurian dolostone in southern Ontario (Novakowski et al., 2006), a range in velocity from 3 to 388 m/day was observed. The aperture of the fractures ranged from approximately 130 to 1100 μm . Figure 20.20 shows the relation between aperture size and groundwater velocity. Although it might be anticipated that velocity would be proportional to aperture, no such relation is evident. This is attributed to the nature of the fracture interconnections at this site where some smaller aperture fractures act as a limit to the overall groundwater flux through the domain and thus experience large hydraulic gradients and high velocities.

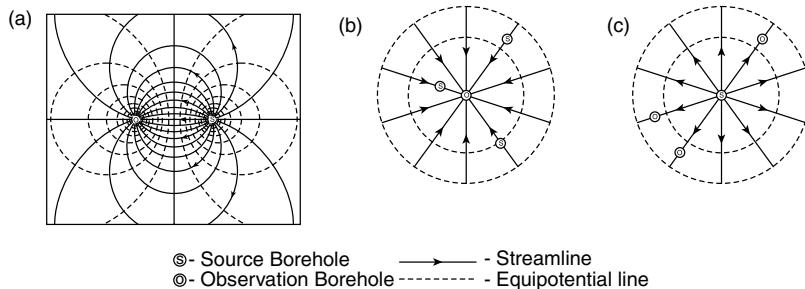


FIGURE 20.21 Flow field geometry commonly used for tracer experiments: (a) injection-withdrawal, (b) radial convergent, and (c) radial divergent.

20.4.8 Tracer Experiments

Tracer experiments conducted in fractured rock provide a means of determining solute transport parameters (i.e., velocity, dispersivity, and matrix porosity) of individual fractures or fracture zones at the field scale. Similar to tests in porous aquifers, tracer experiments conducted in fractures can be carried out using either forced hydraulic gradient or natural gradient flow conditions. Because of the difficulty and cost associated with natural gradient tracer experiments, very few of these have been conducted and no further discussion of this method will be presented here. In the following sections, the common methodologies used to carry out multiwell tracer experiments under conditions of forced gradient are outlined.

The three flow field geometries that are commonly used when conducting tracer experiments under forced gradient in fractured rock are (1) injection-withdrawal, (2) radial convergent, and (3) radial divergent. In the injection-withdrawal format, an artificial steady-state flow field between two wells is established where the injection flow rate is equal to the withdrawal flow rate (Figure 20.21a). Tracer is introduced either as a finite slug or continuously into the injection well and the concentration is monitored at the withdrawal well. An example of this type of experiment can be found in Novakowski et al. (2004). The water removed from the withdrawal well can be re-circulated back into the system via the injection well, but this will complicate the interpretation of the test. The advantage to this test method is the development of a well-controlled flow field where recovery of the tracer should approach 100%. A radial-convergent experiment is conducted by passively injecting tracer into one borehole and withdrawing by pumping at a second borehole (Figure 20.21b). Multiple boreholes can be used by introducing a different tracer in each borehole. An example of the use of this experimental method is outlined in Shapiro and Nicholas (1989). Radial-divergent experiments are conducted by injection of tracer in a single well and passively monitoring the tracer arrival in one or more observation wells (Figure 20.21c). An example of this experiment is presented in Novakowski and Lapcevic (1994). The choice of methodology will be dependent on the objectives of the experiment, the number of boreholes available for study, the ease of pumping and sampling, and the type of tracer used.

To date most forced-gradient experiments have been conducted in discrete fractures or fracture sets in which the geometry of the fracture is well known (e.g., Raven et al., 1988; Shapiro and Nicholas, 1989; Abelin et al., 1991; McKenna et al., 2001; Mazurek et al., 2003). The scale of these experiments has been limited to interwell travel distances of 50 m or less. In future, tracer experiment methodology should prove very useful in confirming the geometry of fracture intersections over larger scales, the presence of which may be surmised from the results of interwell hydraulic tests. Interpreting these experiments is limited at present due to computational requirements for managing the matrix diffusion process (more discussion of this in the modeling section).

20.4.9 Borehole Instrumentation

Permanent completion of boreholes drilled in bedrock is generally carried out in one of two ways (1) using technology designed for piezometer construction in unconsolidated porous media or (2) through the use

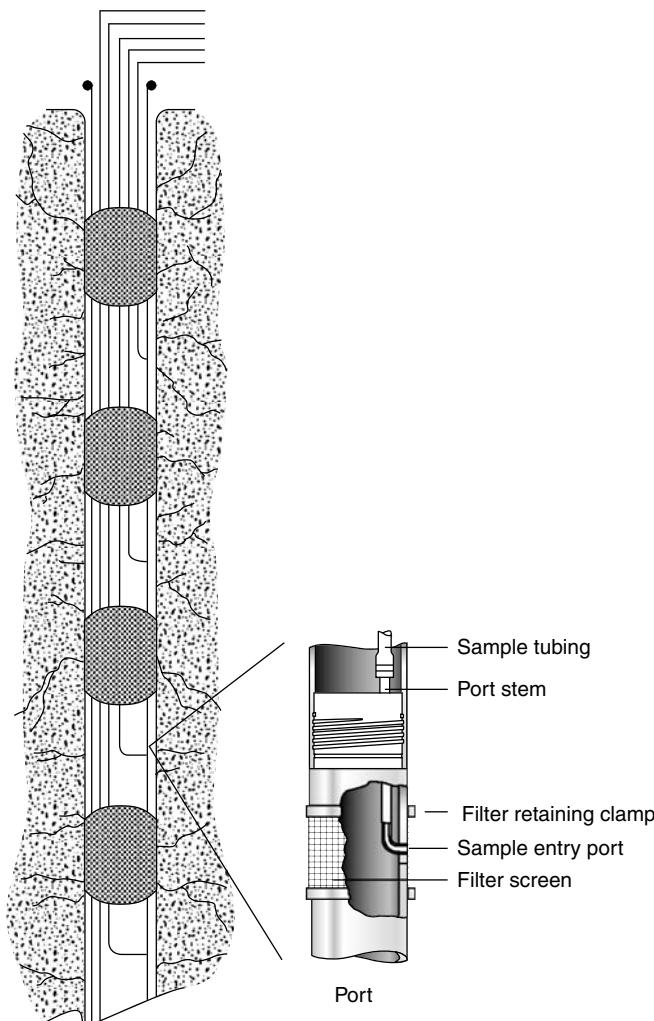


FIGURE 20.22 Components of the Solinst® system of permanent multilevel instrumentation for use in the rock boreholes. Access lines may be electrical cables for pressure transducers installed in the ports or tubing to sample the isolated zones. Pumps (triple tube or double valve types) may be dedicated to specific zones or lowered into the sampling tubing as required. Manual water level measurements are obtained through the tubing.

of multilevel borehole casing or liners. Only the latter will be discussed in this chapter as the former is covered in discussions of well completions for unconsolidated porous media. Multilevel completions are primarily designed to (1) obtain measurements of hydraulic head and (2) obtain representative samples of groundwater. This is accomplished by isolating sections of the borehole using a series of permanent packers joined by casing or riser pipe or by the installation of a liner material that directly contacts the wall of the borehole. Access to the isolated zones is through manometers or access ports. Currently, two commercial systems for multilevel casing completions are widely used around the world. These are the Solinst® system (Cherry and Johnson, 1982) and the Westbay® system (Black et al., 1987). The FLUTE® liner system is a relatively new product that provides similar advantages as the more traditional multilevel casing with the added flexibility of temporary completions.

The Solinst® system (Figure 20.22) is based on packer elements that are filled with a sealant chemical that is activated with the introduction of water into the casing after the complete casing string has been lowered into the borehole. Sampling ports and manometers of thin nylon or teflon tubing provide access

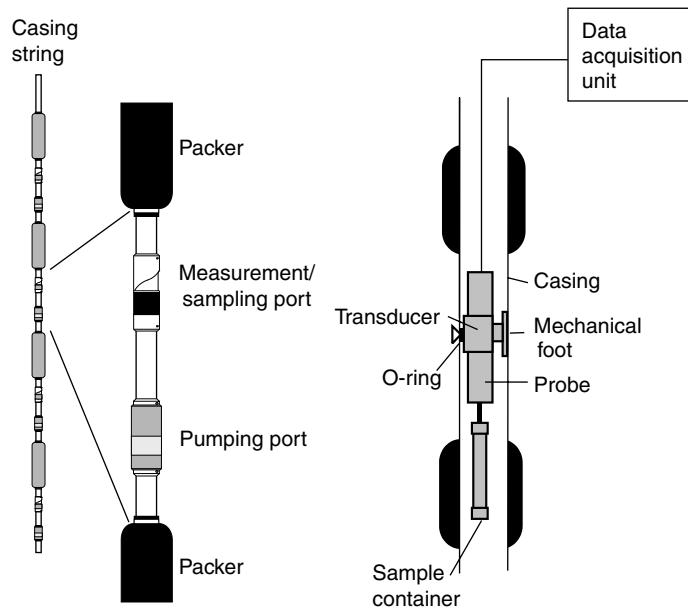


FIGURE 20.23 Components of the Westbay® system of permanent multilevel instrumentation for use in the rock boreholes. Note that no tubing or electrical cable is permanently installed in the casing string as the probe is lowered to each zone to either measure water pressure or obtain a groundwater sample.

to each sealed off zone to measure water levels and withdraw groundwater samples. Triple-tube sampling pumps can either be installed in each zone or lowered into the manometers. These pumps allow sampling in zones where pumping from the surface is not feasible. Additionally, pressure transducers or transmitters can be included in each zone to electronically measure water pressure. Depending on the objectives of the study and specific site conditions, the modular instrumentation system can be customized to maximize the amount of data obtained. The number of zones sampled or measured is limited by the diameter of the borehole (and thus the diameter of the casing) as each transducer, pump, or sampling port requires access to the surface through tubing or electrical cable.

The Westbay® System is also a modular design having water-filled packers connected by casing elements and specially designed pumping and measurement/sampling ports (Figure 20.23). Water pressure is measured and representative groundwater samples are obtained using a submersible probe that is lowered into the casing and connected to an electronic data acquisition device on the surface. The probe has a small arm that is used to locate each measurement port inside the casing. Once positioned on the port, a mechanical foot on one side of the probe is activated and presses against the inner casing wall. This causes an o-ring on the opposite side of the probe to seal around a ball bearing and exposes the pressure transducer in the probe to water pressure outside the casing. Groundwater samples can be obtained using the same port by opening a valve from the surface unit and allowing water to fill a sample container attached below the probe (Figure 20.23). Multiple probes can be installed temporarily in individual casing strings allowing for simultaneous sampling or water level observation.

The FLUTe® liner system was developed initially to provide a temporary seal in open boreholes in bedrock (i.e., to prevent cross contamination between fractures via the borehole). The liner is constructed of a flexible polyurethane-coated nylon fabric and is inverted into the borehole from a reel. The liner is inflated by providing an interior water head greater than that in the fractures and borehole. Sampling is conducted through interior tubing that access a port in the wall of the liner. The port is located behind a geotextile that provides a fluid pathway from the fracture to the sample tubing.

Multilevel completions provide an excellent measurement platform for interwell hydraulic tests. For example, a borehole instrumented with several zones, used as an observation well during a pumping test

can provide information on the vertical connection of fracture zones. To maximize the data obtained from permanently instrumented boreholes, care must be taken in designing appropriate isolated zones. Careful examination of rock core, geophysical surveys, and hydraulic testing results prior to installation of the instrumentation will ensure that the isolated zones appropriately define the three-dimensional nature of the groundwater flow system.

20.5 Conceptual Models

In formulating an understanding of groundwater flow and contaminant transport for a given site or region at which flow is predominated by the presence of fractures, it is important to develop a reasonable conceptual model as a starting point. In many cases, the conceptual model may incorporate elements of the flow system at a variety of scales. This may include processes at the microscopic scale such as matrix diffusion and processes at the single-continuum scale such as recharge or discharge. In the following section, a variety of conceptual models for flow and transport at the scale of a discrete fracture and a fracture network are discussed. In all cases, it is assumed that the cubic law, hydrodynamic dispersion, and matrix diffusion apply at the microscopic scale.

Bear (1993) defined four operational scales in a fractured medium, at which different conceptual approaches might be applied (Figure 20.24). At the very-near field scale, flow and solute transport are dominated by a single fracture. At the near field, flow and transport are dominated by a few well-defined fractures and interaction between the fractures and the matrix may play a role. At this scale discrete fracture models may be used in which the major fractures are defined deterministically and the minor fractures are specified stochastically. Depending on the type of fractured rock and the nature of the problem, it is likely the near-field scale is of most interest to practising hydrogeologists. At the far-field scale, multiple continua are defined using at least one continuum for the unfractured matrix, possibly one for the minor fractures and another for the major fractures. At the very-far-field scale, a single continuum can be applied. In particular cases, it may be found that a mixture of conceptual approaches is required. For example, in some sedimentary rock environments, sheeting fractures along bedding planes may predominate in one stratigraphic horizon, while in another the fracturing is so frequent as to warrant the use of a continuum or a multi-continuum approach.

20.5.1 Conceptual Models for a Single Fracture

Numerous studies have been conducted in which the roughness of natural fracture walls and the distribution of fracture aperture have been measured (e.g., Brown et al., 1986; Gentier and Billaux, 1989; Piggott, 1990; Konzuk and Kueper, 2004). These studies have been conducted on fracture samples ranging in scale between 0.1 and 1 m. The results indicate that the surfaces of fracture walls can be rough and undulating with numerous contact points between the walls. In general, it is concluded that (1) fracture aperture may fit into one of several statistical distributions, (2) aperture is correlated spatially, and (3) aperture distributions are scale invariant.

The distribution of apertures at the laboratory scale has been observed to follow either a log-normal distribution (Gentier and Billaux, 1989; Hakami, 1995), a Gaussian distribution (Piggott, 1990; Brown, 1995; Hakami, 1995), or a gamma distribution (Tsang, 1984). Additionally, spatial distributions of aperture are often observed to have a large variance, indicating a high level of heterogeneity in the fracture plane (e.g., Hakami, 1995). The roughness of fracture surfaces has also been determined to display fractal properties implying that variable aperture distributions may be scale invariant (e.g., Brown et al., 1986; Wang et al., 1988).

Fractures with similar statistical distributions of aperture may have very different flow properties due to the spatial relation of the aperture variations. Fracture surfaces have been determined to be correlated spatially with the degree of correlation related to the scale of measurement. Brown et al. (1986) determined

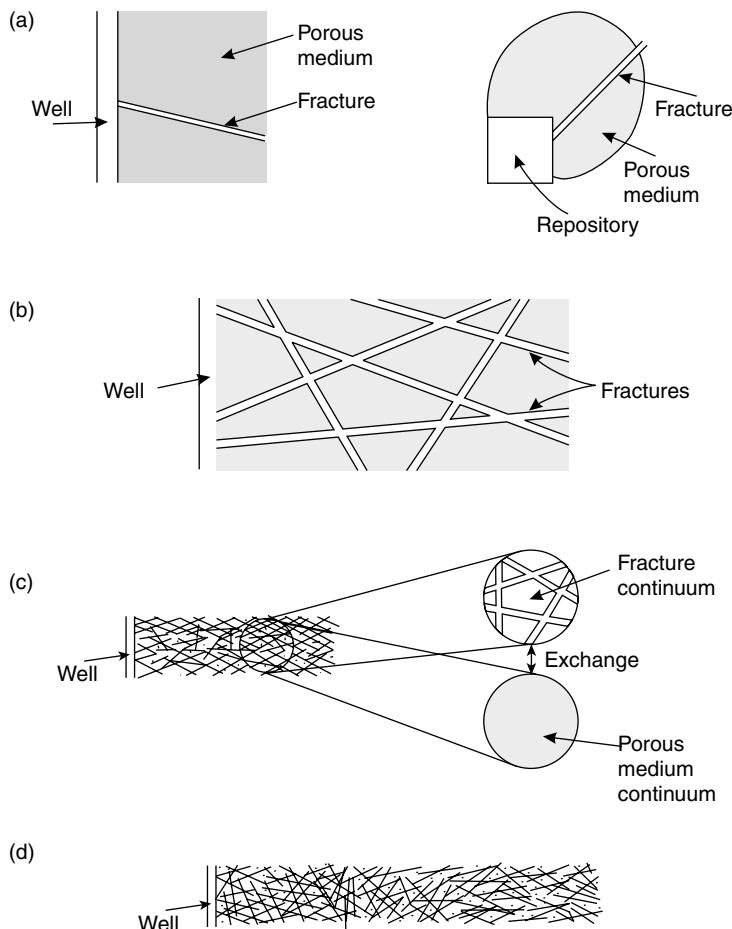


FIGURE 20.24 Operational scales in a fractured medium: (a) very-near field, (b) near field, (c) far field, and (d) very-far field. (From Bear, J., 1993. Modeling Flow and Contaminant Transport in Fractured Rocks. In *Flow and Contaminant Transport in Fractured Rocks*, eds. J. Bear, C.-F. Tsang, and G. de Marsily, pp. 1–37. Academic Press Inc., San Diego, CA.)

correlation lengths from between 0.5 and 5.0 mm for sample lengths and diameters of 25.4 mm. Vickers et al. (1992) studied surface roughness of a single fracture in a block of welded tuff (0.15×0.40 m) and found the resulting aperture distributions close to normal except at the tails. The apertures were also found to be correlated at two scales, one on the order of millimeters and the other on the order of tens of centimeters. In addition, Vickers et al. (1992) found that apertures increased consistently along the entire length of the fracture suggesting that a third spatial correlation occurs at a scale well in excess of the sample dimensions. Hakami (1995) compiled experimental results of aperture measurements at the lab scale and concluded that both the variance and spatial correlation length increase with increasing mean aperture.

It is also well recognized that the regions in which the fracture surfaces are in contact and closed to water flow will strongly influence flow and transport. Experiments conducted by injecting wood's metal into fractures in granitic cores, have shown contact areas ranging from 8% to 15% with a normal stress of 3 MPa (Pyrak-Nolte et al., 1987). At the field scale, hydraulic aperture measurements in two fractures at depths less than 15 m in sedimentary rock (shale/limestone) in an area roughly $30 \text{ m} \times 30 \text{ m}$, indicate fracture closure between 20 and 30% (Lapcevic et al., 1990). Smooth fractures will display many small areas of contact with complex outlines resulting in a complex distribution of fluid flow among many

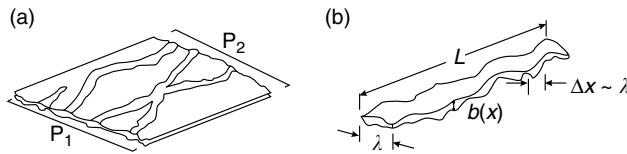


FIGURE 20.25 (a) Channel representation of fluid flow in a single fracture, and (b) schematic sketch of one channel, where $b(x)$ is the aperture distribution, and λ is the spatial correlation length of the distribution. (From Tsang, Y.W., and Tsang, C.F., 1987. *Water Resour. Res.*, 23(3): 467–479).

small channels. In contrast, rough fractures display fewer, larger contact areas with smoother outlines concentrating flow in a few large channels (Odling, 1994). Tsang (1984), observed that increases in contact area results in a reduction in mean aperture, an increase in tortuosity and a decrease in the connectivity of the fluid flow paths. Tortuosity factors determined through comparisons of measured hydraulic versus physical apertures have been determined to be a good approximation to account for closure (e.g., Piggott and Elsworth, 1993).

Experimental studies of flow and transport in laboratory scale fractures have been used to investigate the relationship between measured apertures and resultant flow characteristics within the fracture. Results of flow experiments on fractures in laboratory studies with known aperture distribution show that the ratio between mean aperture and hydraulic aperture is 1.1–1.7, for $0.1 < 2b < 0.5$ mm (Hakami, 1995). The lower value of the hydraulic aperture is to be expected as the variation in aperture of natural fractures forces the flow to be tortuous. It is well recognized that fluid flow in natural fractures will be controlled by the distribution of asperities leading to tortuous paths and areas in which no flow occurs.

There are several conceptual models for transport in a fracture of variable aperture. For example, the “channel model” for flow and transport through a single fracture is based on a series of one-dimensional channels of constant aperture oriented in the direction of flow (Tsang and Witherspoon, 1983). Tsang and Tsang (1987) extended this model to include a limited number of tortuous and intersecting channels each characterized by an aperture density distribution, effective channel width, and correlated in length and aperture (Figure 20.25). A gamma function was used to characterize the aperture distribution and standard geostatistical techniques used to calculate distributions. Johns and Roberts (1991) presented a two-aperture channel model in which lateral transfer of mass from large to small aperture regions of the fracture plane and vertical diffusion into the rock matrix were considered. Moreno et al. (1988) presented a stochastic model where flow and transport in the entire fracture plane was considered. The aperture of the fracture was lognormally distributed and possessed a spatial correlation length (λ). From this it was concluded that a broad distribution of apertures with spatial correlation length on the same order of magnitude as the scale of measurement is responsible for flow channeling phenomena.

Determining the spatial distribution and correlation of fracture aperture at the field scale is difficult due to the extreme cost required to obtain sufficient field data. Consequently, conceptual models of variable aperture based on surface roughness are well verified at scales up to a meter but have yet to be verified with field data at scales in the tens to hundreds of meters. Thus, we are left with using simple correction factors such as macroscopically defined tortuosity to account for this variability.

20.5.2 Conceptual Models for a Fracture Network

In all rock types, fractures of various sizes and lengths together combine to form three-dimensional networks of interconnected groundwater pathways. At scales less than that defined for a single continuum, determining an appropriate conceptual model for the three-dimensional arrangement of fractures is essential in understanding and predicting groundwater flow and solute transport in fractured rock. In general, a distinction must be made between conceptual network models for sedimentary rock and those for crystalline rock. In most sedimentary rock environments, the directions of the principal fractures follow bedding planes and other pre-existing planes of weakness related to deposition. These fractures

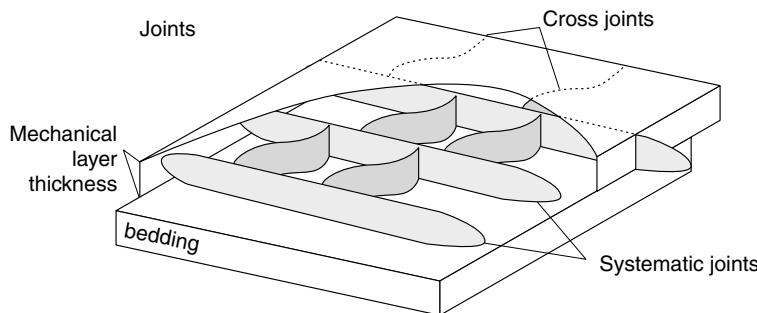


FIGURE 20.26 Typical three-dimensional view of conceptual fracture framework for a layer-cake stratigraphy. (From Engelder, T., Fischer, M.P., and Gross, M.R., 1993. *Geological Aspects of Fracture Mechanics, A Short Course Manual*, Geological Society of America, 281pp.)

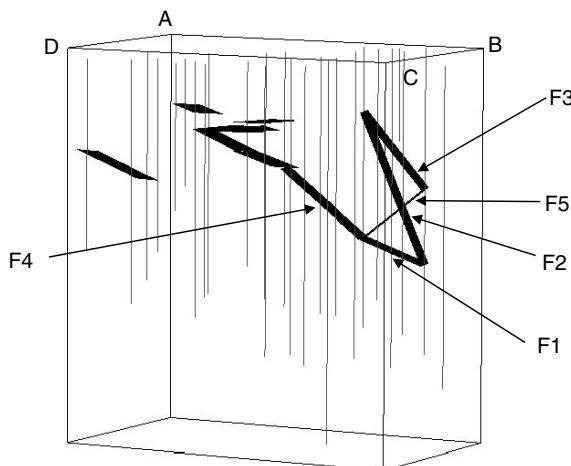


FIGURE 20.27 A three-dimensional depiction of the conceptual model developed using pulse interference tests for fracture planes (dark) and interconnections in a complex limestone site. The light vertical lines indicate the position of the boreholes that were used to conduct the interference testing.

are often connected by near-orthogonal fractures whose geneses are related to paleo- and neotectonic stresses (Holst, 1982; Williams et al., 1985). For flat-lying stratigraphy, this results in a relatively simple fracture framework that can be conceptualized by several discrete horizontal fractures connected by statistically defined sets of sub-vertical fractures. Figure 20.26 illustrates a typical three-dimensional view of a conceptual fracture framework for a layer-cake stratigraphy. Note that the vertical fractures are constrained by the bedding and have a preferred orientation. There are exceptions to this scenario, however, for sedimentary environments that have undergone considerable deformation. In such cases, the stratigraphy tends to be inclined and folded. This concentrates vertical fractures in the vicinity of fold axes and leads to shearing movement on some bedding planes. An example of this is provided in Figure 20.27 which illustrates the primary interconnected fractures in a folded and faulted limestone (Stephenson et al., 2006). This conceptual model was developed primarily from the results of pulse interference tests and was verified by comparison to the results of a steam flood in which the pathways of the hot water front was monitored.

Conceptual models for crystalline rock environments are usually more complex than that for sedimentary rock, although the example shown in Figure 20.27 for limestone is an exception. In uniform granitic rock, fractures may be oriented in preferred directions according to magma emplacement and cooling



FIGURE 20.28 Three-dimensional conceptual model of disc-shaped fractures that are of random size and oriented orthogonally. (From Long, J.C.S., Gilmour, P., and Witherspoon, P.A., 1985. *Water Resour. Res.* 21(8): 1105–1115.)

(polygonal fracture sets), local deformation, regional deformation, and erosional unloading (sheeting fractures).

One of the earliest conceptual models for flow in crystalline rock was based on the assumption of an equivalent porous medium (Snow, 1969). The concept was developed to relate the results of permeability measurements obtained from boreholes to a hydraulic conductivity ellipsoid that defines the anisotropic permeability of the bulk rock. To conduct the interpretation, it was assumed that flow was predominated by the fractures, all fractures contributed to the flow system, and the orientations of the fractures intersecting the boreholes were known. The hydraulic conductivity ellipsoid was constructed by summing the permeability contributions of each fracture intersecting each test interval.

It has long been recognized that fractures form disc-shaped discontinuities, particularly in granitic environments. Figure 20.28 shows a three-dimensional conceptual model in which disc-shaped fractures are of random size but oriented orthogonally. This concept can be extended to the general case for more

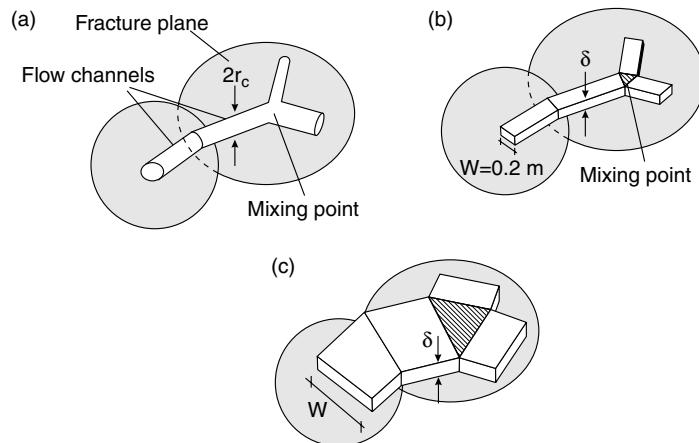


FIGURE 20.29 Three types of interconnecting flow channels in a fracture network: (a) tubular flow channels, (b) parallel plate channels with constant width, and (c) parallel plate channels with width equal to the fracture intersection line. (From Dverstorp, B., Andersson, J., and Nordqvist, W., 1992. *Water Resour. Res.* 28(9): 2327–2343.)

random orientations of discs. For example, Cacas et al. (1990) and Dverstorp et al. (1992) developed numerical flow and transport models based on a three-dimensional distribution of discs randomly placed in space and of random radius, but given a specific spatial orientation to simulate fracture sets. Groundwater flow and solute transport were then assumed to follow one of three types of linear flow channels that interconnect the individual fractures in the network (Figure 20.29).

Recently, scaling relations have been introduced into these types of conceptual models through the use of fractals and other methods (Piggott et al., 1997; Mazurek et al., 2003) thus providing a means of using field data in the generation of artificial fracture networks. Figure 20.30 illustrates the generation of a three-dimensional network of fractures based on fracture traces measured from outcrop scan lines, air photographs, and geophysical data (Piggott et al., 1997). From the trace data (Figure 20.30a) statistically generated two-dimensional traces are extended to produce the three-dimensional network of fractures (Figure 20.30b). A detailed three-dimensional view of a portion of the fracture network in Figure 20.30b is shown in Figure 20.30c.

20.6 Modeling Flow and Transport in Fractures and Fracture Networks

In some cases, such as where sheeting fractures dominate the groundwater flow system, it may be necessary only to model these discrete features and not the flow system as a whole. Simple analytical models for flow and transport in a single fracture (e.g., Tang et al., 1981) or a set of parallel fractures (Sudicky and Frind, 1982; West et al., 2004) may suffice. For example, transport calculations such as that shown in Figure 20.7 or Figure 20.8 may provide the information necessary in illustrating the potential importance of matrix diffusion at a given site.

Non-uniform aperture in the fracture plane, as discussed in the previous section, will influence solute transport in single fractures. Direct solution of the flow and transport in a fully defined aperture distribution can be conducted using numerical models (e.g., Lapcevic, 1997). For example, Figure 20.31 compares a tracer plume created assuming a constant aperture to one simulated assuming spatially variable aperture. Note that the variability in aperture leads to a plume of much greater irregularity in shape and of lower concentration. Thus, for some distributions of aperture it may be necessary to consider aperture variable in simulation at the discrete fracture scale. Unfortunately, in almost all field studies there is insufficient information to generate a defensible model of variable aperture.

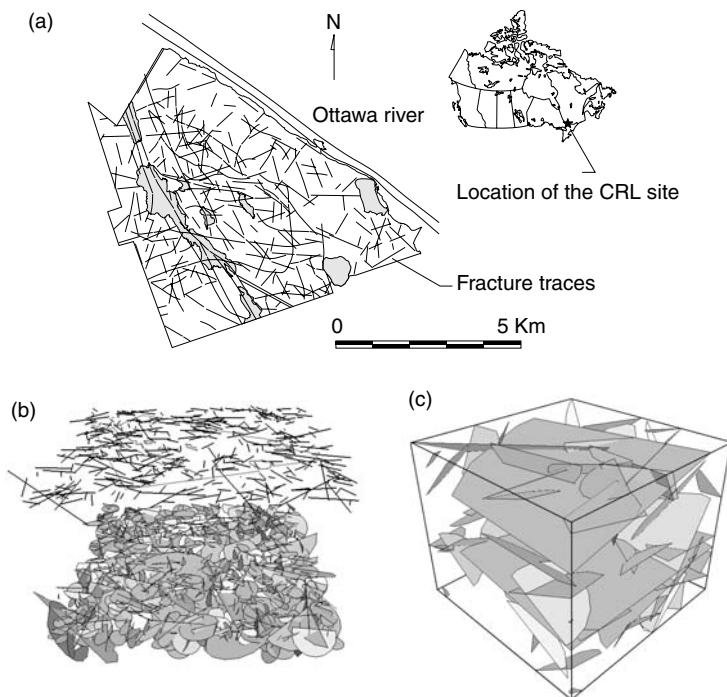


FIGURE 20.30 Generation of a conceptual model of fracture network using field data and fractal relations: (a) fracture traces measured from outcrop scan lines, air photographs, and geophysical surveys, (b) statistically generated two-dimensional traces (top) extended to form three-dimensional network of fractures (bottom), and (c) detailed portion of network shown in bottom part of (b). (From Piggott, A., Moltyaner, G., Yamazaki, L., and Novakowski, K., 1997. Preliminary characterization of fracturing at the site of the Chalk River Laboratories, In *Proceedings of 50th Canadian Geotechnical Conference, Ottawa, Canada* [also NWRI No. 97-126].)

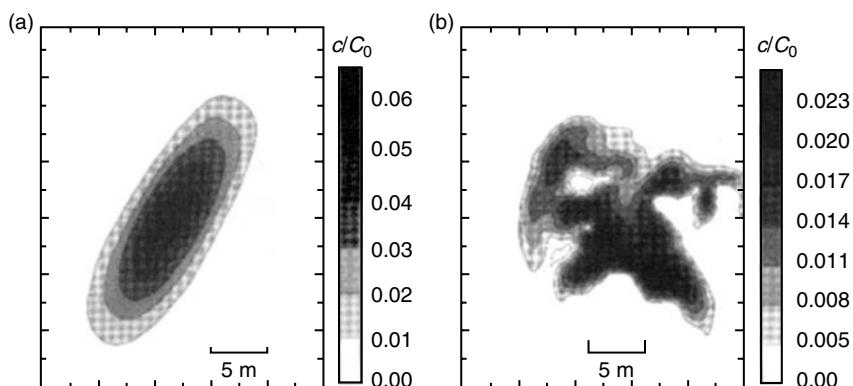


FIGURE 20.31 Two-dimensional concentration plumes in a single fracture illustrating the effect of variable aperture on plume dimensions: (a) tracer plume assuming a constant aperture and (b) tracer plume assuming aperture of the fracture is variable.

Modeling approaches to simulate flow and transport in fracture networks fall into one of three categories within the range of conceptual models for fractured rock (1) equivalent porous medium (EPM), (2) dual porosity, and (3) discrete fracture representation. In addition, fracture network models may be implemented in either two or three dimensions.

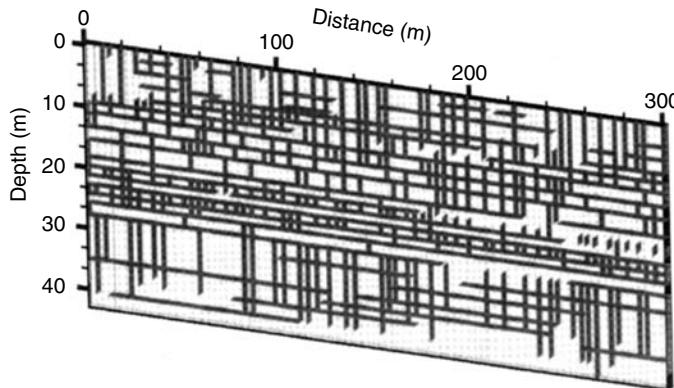


FIGURE 20.32 Network of orthogonal fractures generated based on a conceptual model of layered sedimentary rock where each layer has differing fracture spacing and connectivity parameters. System is divided into three layers: a densely fractured zone (0–26 m), a relatively unfractured zone (26–28 m) and a more sparsely fractured zone (28–45 m).

Models based on equivalent porous medium treat the fractured porous rock as equivalent to a non-fractured continuum. Bulk parameters for the permeability of the rock mass are used and the geometry of individual fractures or the rock matrix is not considered. This is a reasonable approach for the simulation of flow if fracturing is intense or the study domain is sufficiently large that individual fractures have no influence on the overall flow system (e.g., some regional systems of the scale of kilometers). Berkowitz et al. (1988) and Schwartz and Smith (1988) discuss the fundamentals of continuum models for fractured rock systems and Min et al. (2005) provide a recent example of the application at large scale.

Double porosity models are used to attempt to bridge the gap between the simplifications of EPM models and the details of the discrete fracture models by treating the fracture system and the porous matrix as two separate inter-related continua (Barenblatt et al., 1960). In this approach, equations of flow and transport for each system are linked by a source/sink term that describes the fluid or solute exchange between the two systems each of which may have very different properties relative to the other. Examples of the use of semi-analytical and numerical models having this approach can be found in Huyakorn et al. (1983); Rowe and Booker (1990); Sudicky (1990); Hassan and Mohamed (2003); and Nair et al. (2004). Some limitations to the double-porosity approach include the assumption that the matrix blocks are of simple (and possibly unrealistic) geometry and that advection of solutes in each block is ignored.

The discrete fracture approach will result in the most physically representative simulation of flow and transport processes at the sub-continuum scale. However, discrete fracture models require the generation of fracture networks based on a conceptual model developed using information on both the individual fractures and the geometry of inter-fracture relationships. A network of fractures is characterized by a distribution of fractures of fixed or variable aperture, finite length, regular, or random orientations, and some degree of connectivity with other individual fractures. Thus, the development of the conceptual model strongly influences the outcome of the model simulations. Figure 20.32 illustrates a fracture network generated using a conceptual model of fracturing in a layered sedimentary sequence of dolostone and shale. To generate this network minimum horizontal and vertical fracture spacings of 1.5 and 3.5 m were set for the entire sequence. Within each geologic unit a range of fracture lengths and fracture density were set. Note the denser fracture zones at the top of the sequence (0–26 m), a relatively unfractured zone at 26 m and a more sparsely fractured zone from 28 to 45 m.

20.6.1 Flow in Discrete Fracture Networks

The modeling of flow in discrete fracture networks may be conducted to estimate groundwater flux for the purpose of groundwater resource evaluation or to provide an estimate of the distribution of groundwater

velocity in individual fractures as input to a solute transport model. In the former case, transient hydraulic head conditions may be of interest, whereas in the latter, only steady conditions are considered. In the following, two approaches to the modeling of flow in discrete features are described, both of which are suitable for either transient or steady flow conditions.

Barker (1991) developed a robust semi-analytical method that can be used to determine the distribution of groundwater flux and velocity in a two-dimensional network of fractures having random orientations. The fracture network is conceptualized as a linked system of linear elements for both steady and transient flow conditions (e.g., Dverstorp et al., 1992). Flux into each node is calculated by summing the contribution from each fracture using Darcy's law. The cubic law is used to relate the fracture aperture to elemental hydraulic conductivities. The system of Laplace transformed equations is solved by direct or iterative methods depending on the size of the network and the results numerically inverted from the Laplace domain to obtain values in real space.

Modeling groundwater flow in three-dimensional fracture networks is considerably more difficult than for two-dimensional slices. For conceptual models involving sparsely distributed fractures in impermeable rock, randomly oriented fracture discs and fracture intersections of various orientations must be discretised. However, the mesh generation required for solution with standard finite element or finite difference techniques is very difficult, even for small solution domains. Alternative solution methods have been attempted including a hybrid analytical-numerical scheme (Long et al., 1985) and a boundary element formulation (Elsworth, 1986), but were found not to be generally useful for large complex networks. At present the most widely used model for simulating flow in networks is the proprietary code package known as FracMan/Mafic (Golder Associates Inc.). Example applications using this code are found in Kim et al. (2004) and Pashin (2005).

Therrien and Sudicky (1996) derived a model for a variably saturated fracture–matrix system having three dimensions in which advective fluid exchange is allowed between the fractures and matrix. A modified form of the Richards' equation was used for determining hydraulic head in the matrix and an extension to the variably saturated flow equations was used for hydraulic head in the fracture. The fractures were idealized as two-dimensional parallel plates and fluid leakage flux was used to link the equations for the fracture with that of the matrix. The resulting system of governing equations was solved using the control-volume finite-element method in conjunction with Newton–Raphson linearization. Oxtobee and Novakowski (2003) used this model to simulate groundwater–surface water interchange in a stream flowing directly on bedrock. Connection between the stream and the underlying orthogonal fracture network was simulated through sparsely distributed vertical fractures. The results showed that capture zones in the fracture network could develop over a substantial width (several hundreds of meters) even though discharge to the stream was limited to a single fracture no larger than $1000 \mu\text{m}$. Although there is no technical limitation to the method in the simulation of flow in a non-orthogonal fracture system, as aforementioned, any problem simulated using this model would be significantly limited in scale. Recent advances in the formulation of the model, however, have shown that large-scale non-orthogonal fractures can be accommodated under some conditions (Graf and Therrien, 2005).

20.6.2 Solute Transport in a Discrete Fracture Network

Solute transport in fracture networks can be simulated in two ways (1) using particle tracking, and (2) direct solution of the governing equations for solute transport. Of the two methods, the former has received more widespread use. In the following sections, both particle tracking and methods for direct solution will be described, with focus on the former.

In modeling solute transport in fracture networks, the presence of fracture intersections where solute must be apportioned according to the flux entering and leaving the intersection node, must be incorporated. There are two approaches to this (1) assume complete mixing at the node, and (2) use streamtube routing. For the former, concentration entering the intersection is assumed to completely mix resulting in a uniform distribution of concentration leaving the intersection. For the latter, mass is distributed at the intersection according to the distribution of streamtubes and no mixing is assumed to occur (Figure 20.33).

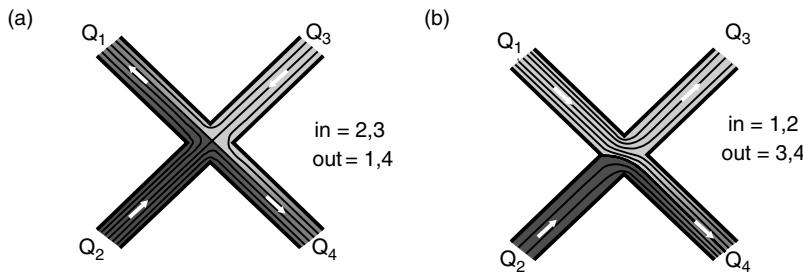


FIGURE 20.33 Schematic diagram showing mass transfer at a fracture intersection according to the distribution of streamtubes. (a) $Q_2 = Q_3$ and transport is split evenly into Q_1 and Q_4 and (b) $Q_1 > Q_2$ and transport is split unevenly between Q_3 and Q_4 .

Complete mixing at an intersection having n incoming fractures is described by (Küpper et al., 1995a):

$$c = \frac{\sum_{i=1}^n c_i Q_i}{\sum_{i=1}^n Q_i} \quad (20.25)$$

where c is the concentration in the intersection, and Q_i and c_i are the volumetric flux and concentrations in the incoming fractures, respectively. The development required for a rigorous implementation of streamtube routing is considerably more complex (e.g., Philip, 1988).

Berkowitz et al. (1994) suggested that diffusional transfer between streamtubes at intersections may smear the distinction between streamtubes, resulting in a distribution of concentration more like that observed with complete mixing. However, Küpper et al. (1995b) showed that, even in the absence of diffusion, for three of four possible flow conditions at intersections, the complete mixing assumption and streamtube routing produce identical results. In addition, it was also shown that for some types of flow (i.e., radial flow in a network), the flow condition giving rise to the difference is rare. Thus, in practical terms, the degree of error introduced into a model by assuming complete mixing at every intersection, is likely not significant. However, it is important to note that these concepts apply only to mixing at linked one-dimensional elements. In the more realistic case where two-dimensional planes intersect, the hydraulic head will be non-constant along the intersection resulting in much more complex mixing arrangements. To date, this issue has not been rigorously addressed.

Particle tracking is a means to simulate solute transport by following the residence times of particles released into the flow system at a given boundary or internal location. In the most rudimentary implementation, a given number of particles, usually 20,000–30,000, are released into the flow system, routed at fracture intersections according to volumetric flux (i.e., assume perfect mixing at the intersection node) and eventually tracked to an exit boundary (Schwartz et al., 1983). The number of particles are summed at the exit boundary resulting in a breakthrough curve equivalent to a dirac input. Integration of the breakthrough curve with respect to time yields a breakthrough curve equivalent to a step-function input.

The use of particle tracking methods was expanded in the 1990s to account for solute retardation (Dverstorp et al., 1992), and variable aperture fracture elements (Nordqvist et al., 1992). Retardation is accounted for by manipulating the travel time in individual fracture elements using the expression:

$$t_R = t_C R_a \quad (20.26)$$

where t_R is the travel time adjusted for sorption, t_C is determined from the cubic law, and R_a is as defined in Equation 20.10. Nordqvist et al. (1992) also manipulated residence time in fracture elements to account for the influence of variable-aperture fractures. This was conducted by generating a spectrum of residence times using a variety of realizations of a variable aperture fracture having different mean apertures and

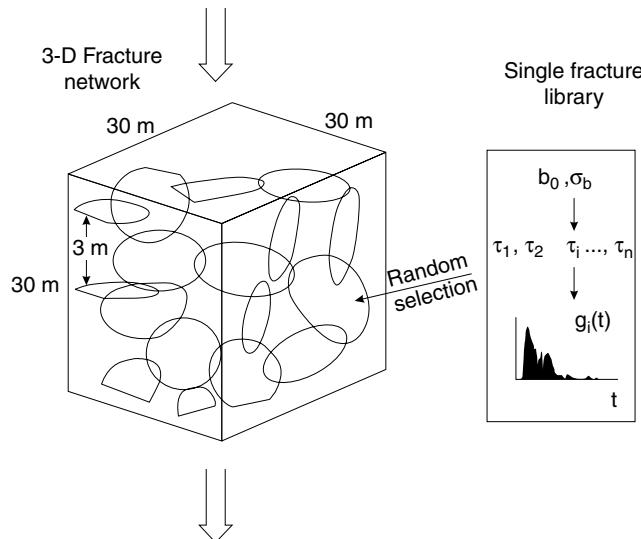


FIGURE 20.34 The incorporation of variable aperture fractures is introduced in a fracture network model. Here, b_0 and σ_b are the two parameters describing the lognormal aperture distribution (mean aperture and standard deviation). The τ_i are effective transmissivities of the fractures in the library. The $g_i(t)$ plots are graphical illustrations of the residence time spectra for the fractures in the library. (From Nordqvist, A.W., Tsang, Y., Tsang, C.-F., Dverstorp, B., and Andersson, J., 1992. *Water Resour. Res.* 28(6): 1703–1713).

different properties of the log-normal aperture distribution. To develop the residence time distribution for the system of linked linear elements, the residence time spectrum was randomly sampled for each fracture element (Figure 20.34). Recent work by Tsang and Doughty (2003) showed that matrix diffusion along with heterogeneity of the aperture distribution, the presence of fracture gouge, and mineral alteration can also be accommodated using particle tracking techniques.

For two-dimensional networks of linked linear fracture elements, direct solution offers the most robust means of simulating transport in the fractures and interactions with the matrix. Sudicky and McLaren (1992) and Novakowski and Bogan (1999) describe finite-element and semi-analytical approaches, respectively.

The solution method described by Sudicky and McLaren (1992) accommodates advection-dispersion, retardation, and decay in the fracture and the same in the matrix. A third-type boundary condition is used for the exchange between fracture and matrix. This solution method can therefore simulate fractured domains having a permeable matrix (e.g., fractured clays). The finite element method is used to discretise the equations in the spatial coordinates and the Laplace transform is used to manage the time derivative. To minimize the difficulties in discretisation and to maximize the domain size for tractable problems, fracture geometry has been limited to orthogonal systems. For two-dimensional networks having more random fracture orientations, the solution method of Novakowski and Bogan (1999) offers a practical alternative. In this case, only diffusive transport is allowed between the fracture and matrix. A semi-infinite domain is also defined for the diffusive transport in the matrix. Clearly, this is an approximation that is viable only for geological material having sparse fractures and low matrix porosity.

In some settings, the conceptual model for transport has been developed to the degree where a three-dimensional discrete fracture simulation is viable. In three dimensions, the direct solution method of Therrien and Sudicky (1996) for a variably saturated fractured medium can be employed. The two governing equations for the fracture planes and matrix blocks are linked using continuity between the concentration in the fracture and that in the matrix at the fracture wall. As the transport equations are linear, solution is based on a standard time-marching Galerkin scheme. Two-dimensional rectangular elements are used for the fractures and three-dimensional rectangular prisms were used for the matrix.

As with the flow case, non-orthogonal fractures present a difficulty, although the newest version of the model (Graf and Therrien, 2005) can accommodate these.

At present, in the general practice of hydrogeology, the modeling of flow and transport in fractured rock systems is often conducted using EPM models. As we have suggested above, this approach will lead to error in the prediction of both groundwater flux and solute travel time, depending on the size of the domain modeled. At a domain scale of several hundreds of meters, or more, however, prediction of groundwater flux may be conducted with only minor error using EPM models, provided the characteristics of the major fracture zones and unfractured rock are known and appropriately represented. Thus, the hydraulic effects of short or long-term pumping or changes in boundary conditions such as recharge, can be reliably explored using this approach.

Unfortunately, the prediction of solute transport using the EPM approach at this scale can be substantially erroneous. Thus, as a general rule of thumb, all predictions of contaminant transport and the design of remedial facilities at the scale of a contaminated site should be conducted using a model which, at a minimum, incorporates multi-continuum concepts. In some cases, a defensible modeling strategy may be to simulate transport in the fracture system with simple analytical models that incorporate a parallel system of fractures and matrix diffusion. Although the fracture system may not be well represented by this approach, it is likely that this approach will lead to substantially more accurate predictions of contaminant transport than were an EPM model employed.

For conditions where more accurate simulations of contaminant transport are required, such as for large-scale groundwater remediation program, the simulation of controlled tracer experiments conducted in fracture networks, or investigations of intrinsic bio-remediation, there are now very robust modeling tools available that account for fractures on a discrete basis. Unfortunately, there are significant practical limitations in the application of some models that are directly related to the ability to discretise the fractures and matrix in detail sufficient enough to provide the necessary accuracy in the predicted transport rate and concentration. For example, a radially divergent tracer experiment conducted in a discrete fracture in a slightly porous dolostone over a distance of approximately 125 m (Novakowski et al., 1999) has yet to be successfully interpreted using a direct numerical solution. This is because of the computational limitations which, for this problem, are related to the excessive number of elements required (over 60 million estimated).

Glossary

Aperture: Separation distance between two fracture surfaces used as a measure of fracture width.

Borehole: Drilled open hole in rock. Boreholes may be of any orientation and plunge from vertical to horizontal in relation to the earth's surface.

Channel: Preferred pathway in a single fracture plane or fracture network.

Conceptual Model: Geological or hydrogeological image representing natural system.

Fracture Network: Two- or three-dimensional arrangement of intersecting fractures. Fractures in a network may be of identical aperture, length, or orientation or may represent a distribution of fractures of varying parameters.

Fracture: Any planar or curviplanar discontinuity or break in a rock mass that has formed as a result of a brittle deformation process. Joints, shear fractures, faults, microcracks etc. are all examples of fractures.

Hydraulic Test: Test involving the injection or withdrawal of water into a given test zone to determine the permeability of the zone. Either the transient or steady-state response can be used.

Interference Test: Hydraulic test conducted between two or more hydraulically connected zones. A vertical interference test can be conducted between two isolated zones in a single borehole.

Joint: A mesoscopic fracture in rock exhibiting purely opening mode displacement, with no appreciable shear offset.

Matrix Diffusion: Process whereby solute from water in fracture is transferred to water in pore spaces of rock through chemical diffusion.

Sheeting Structure: Mesoscopic fractures formed due to the vertical expansion of the rock resulting from erosional unloading. Sheeting structures are often found parallel to earth's surface in all types of rock and are common conduits for groundwater flow at the regional scale.

Skin: Zone around borehole that has different properties than the bulk rock mass. A positive skin suggests the permeability around the borehole is less than the bulk rock whereas a negative skin suggests greater permeability. Skin zones may be caused by fracture infilling due to rock flour, drill bit damage or invasion of drilling muds.

Surface Roughness: Texture of fracture surface due to asperities in the plane. Roughness can range from smooth planes to highly irregular surfaces.

Rock Matrix: Structure of unfractured bulk rock mass composed of mineral grains, cement, pore water, and pore space.

Wellbore Storage: Volume of water in the borehole which, when pumping is initiated, is first depleted prior to water stored in fracture or rock matrix.

References

- Abelin, H., Birgersson, L., Gidlund, J., and Neretnicki, I., 1991. A large-scale flow and tracer experiment in granite 1. Experimental design and flow distribution, *Water Resour. Res.*, 27(12): 3107–3117.
- Atkinson, B.K., 1987. *Fracture Mechanics of Rock*, Academic Press, Orlando, FL, 534 pp.
- Barenblatt, G.I., Zheltov, I.P., and Kochina, I.N., 1960. Basic concepts in the theory of seepage of homogeneous liquids in fissured rocks, *J. Appl. Math. Mech.*, (Engl. Transl.) 24(5): 852–864.
- Barker, J.A., 1991. Reciprocity principle and analytical solution for Darcian flow in a network, *Water Resour. Res.*, 27(5): 743–746.
- Bear, J., 1972. *Dynamics of Fluids in Porous Media*, Elsevier, New York, 764 pp.
- Bear, J., 1993. Modeling flow and contaminant transport in fractured rocks. In *Flow and Contaminant Transport in Fractured Rocks*, eds. J. Bear, C.-F. Tsang, and G. de Marsily, pp. 1–37. Academic Press Inc., San Diego, CA.
- Bear, J., Tsang, C.-F., and de Marsily, G., 1993. *Flow and Contaminant Transport in Fractured Rocks*, Academic Press Inc., San Diego, CA, 560 pp.
- Berkowitz, B., Naumann, C., and Smith, L., 1994. Mass transfer at fracture intersections: an evaluation of mixing models. *Water Resour. Res.*, 30(6): 1765–1773.
- Berkowitz, B., Bear, J., and Braester, C., 1988. Continuum models for contaminant transport in fractured porous formations, *Water Resour. Res.*, 24(8): 1225–1236.
- Black, W.H., Smith, H.R., and Patton, F.D., 1987. Multiple-level groundwater monitoring with the MP system. In *Proc. NWWA-AGU Conf. on Surface and Borehole Geophysical Methods and Groundwater Instrumentation*, NWWA, Dublin, Ohio.
- Bliss, J.C. and Rushton, K.R., 1984. The reliability of packer tests for estimating the hydraulic conductivity of aquifers, *Q. J. Engrg. Geol.*, 17: 88–91.
- Bredehoeft, J.D. and Papadopoulos, S.S., 1980. A method for determining the hydraulic properties of tight formations, *Water Resour. Res.*, 16(1): 233–238.
- Brown, S.R., Kranz, R.L., and Bonner, B.P., 1986. Correlation between the surfaces of natural rock joints, *Geophys. Res. Lett.*, 13(13): 1430–1433.
- Brown, S.R., 1995. Simple mathematical model of a rough fracture, *J. Geophys. Res.*, 100(B4): 5941–5952.
- Cacas, M.C., Ledoux, E., de Marsily, G., Tillie, B., Barbareau, A., Calmels, P., Gaillard, B., and Margritta, R., 1990. Modelling fracture flow with a stochastic discrete fracture network: calibration and validation 1. The flow model. *Water Resour. Res.* 26(3): 479–489.
- Cherry, J.A. and Johnson, P.E., 1982. A multilevel device for hydraulic head monitoring and groundwater sampling in fractured rock. *Ground Water Monitoring Rev.*, 2(3): 41–44.
- Cooper, H.H., Bredehoeft, J.D., and Papadopoulos, I.S., 1967. Response to a finite diameter well to an instantaneous charge of water, *Water Resour. Res.*, 3(1): 263–269.

- Deltombe, J.-L. and Schepers, R., 2004. New developments in real-time processing of full waveform acoustic televiewer data. *J. Appl. Geophys.*, 55(2004): 161–172.
- de Marsily, G., 1986. *Quantitative Hydrogeology*, Academic Press, Orlando, FL.
- Doe, T. and Remer, J. 1980. Analysis of constant-head well tests in non-porous fractured rock. In *Third Invitational Well-Testing Symposium — “Well testing in Low Permeability Environments”*, Berkeley, California, pp. 84–89.
- Doe, T., Osnes, J., Kenrick, M., Geier, J., and Warner, S., 1987. Design of well testing programs for waste disposal in crystalline rock. In *Proceedings of 6th Congress of the International Society for Rock Mechanics, Montreal, Canada*.
- Dougherty, D.E. and Babu, D.K., 1984. Flow to a partially penetrating well in a double porosity reservoir, *Water Resour. Res.*, 20(8): 1116–1122.
- Drost, W., Klotz, D., Koch, A., Moser, H., Neumaier, F., and Werner R., 1968. Point dilution methods of investigating ground water flow by means of radioisotopes, *Water Resour. Res.*, 4(1): 125–146.
- Dunnicliif, J., 1988. *Geotechnical Instrumentation for Monitoring Field Performance*, John Wiley & Sons, Canada, Ltd.
- Dverstorp, B., Andersson, J., and Nordqvist, W. 1992. Discrete fracture network interpretation of field tracer migration in sparsely fractured rock. *Water Resour. Res.*, 28(9): 2327–2343.
- Earlougher, R.C., Jr., 1977. *Advances in Well Test Analysis, Monogr. 5*, Society of Petroleum Engineers, Dallas, TX.
- Elsworth, D., 1986. A model to evaluate the transient hydraulic response of three-dimensional sparsely-fractured rock masses. *Water Resour. Res.*, 22(13): 1809–1819.
- Engelder, T., Fischer, M.P., and Gross, M.R., 1993. *Geological Aspects of Fracture Mechanics, A Short Course Manual*, Geological Society of America, 281 pp.
- Faust, C.R. and Mercer, J.W., 1984. Evaluation of slug tests in wells containing a finite-thickness skin, *Water Resour. Res.*, 20(4): 504–506.
- Ford, D.C., Palmer, A.N., and White, W.B., 1988. Landform development:karst. In *Hydrogeology*, eds. Back, W., Rosenshein, J.S., and Seaber, P.R. Vol. O-2, pp. 401–412, Geological Society of America, Boulder, CO, The Geology of North America.
- Gentier, S. and Billaux, D., 1989. Caracterisation en laboratoire de l ’espace fissural d’une fracture, In *Proceedings of the International Symposium on Rock at Great Depth*, Vol.1, pp. 425–431, A.A. Balkema, Rotterdam, The Netherlands.
- Graf, T. and Therrien, R. 2005. Variable-density groundwater flow and solute transport in porous media containing nonuniform discrete fractures *Adv. Water Resour.*, 28: 1351–1367.
- Grisak, G.E., Merritt, W.F., and Williams, D.W., 1977. A fluoride borehole dilution apparatus for groundwater velocity measurements, *Can. Geotech. J.*, 14: 554–561.
- Hakami, E. 1995. *Aperture Distribution of Rock Fractures*, Ph.D. thesis, Dept. of Civil and Environmental Engineering, Royal Institute of Technology, Stockholm, Sweden.
- Hassan, A.E. and Mohamed, M.M., 2003. On using particle tracking methods to simulate transport in single-continuum and dual continua media. *J. Hydrol.*, 275: 242–260.
- Hess, A.E. and Paillet, F.L., 1990. Applications of the thermal-pulse flowmeter in the hydraulic characterization of fractured rocks, *ASTM STP*, 1101: 99–112.
- Holst, T.B., 1982. Regional jointing in the northern Michigan Basin. *Geology.*, 10: 273–277.
- Holzhauser, G.R., 1989. Origins of sheet structure, 1. Morphology and boundary conditions. *Engng. Geol.* 27: 225–279.
- Huyakorn, P.S., Lester, B.H., and Mercer, J.W. 1983. An efficient finite element technique for modeling transport in fractured porous media, 1. Single species transport, *Water Resour. Res.*, 19(3): 841–854.
- Johns, R.A. and Roberts, P.V., 1991. A solute transport model for channelized flow in a fracture, *Water Resour. Res.*, 27(8): 1797–1808.
- Jouanna, P., 1993. A summary of field test methods in fractured rocks, In *Flow and Contaminant Transport in Fractured Rocks*, eds. J. Bear, C.-F. Tsang, and G. de Marsily, pp. 437–543. Academic Press Inc., San Diego, CA.

- Karasaki, K., Long, J.C.S., and Witherspoon, P.A. 1988. Analytical models of slug tests, *Water Resour. Res.*, 24(1): 115–126.
- Kim, H.-M., Inoue, J., and Hideyuki, H., 2004. Flow analysis of jointed rock masses based on excavation-induced transmissivity change of rough joints. *Int. J. Rock Mech. Mining Sci.*, 41(6): 959–974.
- Konzuk, J. and Kueper, B., 2004. Evaluation of cubic law based models describing single-phase flow through a rough-walled fracture. *Water Resour. Res.*, 40(2), W02402, doi:10.1029/2003WR002356.
- Kulander, B.R., Dean, S.L., and Ward, B.J., 1990. *Fractured Core Analysis: Interpretation, Logging and Use of Natural and Induced Fractures in Core: AAPG Methods in Exploration Series*. No. 8, American Association of Petroleum Geologists, Tulsa, Ok, 88 pp.
- Küpper, J.A., Schwartz, F.W., and Steffler, P.M., 1995a. A comparison of fracture mixing models, 1. A transfer function approach to mass transport modeling. *J. Contam. Hydrol.* 18: 1–32.
- Küpper, J.A., Schwartz, F.W., and Steffler, P.M., 1995b. A comparison of fracture mixing models, 2. Analysis of simulation trials. *J. Contam. Hydrol.* 18: 33–58.
- La Pointe, P.R. and Hudson, J.A., 1985. Characterization and interpretation of rock mass joint patterns, *Geological Society of America Special Paper* 199, 37pp.
- Lapcevic, P.A., Novakowski, K.S., and Cherry, J.A., 1990. The characterization of two discrete horizontal fractures in shale, In *Proc. Technology Transfer Conference, Ontario Ministry of Environment, Nov., Toronto, Ontario. Vol II*, pp. 486–495.
- Lapcevic, P.A., 1997. *Tracer Experiments Conducted in a Discrete Horizontal Fracture under Conditions of Forced Hydraulic Gradient and Natural Groundwater Flow*, MSc. Thesis, University of Waterloo, Waterloo, ON, 98 pp.
- Lau, J.S.O., 1983. The determination of true orientations of fractures in rock cores, *Canad. Geotech. J.*, 20: 221–227.
- Long, J.C.S., Gilmour, P., and Witherspoon, P.A., 1985. A model for steady fluid flow in random three-dimensional networks of disc-shaped fractures. *Water Resour. Res.*, 21(8): 1105–1115.
- Mazurek, M., Jakob, A., and Bossart, P., 2003. Solute transport in crystalline rocks at Aspo I: Geological basis and model calibration. *J. Contam. Hydrol.*, 61: 157–164.
- McKenna, S.A., Meigs, L.C., and Haggerty, R., 2001. Tracer tests in fractured dolomite, 3. Double porosity, multiple-rate transfer processes in two-well convergent-flow tests. *Water Resour. Res.*, 37(5): 1143–1154.
- Min, K.-B., Rutqvist, J., Tsang, C.-F., and Jing, L., 2005. Thermally induced mechanical and permeability changes around a nuclear waste repository — a far-field study based on equivalent properties determined by a discrete approach. *Int. J. Rock Mech. Mining Sci.*, 42: 765–780.
- Moench, A.F. and Hsieh, P.A., 1985. Analysis of slug test data in a well with finite thickness skin, In *Memoirs of the 17th International Congress on the Hydrogeology of Rocks of Low Permeability*, Vol. 17, pp.17–29, International Association of Hydrologists, Tucson, AZ.
- Moreno, L., Tsang, Y.W., Tsang, C.F., Hale, F.V., and Neretnieks, I., 1988. Flow and tracer transport in a single fracture: a stochastic model and its relation to some field observations, *Water Resour. Res.*, 24(12): 2033–2048.
- Nair, R., Abousleiman, Y., and Zaman, M., 2004. A finite-element porothermoelastic model for dual-porosity media. *Int. J. Numer. Anal. Meth. Geomech.*, 28: 875–898 (DOI: 10.1002/nag.336)
- Neuzil, C.E., 1982. On conducting the modified > slug = test in tight formations, *Water Resour. Res.*, 18(2): 439–441.
- Nordqvist, A.W., Tsang, Y., Tsang, C.-F., Dverstorp, B., and Andersson, J., 1992. A variable aperture fracture network model for flow and transport in fractured rocks. *Water Resour. Res.* 28(6): 1703–1713.
- Novakowski, K.S., 1989. Analysis of pulse interference tests, *Water Resour. Res.*, 25(11): 2377–2387.
- Novakowski, K.S., 2005. Use of the Theim equation to interpret constant head tests conducted in boreholes. *J. Hydrol.*, submitted.
- Novakowski, K.S. and Lapcevic P.A., 1994. Field measurement of radial solute transport in fractured rock, *Water Resour. Res.*, 30(1): 37–44.

- Novakowski, K.S. and Bogan, J.D., 1999. A semi-analytical model for the simulation of solute transport in a network of fractures having random orientations. *Int. J. Num. Anal. Meth. Geomech.*, 23: 317–333.
- Novakowski, K.S., Lapcevic, P., Bickerton, G., Voralek, J., Zanini, L., and Talbot, C., 1999. The development of a conceptual model for contaminant transport in the dolostone underlying Smithville, Ontario. Final report prepared for the Smithville Phase IV Remediation Program (Available on CD), 175 pp.
- Novakowski, K.S., Bickerton, G., and Lapcevic, P. 2004. Interpretation of injection-withdrawal tracer experiments conducted between two wells in a large single fracture. *J. Contam. Hydrol.*, 73: 227–247.
- Novakowski, K.S., Bickerton, G., Lapcevic, P.A., Voralek, J., and Ross, N., 2006. Measurement of groundwater velocity in discrete fractures. *J. Contam. Hydrol.*, 82: 44–60.
- Odling, N.E., 1994. Natural rock profiles, fractal dimensions and joint roughness coefficients, *Rock Mech. Rock Engng.*, 27(3): 135–153.
- Oxtobee, J. and Novakowski, K.S., 2003. A numerical investigation of the principal processes governing the exchange of groundwater between a bedrock stream and a fractured rock aquifer. *J. Groundwater*, 41(5): 667–681.
- Paillet, F.L. and Kapucu, K., 1989. Characterization of fracture permeability and fracture flow modeling at Mirror Lake, New Hampshire, *U.S. Geological Survey Water Resources Investigations Report 89-4058*, 49p.
- Paillet, F.L., 2004. Borehole flowmeter applications in irregular and large-diameter boreholes. *J. Appl. Geophys.*, 55: 39–59.
- Pashin, J.C., 2005. Compartmentalization analysis of coalbed methane resources in the Black Warrior Basin; using 3-D computer models to balance industrial and environmental concerns. *AAPG Bull.*, 89(1): 151.
- Philip, J.R., 1988. The fluid mechanics of fractures and other junctions. *Water Resour. Res.*, 24(2): 239–246.
- Piggott, A.R. 1990 *Analytical and Experimental Studies of Rock Fracture Hydraulics*, Ph.D. thesis, Pennsylvania State Univ., University Park, PA.
- Piggott, A.R. and Elsworth, D., 1993. Laboratory assessment of the equivalent apertures of a rock fracture, *Geophys. Res. Lett.*, 20(13): 1387–1390.
- Piggott, A.R. Huynh, T.N.T., Lapcevic, P.A., and Novakowski, K.S. 1996. Automated analysis of hydraulic and tracer tests conducted in fractured rock, *Hydrogeology J.*, 4(3): 84–93.
- Piggott, A., Moltyaner, G., Yamazaki, L., and Novakowski, K.S. 1997. Preliminary characterization of fracturing at the site of the Chalk River Laboratories, In *Proceedings of 50th Canadian Geotechnical Conference, Ottawa, Canada* (also NWRI No. 97-126).
- Pyrak-Nolte, L.J., Myer, L.M., Cook, N.G.W., and Witherspoon, P.A., 1987. Hydraulic and mechanical properties of natural fractures in low permeability rock, In *Proceedings of 6th Congress of the International Society for Rock Mechanics, Montreal, Canada*, pp. 225–231.
- Ragan, D.M. 1973. *Structural Geology: An Introduction to Geometrical Techniques*, John Wiley & Sons, New York, NY, 208 pp.
- Raven, K.G. 1986. Hydraulic characterization of a small ground-water flow system in fractured monzonitic gneiss, *Nat. Hyd. Res. Inst. Scientific Series No. 149*, No. 30, 133pp.
- Raven, K.G., Novakowski, K.S., and Lapcevic, P.A., 1988. Interpretation of field tracer tests of a single fracture using a transient solute storage model, *Water Resour. Res.*, 24(12): 2019–2032.
- Rowe, R.K., and Booker, J.R. 1990. Contaminant migration through fractured till into an underlying aquifer, *Can. Geotech. J.*, 27(4): 484–495.
- Sageev, A., 1986. Slug test analysis, *Water Resour. Res.*, 22(8): 1323–1333.
- Sanford, B.V., Thompson, F.J., and McFall, G.H., 1985. Phanerozoic and recent tectonics of the Canadian Craton, in *The Geoscience Program — Proceedings of the Seventeenth Information Meeting of the Nuclear Fuel Waste Management Program*, Atomic Energy of Canada Limited, TR-299, pp. 334–352.
- Schwartz, F.W., Smith, L., and Crowe, A., 1983. A stochastic analysis of macroscopic dispersion in fractured media. *Water Resour. Res.* 19(5): 1253–1265.

- Schwartz, F.W. and Smith, L. 1988. A continuum approach for modeling mass transport in fractured media, *Water Resour. Res.*, 24(8): 1360–1372.
- Shapiro, A.M. and Nicholas, J.R. 1989. Assessing the validity of the channel model of fracture aperture under field conditions, *Water Resour. Res.*, 25(5): 817–828.
- Smith, L. and Schwartz, F.W., 1984. An analysis of the influence of fracture geometry on mass transport in fractured media. *Water Resour. Res.* 20(9): 1241–1252.
- Snow, D.T., 1969. Anisotropic permeability of fractured media. *Water Resour. Res.*, 5(6): 1273–1289.
- Stephenson, K., Novakowski, K.S., Davis, E., and Heron, G., 2006. Hydraulic characterization for steam enhanced remediation in fractured rock. *J. Contam. Hydrol.*, 82: 220–240.
- Sudicky, E.A. and Frind, E.O. 1982. Contaminant transport in fractured porous media: analytical solution for a system of parallel fractures, *Water Resour. Res.*, 18(6): 1634–1642.
- Sudicky, E.A., 1990. The Laplace transform galerkin technique for efficient time-continuous solution of solute transport in double-porosity media, *Geoderma*, 46: 209–232.
- Sudicky, E.A. and McLaren, R.G., 1992. The Laplace transform Galerkin technique for large-scale simulation of mass transport in discretely fractured porous formations. *Water Resour. Res.*, 28(2): 499–514.
- Tang, D.H., Frind, E.O., and Sudicky, E.A., 1981. Contaminant transport in fractured porous media: analytical solution for a single fracture, *Water Resour. Res.*, 17(3): 555–564.
- Therrien, R. and Sudicky, E.A., 1996. Three-dimensional analysis of variably-saturated flow and solute transport in discretely-fractured porous media, *J. Contam. Hydrol.*, 23(1–2): 1–44.
- Trainer, F.W. 1988. Hydrogeology of the plutonic and metamorphic rocks, In *Hydrogeology*, eds. Back, W., Rosenshein, J.S., and Seaber, P.R. Vol. O-2, pp. 367–380, Geological Society of America, Boulder, CO, The Geology of North America.
- Tsang, Y.W. and Witherspoon, P.A., 1983. The dependence of fracture mechanical and fluid flow properties on fracture roughness and sample size, *J. Geophys. Res.*, 88(B3): 2359–2366.
- Tsang, Y.W., 1984. The effect of tortuosity on fluid flow through a single fracture, *Water Resour. Res.*, 20(9): 1209–1215.
- Tsang, Y.W. and Tsang, C.F., 1987. Channel model of flow through fractured media, *Water Resour. Res.*, 23(3): 467–479.
- Tsang, C.-F. and Doughty, C., 2003. A particle-tracking approach to simulating transport in a complex fracture. *Water Resour. Res.*, 39(7): 1174, doi:10.1029/2002WR001614.
- Vickers, B.C., Neuman, S.P., Sully, M.J., and Evans, D.D., 1992. Reconstruction and geostatistical analysis of multiscale fracture apertures in a large block of welded tuff, *Geophys. Res. Lett.*, 19(10): 1029–1032.
- Wang, J.S.Y., Narasimhan, T.N., and Scholz, C.H. 1988. Aperture correlation of a fractal fracture, *J. Geophys. Res.*, 93(B3): 2216–2224.
- West, M.R., Kueper, B.H., and Novakowski, K.S., 2004. Analytical solutions for solute transport in fractured porous media using a finite width strip source. *Adv in Water Resour.*, 27: 1045–1059.
- West, A.C.F., Novakowski, K.S., and Gazor, S., 2005. Usefulness of core logging for the identification of conductive fractures in bedrock. *Water Resour. Res.*, 41: W03018, doi:10.1029/2004WR003582.
- Williams, H.R., Corkery, D., and Lorek, E.G., 1985. A study of joints and stress-relief buckles in Paleozoic rocks of the Niagara Peninsula, southern Ontario. *Can. Geotech. J.*, 22: 296–300.
- Zanini, L., Novakowski, K., Lapcevic, P., Bickerton, G., Voralek, J., and Talbot, C., 2000. Regional groundwater flow in the fractured carbonate aquifer underlying Smithville, Ontario, inferred from combined hydrogeological and geochemical measurements. *J. Groundwater*, 38(3): 350–360.
- Zeigler, T.W., 1976. *Determination of Rock Mass Permeability*, Waterways Experiment Station, Technical Report S-76-2, Vicksburg, MI.
- Zemanek, J., Caldwell, R.L., Glenn, E.E., Holcomb, S.V., Norton, L.J., and Strauss, A.J.D., 1969. The borehole televIEWER — A new logging concept for fracture location and casing inspection, *J. Pet. Technol.*, 21(12): 762–774.

Further Information

Englede et al. (1993) presents a thorough overview of the geological aspects of fracture mechanics and includes a comprehensive glossary of terms.

deMarsily (1986) is a general text in quantitative hydrogeology that discusses flow and transport in fractures in an easy to follow manner.

Kulander et al. (1990) discusses in detail core analysis and the identification and classification of breaks in rock core.

Doe et al. (1987) discusses various techniques and strategies for hydraulic testing in fractured rock.

Earlougher (1977) provides a comprehensive look at standard well testing methods in the petroleum industry that can be adapted to hydrogeological investigations.

Bear et al. (1993) provides a compilation of chapters investigating flow and contaminant transport in fractured rock and in addition to theoretical discussions has sections of flow and solute transport in fracture networks.

21

Groundwater Flow in Karstic Aquifers

21.1	Introduction.....	21-2
21.2	Karst Aquifers: The Conceptual Framework	21-2
	Groundwater Basins • Allogenic Recharge • Internal Runoff • Dispersed Infiltration and the Epikarst • Infiltration Pathways in the Vadose Zone • Subsurface Drainage Patterns • Groundwater Discharge: Karst Springs • The Geologic Boundary Conditions • The Conceptual Model	
21.3	Permeability and Flow Dynamics in Karst Aquifers ...	21-8
	Matrix and Fracture Flow • Conduit Permeability and Flow Dynamics • Sediment Fluxes and Transport Mechanisms • Caves and Conduits: Hydrologic Interpretation of Cave Patterns • Paleohydrology	
21.4	Chemistry of Carbonate Rock Dissolution	21-17
	Chemical Equilibrium • Chemical Parameters for Karst Aquifers • Dissolution Kinetics of Limestone and Dolomite	
21.5	Evolution of Karst Aquifers	21-24
	Life History of a Conduit • Initiation and Critical Thresholds • Enlargement, Maturity, and Decay	
21.6	Quantitative Hydrology of Karst Aquifers	21-27
	Water Budget • Base Flow/Runoff Relations • Tracer Studies • Well Tests and Water Table Mapping • Hydrographs of Karst Springs • Chemographs of Karst Springs	
21.7	Modeling of Karst Aquifers	21-37
	Evolutionary Models • Equivalent Porous Media Models • Pipe Flow Models • Coupled Continuum Pipe Flow Models • Input–Output Models	
21.8	Water Resource Issues in Karst Hydrology	21-39
	Water Supplies in Karst • Flood Flows in Karst: Sinkhole Flooding • Contaminants and Contaminant Transport in Karst Aquifers • Water Quality Monitoring	
21.9	Land Instability and Sinkhole Development.....	21-41
	Sinkholes as Land Use Hazards • Mechanisms of Sinkhole Development • Hydrologic Factors in Sinkhole Development • Sinkhole Remediation	
	References	21-43
	Further Information	21-47

William B. White
The Pennsylvania State University

21.1 Introduction

Karstic aquifers are those that contain dissolution-generated conduits that transport groundwater, often localized, fast-moving and in a turbulent regime. The ability of the conduit system to transport a sediment load, sometimes with associated contaminants, is an additional feature. Such aquifers are found mainly in limestone, dolomite and gypsum as the commonly occurring soluble rocks. In most aquifers there is a mismatch of many orders of magnitude between groundwater flow rates in the aquifer and stream flow rates on the land surface above. In karstic aquifers there is a continuum of flow rates such that the groundwater system in karstic aquifers takes on the characteristics of both surface water and groundwater. Further, because of the presence of sinking streams and large springs, there is a complex interplay between surface water and groundwater.

Karst regions display a characteristic suite of landforms. On the surface are sinking streams and dry channels linked by underground routes to large springs. There are closed depressions on many size scales and elaborately sculptured bedrock surfaces that may or may not be mantled with soil. Underground, there are cave systems which may or may not be accessible through entrances from the surface. Some caves are dry and abandoned except for drip water, some carry active streams, and some may lie below local base levels and be completely water-filled. Karst regions are often classified by their dominant landforms such as fluviokarst for mixed carbonate — clastic rock regions, cone-and-tower karst, pavement karst, sinkhole karst, and others. They are also sometimes classified by climate as in alpine karst, tropical karst, arid karst, and others. The emphasis of this chapter is on fluviokarst because this is the form of karst most commonly encountered by engineers and resource managers in the United States.

21.2 Karst Aquifers: The Conceptual Framework

21.2.1 Groundwater Basins

The physical framework for discussion of groundwater is the aquifer. Usually, an aquifer is a specific rock formation with a well-defined thickness but an indefinite lateral extent. One can speak of the “area” of an aquifer only in a rough sense. The physical framework for discussion of surface water is the drainage basin. Drainage basins have an arbitrary downstream limit determined by the choice of gauge point but once the gauge point is fixed, a drainage divide can be drawn and the basin has a well-defined area. Karst aquifers share the properties of both groundwater and surface water which lead to the concept of groundwater basins with well-defined boundaries which do not necessarily coincide with the boundaries of overlying surface water basins.

The essential components are sketched in Figure 21.1. The sketch is appropriate to the mixed carbonate/clastic rock terrain that makes up the fluviokarst of the eastern United States. There are three main components to the catchment system: (1) allogenic recharge, (2) internal runoff, and (3) dispersed (diffuse) infiltration. The groundwater basin arises when the aquifer acquires a well-developed conduit system. The conduit system provides a flow path of low hydraulic resistance. It also acts as a drain for matrix and fracture flow so that there is the subsurface equivalent of a drainage divide. However, unlike surface divides which are fixed by topography, groundwater basin divides may shift depending on groundwater levels. High-level spillovers between adjacent basins are common.

21.2.2 Allogenic Recharge

Streams rising on adjacent or overlying clastic rocks often sink underground near the clastic rock/carbonate rock contact when they flow onto the karst. These sinkpoints are called swallow holes or swallets. Each sinking stream derives water from its own small surface catchment with a well-defined area. The sink point acts as the gauge point. Sinking streams inject large quantities of water into the carbonate aquifer at single points. Allogenic recharge often has a different chemistry, is exceptionally vulnerable to contamination, and exhibits the rapid response to storm flow characteristic of surface runoff. The total area providing

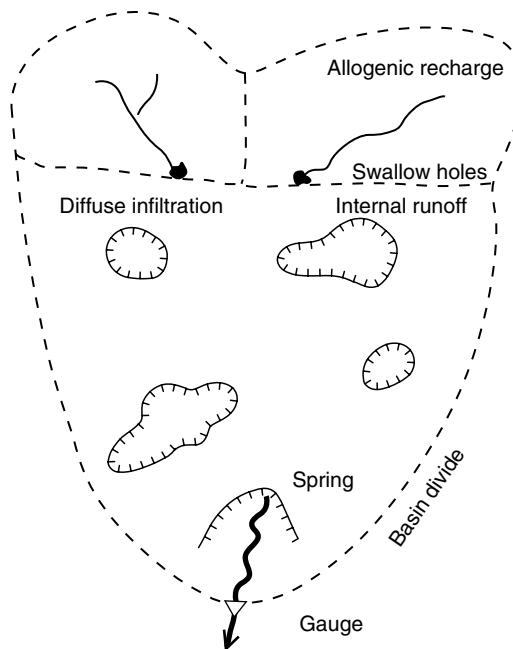


FIGURE 21.1 Map showing a groundwater basin with areas of allogenic recharge, internal runoff and diffuse (or dispersed) infiltration. The area of the basin is defined by the dashed line showing the basin divide which in turn is determined by the choice of gauge location.

allogenic recharge, A_a , is simply the sum of the areas of the individual sinking stream basins A_i :

$$A_a = \sum_i A_i \quad (21.1)$$

Swallets take many forms. In some cases the allogenic streams have downcut upstream from the swallet forming blind valleys so that all allogenic runoff feeds into the groundwater regardless of the magnitude of flow. In other cases, a dry surface channel continues downstream from the swallet so that during flood flows, only a portion of the flow enters the aquifer and the remainder remains as surface flow. Some streams lose a portion of their water through the channel but do not go completely dry even during base flow. In physical appearance, swallets range from open cave mouths with streams flowing into them to inconspicuous loosing reaches of stream channel.

In regions such as the Appalachian Plateaus, the Illinois Basin, and the Ozark Plateau, karstic carbonate aquifers underlie extensive plateaus capped with clastic rocks. Surface water catchments on the plateaus provide allogenic runoff that enters the carbonate aquifer from the top. These waters tend to be chemically aggressive and enlarge fractures into open shafts which in turn permit very rapid movement of water through the unsaturated zone.

21.2.3 Internal Runoff

Some karst terrains are so densely packed with closed depression that overland storm flow is entirely captured by the depressions and does not reach surface stream channels. Closed depressions — also known as sinkholes and dolines — vary widely in their ability to accept storm flow. In the eastern United States, most closed depressions are mantled with soil of widely varying thicknesses. Some have open drains at the bottom which carry storm flow directly into tributary channels of the conduit drainage system. In some, the storm flow must infiltrate through the soil mantle, a situation that can change rapidly if a soil

piping failure opens a throat in the bottom of the depression. Some depressions are mantled with clay-rich soils that act as effective seals thus producing perennial sinkhole ponds.

Calculation of the internal runoff contribution to the overall water balance is difficult. In principle, the collection area for internal runoff would be the sum of the areas of enclosed depressions. However, intense storms may be required to saturate the soils and produce enough overland flow to reach the bottom of the depression. Less intense storm flow infiltrates the soils on the walls of the depression and becomes part of the dispersed infiltration. The distinction between internal runoff and dispersed (or diffused) infiltration is based on transit time. Overland flow directly into open drains will have a different chemistry from water that permeates slowly through the soil. There is a continuum between the two recharge types.

21.2.4 Dispersed Infiltration and the Epikarst

The entire area of the karst groundwater basin receives rainfall. What does not drain into closed depressions by overland flow infiltrates slowly as the soil gradually becomes saturated. In karst terrain, the soils are strongly leached, with few carbonate rock fragments remaining in the soil. The C-horizon is usually absent; there is a sharp contact between the B-horizon and limestone or dolomite bedrock. The bedrock surface is often highly irregular with deep channels and trenches dissolved along fractures, sometimes known as “cutters” (or “grikes” in the UK). Between the fractures are irregular pinnacles of residual bedrock. Cavities may occur within the bedrock mass. The zone of regolith plus the cutter and pinnacle zone in the bedrock has a thickness of meters to tens of meters and is known collectively as the epikarst (Jones et al., 2003) (Figure 21.2).

The bulk bedrock usually has a very low permeability so that the infiltrating water can reach the subsurface only by moving laterally along the crevices and trenches until it reaches an open fracture or shaft descending into the aquifer below. Because the irregular cutters are soil-filled, flow rates are slow

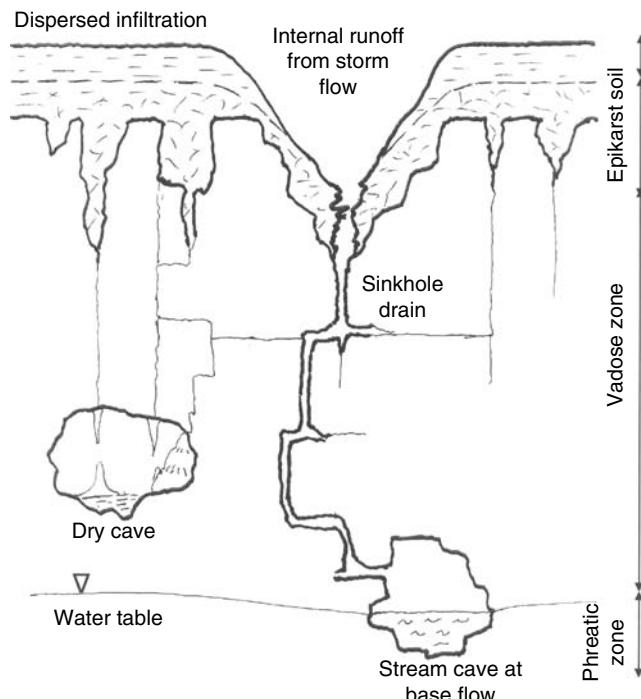


FIGURE 21.2 Cross-section of a karst surface showing a closed depression collecting internal runoff and the epikarst collecting dispersed infiltration.

and storm water may be stored for periods of days to weeks. The epikarst is the reservoir that supplies cave dripwater during drought periods.

21.2.5 Infiltration Pathways in the Vadose Zone

Movement of water through the vadose zone is highly anisotropic and often very rapid. Pathways for the vertical movement of water that has penetrated the epikarst include:

1. Slow percolation through the matrix porosity of the bedrock. This component is important in young limestones, such as those in Florida and the Bahamas, for example, that have not been compacted by deep burial or compressed by orogenic processes. Infiltration through the matrix is usually negligible in the dense Paleozoic limestones of eastern United States.
2. Vertical flow through fractures at the base of the epikarst.
3. Dissolutionally widened fractures and open throats that serve as sinkhole drains.
4. Larger shafts that carry the runoff from caprocked upland surfaces and discharge from overlying perched aquifer system.
5. Open channels that carry water from sinking streams.

Allogenic streams often sink into open cave systems at elevations well above regional base levels. The underground streams descend rapidly, often in open channels with a stair-step profile. In high-relief terrains, the profile may consist of a sequence of steep channels with interspersed waterfalls. These input channels are often large enough for human exploration and many have been mapped. The explorable allogenic stream inputs and some large closed depression drains have been classified as vadose caves. The internal relief varies from a few meters in low-relief terrains to more than 1000 m in high-relief terrains.

Water draining from upland areas underlain by clastic rock is injected into vertical pathways when it reaches the top of the carbonate aquifer. Because allogenic recharge tends to be highly aggressive, dissolution by the water moving along vertical pathways enlarges the initial fractures into shafts with diameters of meters to tens of meters and depths of tens to hundreds of meters. Some shafts have irregular shapes dictated by the shape of the fracture system that provided the initial pathway. With time, the shafts take the form of near-perfect cylinders with vertical walls that slice across beds of varying lithology (Brucker et al., 1972).

Drains from closed depressions have dimensions from centimeters to meters and are sometimes large enough to permit human exploration. Flow velocities are high and sufficient to transport sediment. Soil piping processes in closed depressions and above open fractures are an important source of sediment injected into the conduit system. The walls of closed depression drains are often jagged and corroded showing the highly aggressive chemistry of the water collected in the closed depressions and injected into the drain.

The Paleozoic carbonate aquifers of the eastern United states are highly fractured. The fractures provide effective pathways for infiltration of water into the subsurface. If the chemical reaction between the aggressive water of the epikarst with the carbonate bedrock goes to equilibrium, there may be relatively little dissolutional enlargement of these fractures. If the fractures are closely spaced, they provide a very important pathway for dispersed infiltration. In contrast to the open shafts and drains, because of the long residence time at the base of the epikarst, the fracture water is often near saturation with calcium carbonate.

21.2.6 Subsurface Drainage Patterns

The mixed groundwater/surface water character of karst aquifers is best revealed in the flow pattern in the subsurface. Instead of a uniform flow field mapped by isopotential contours, groundwater in karstic aquifers moves through an interconnected system of fractures, small solution channels, and large conduits. Travel times in the conduit system are often very short, hours to days per kilometer. The water may flow

as open channel underground streams or as pipe flow with varying degrees of pressure head. Gradients within the matrix and fracture flow portions of the aquifer are usually oriented toward the conduits.

Sinking surface streams for the most part inject water directly into the conduit system. Drains from closed depressions also feed into the conduit system. Dispersed infiltration recharges the matrix and fracture system which in turn discharges into the conduit system. Each of these subsystems has different storage capacities and travel times.

21.2.7 Groundwater Discharge: Karst Springs

Most groundwater discharge from karst aquifers is through high volume springs. Sometimes a groundwater basin discharges through a single spring and sometimes the water is distributed over multiple springs including wet-weather spillover springs. As more tracer data accumulates, groundwater distributed to multiple springs is being shown to be common. Diffuse discharge of groundwater in valley bottoms certainly occurs, but in well-developed karst aquifers, it constitutes only a small portion of the total discharge.

As an example, measurements show that 95% of the discharge from groundwater basins in the south-central Kentucky karst aquifer is through a few large springs (Hess and White, 1989). The springs along the Green River in south-central Kentucky can be divided into a small number of large regional springs draining groundwater basins of more than 100 km^2 , a somewhat larger number of valleys drain springs with catchments of a few tens of km^2 , and a large number of very small local springs are fed by small catchment areas near the river. Water balance shows, however, that the large springs account for most of the discharge.

21.2.8 The Geologic Boundary Conditions

The hydrologic components of the karst groundwater system must be superimposed on the geologic framework of the drainage basin because much of the detail in the pattern and development of the internal drainage system is guided and controlled by the local geology. The details of the geologic framework are highly site-specific but the main factors can be listed.

1. *Physiographic factors*
 - (a) Placement of the karst aquifer within the regional drainage basin
 - (b) Relief between recharge area and springs
2. *Structural factors*
 - (a) Folded vs. non-folded rocks
 - (b) Dip with respect to groundwater flow directions
 - (c) Faults
 - (d) Regional scale fracture systems
 - (e) Local pattern of joints and bedding plane partings
3. *Stratigraphic and lithologic factors*
 - (a) Thickness of carbonate or gypsum rocks
 - (b) Location of soluble rock sequences with respect to clastic rock sequences
 - (c) Limestone vs. dolomite vs. gypsum
 - (d) Lithologic characteristics: pure limestones, shaly limestones, marls
 - (e) Presence and location of confining layers — thin shales and sandstones — within the soluble rock sequence
 - (f) Confined vs. unconfined aquifer

Conduit development seeks out optimum flow paths through the available fractures and bedding plane partings. Conduits are deflected by relatively minor impermeable layers and by unfavorable lithologies. Fine-grained (micritic) pure limestone are favorable; shaly limestones, dolomitic limestones, and dolomites are unfavorable. Folds and faults guide possible flow paths by controlling the placement of favorable

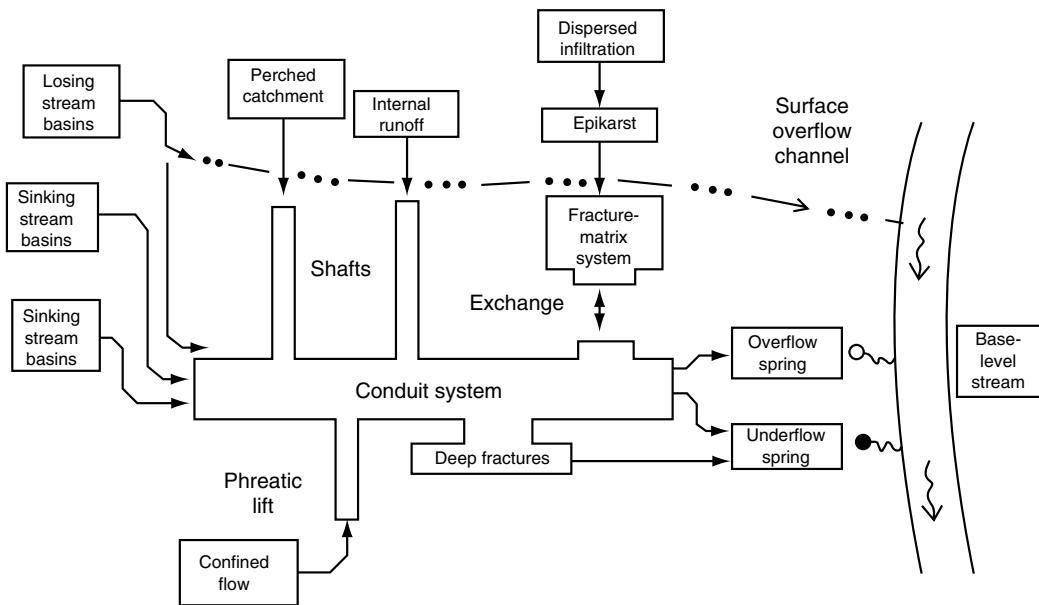


FIGURE 21.3 Conceptual model for the internal structure and flow paths in a karst aquifer. This is the current version of a sketch that has been drawn and re-drawn by the author and by J.A. Ray of the Kentucky Division of Water since its first publication (White, 1999).

rock lithologies. The final arrangement of internal flow paths represents a balance between hydraulic gradients and geologic constraints.

21.2.9 The Conceptual Model

The hydrogeologic concepts can be reduced to a conceptual model based on overall water balance and internal flow paths (White, 1999). A variety of these models can be drawn. One that applies to many fluviokarst groundwater basins in eastern United States is shown in Figure 21.3. The spring (or springs) where groundwater and spillover surface water are recombined into a single surface stream is the effective gauge point for the system.

The flow paths shown in the conceptual model are schematic and generalized. Allogenic recharge and internal runoff feed mainly into the conduit system. Throughput times are short. A well-developed conduit system exhibits a fast response to storm recharge. Rising and falling stage levels in the conduit system force the exchange of water into and out of the matrix and fracture systems. Dispersed infiltration recharges mainly the matrix and fracture system. There is a substantial time lag due to water stored in the epikarst. Overall throughput times for the dispersed infiltration are long, generally longer than the mean spacing between storms.

Not all of these components are present in all karst aquifers. There are carbonate aquifers to which there is no allogenic recharge and those in which the conduit system is absent. These have been called diffuse flow aquifers. In some aquifers, notably, aquifers in young limestones of Florida and the Caribbean, matrix flow is a dominant component of the groundwater system. In the dense Paleozoic limestones of the eastern United States, matrix flow rarely makes much of a contribution.

Beyond visualizing the internal workings of a karst aquifer, the principal value of the conceptual model is that it forms the framework for writing water budgets. Karst aquifers have the virtue that they provide many sites for flow measurements. Flow rates can be measured in sinking streams, in cave streams, and in springs. Measurable areas within the karst groundwater basin allow additional estimates of input. Such measurements provide the necessary data for input-output type of water balance calculations.

TABLE 21.1 The Component of the Triple Permeability Model for Karstic Aquifers

Permeability	Aperture	Travel Time	Flow Mechanism	Guiding Equation	Distribution
Matrix	μm to mm	Long	Darcian flow field. Laminar	$h_f = \frac{\eta v L}{\rho g (N d^2)}$	Continuous medium
Fracture	$10 \mu\text{m}$ to 10 mm	Intermediate	Cube law. Mostly laminar; may be nonlinear components	$\frac{Q}{h} = \frac{C}{f} b^3$	Localized but statistically distributed
Conduit	10 mm to 10 m and greater	Short	Darcy–Weisbach. Open channel and pipe flow. Turbulent	$h_f = \frac{f L v^2}{4 g r}$	Localized

21.3 Permeability and Flow Dynamics in Karst Aquifers

A central requirement for any theoretical description of groundwater flow is that the permeability remains a fixed-quantity characteristic of particular rock units. Discontinuities in permeability across geologic boundaries can be accommodated. Likewise, expressing permeability in tensor form accounts for the anisotropy of bedded rock units. Within these constraints, permeability remains scale-independent and time independent. Aquifer behavior can be described theoretically or modeled by computer programs with reasonable accuracy.

The permeability in karst aquifers is often described by the “triple porosity” or “triple permeability” model. It is the conduit permeability that sets karst aquifers apart from other water-bearing units (Table 21.1). As a consequence of the permeability distribution, the hydraulic conductivity varies by many orders of magnitude on different size scales (Sauter, 1991). Karst aquifers are not only anisotropic with respect to orientation but they are also anisotropic with respect to measurement scale and with respect to specific points in the aquifer where measurements are to be made.

21.3.1 Matrix and Fracture Flow

Ground water flow through the limestone or dolomite bedrock is not intrinsically different from groundwater flow in any other aquifer. The guiding equation is Darcy’s law.

$$Q = -AK \frac{\Delta h}{\Delta L} = -\frac{A \rho g N d^2}{\eta} \frac{\Delta h}{\Delta L} \quad (21.2)$$

Q is flow volume ($\text{m}^3 \text{ sec}^{-1}$), A is cross-sectional area of aquifer (m^2), K is the hydraulic conductivity (m sec^{-2}), ρ is density of water (kg m^{-3}), Nd^2 is permeability (m^2), and η is viscosity of water (Pa sec). $\Delta h/\Delta L$ is the hydraulic gradient.

The true matrix component of the flow must be based on the hydraulic conductivity of the unfractured bedrock. Such data are sparse. Most hydraulic conductivities are based on pump tests on wells and those data are dominated by the fracture flow. Intrinsic hydraulic conductivities must be determined from laboratory measurements on unfractured core samples. Estimates for the Mississippian limestone in which Mammoth Cave is formed and for a Silurian dolomite in Ontario are $K = 2 \times 10^{-11}$ and $K = 10^{-10} \text{ m sec}^{-1}$, respectively (Worthington, 1999). These may be contrasted with $K = 10^{-8} \text{ m sec}^{-1}$ for the Cretaceous Edwards Aquifer in Texas (Worthington et al., 2002) and $3.65 \times 10^{-6} \text{ m sec}^{-1}$ as a mean value for the Floridan Aquifer (Budd and Vacher, 2002).

What is clear from the observed hydraulic conductivities is that matrix flow is a negligible part of the overall flow field in the compacted Paleozoic limestones. Matrix flow is important in the Floridan aquifer and many other young, undisturbed limestones in the Yucatan and in the Caribbean islands.

The idealized model for fracture flow assumes a fracture with plane parallel walls and a uniform aperture. For this ideal geometry, the flow is described by the cubic law which can be derived theoretically from the fundamental Navier–Stokes equations:

$$Q = -\frac{W \rho g b^3}{12\eta} \frac{\Delta h}{\Delta L} \quad (21.3)$$

In Equation (21.3) W is the total width of fractures (m) and b is the full aperture of the fracture (m). Other symbols are as defined for Equation 21.2.

It has long been recognized that real fractures do not have uniform apertures and that the walls are not parallel. An empirical approach (Witherspoon et al., 1980) was to compact the constants of Equation (21.3) into a single constant, C , and add an empirical fraction factor, f_f .

$$Q = \frac{C}{f_f} b^3 \frac{\Delta h}{\Delta L} \quad (21.4)$$

Good agreement was obtained with laboratory measurements with values of f_f ranging from 1.04 to 1.64.

Real carbonate aquifers contain a complex suite of fractures. These are typically rough-walled, of varying aperture, and which may or may not be interconnected. Describing the flow through such fracture networks is the object of much current research (Brush and Thomson, 2003; Konzuk and Kueper, 2004). See Chapter 20.

21.3.2 Conduit Permeability and Flow Dynamics

For reasons to be discussed later, the aperture marking the transition from fracture permeability to conduit permeability is taken to be 1 cm. Within a mature karst aquifer is a network of solutionally widened fractures and bedding plane partings less than 1 cm in aperture, widened fractures and small solution channels with dimensions of centimeters to tens of centimeters, and conduits with dimensions of meters to tens of meters. All of these are interconnected in complex ways, sometimes in an ordered branchwork pattern as an underground analogy of a surface stream, but more often with multiple interconnections of closed loops. The conduits have very low hydraulic resistance and can carry large volumes of water at high velocities. Knowledge of the conduit system is mainly obtained by tracer tests and by direct observation and survey. Given 0.5 m as a practical lower limit for human exploration, there is a very large portion of the conduit permeability in the scale range from 0.01 to 0.5 m which cannot be directly probed by any presently available technique.

Because hydraulic conductivity has units of velocity, there is a temptation to equate travel time to an effective “hydraulic conductivity.” This is erroneous. Use of hydraulic conductivity implies a linear relationship between hydraulic head and discharge and a continuous medium. Most flow in conduits is turbulent and the flow is localized. Further, the travel time represents an average. The actual flow path may contain reaches of high-velocity flow interspersed with reaches where the water is ponded.

To a certain extent, groundwater flow in conduits can be treated as flow in pipes. However, conduit geometry is highly variable with cross-sections ranging from uniform elliptical shapes, to high and narrow fissures, to irregular shapes partly filled with rock fragments. Most conduits contain a bed of clastic sediments on their floors and some of these are armored with cobble and boulder-size material in the manner of surface streams. Water may flow in conduits in an open-channel regime subject only to gravitational gradients, or as pipe flow under substantial pressure heads. Both regimes frequently occur in different reaches of the same conduit. Rising and falling water levels in response to storm recharge can shift the regime from open channel flow to pipe flow and back again. Passages that flood to the ceiling during storms are well known to cave explorers.

Flow velocities in conduits become sufficiently large to drive the system into a turbulent regime. The flow regime in closed pipes is described by a Reynolds number defined as:

$$N_R = \frac{\rho v R}{\eta} \quad (21.5)$$

Again, ρ and η are the fluid density and viscosity, respectively and v is the flow velocity. R is the hydraulic radius. Although the hydraulic radius for circular pipes is often defined as the pipe diameter, it is more useful to define the hydraulic radius of irregular karst conduits as the cross-sectional area divided by the perimeter. With this definition $R = 2r/4$ ($2r$ = pipe diameter) with a corresponding scaling of the numerical value of the Reynolds number for circular conduits. If Equation 21.5 is written in consistent units, the Reynolds number is dimensionless. At low Reynolds numbers, the flow regime will be laminar; at high Reynolds numbers it will be turbulent. The transition from laminar to turbulent flow occurs over a range of Reynolds numbers. The onset depends on the wall roughness of the conduit and will be in the range of $N_R = 50$ to 500 . At Reynolds numbers greater than 500 (with the definition of hydraulic radius given above), fully developed turbulent flow will occur even in smooth-walled conduits.

For a circular conduit of radius, r , containing water flowing in a laminar regime, the flow volume is described by the Hagen–Poiseuille equation:

$$Q = \frac{\pi \rho g r^4}{8\eta} \frac{\Delta h}{\Delta L} \quad (21.6)$$

There is no corresponding equation for turbulent flow derived from first principles. Instead, the usual starting point is an empirical relation, the Darcy–Weisbach equation. The Darcy–Weisbach equation is shown as a head loss in Table 21.1. As a calculation of flow volume in a circular conduit, it takes the form:

$$Q = 2\pi \left(\frac{g}{f}\right)^{1/2} r^{5/2} \left(\frac{\Delta h}{\Delta L}\right)^{1/2} \quad (21.7)$$

The parameter, f , is a dimensionless friction factor. Its determination is one of the serious problems in modeling conduit flow in karst aquifers. If the flow is laminar, the friction factor is exactly determined by the Reynolds number:

$$f = \frac{64}{N_R} \quad (21.8)$$

In the limiting case of turbulent flow in a smooth-walled conduit, unlikely to be found in a karst aquifer, the friction factor is given by the Prandl–Von Karman equation:

$$\frac{1}{\sqrt{f}} = 2 \log(N_R \sqrt{f}) - 0.8 \quad (21.9)$$

Determination of the Darcy–Weisbach friction factor for real karst conduits has taken two approaches. For uniform conduits, the friction factor can be calculated from the wall roughness. Turbulent flow over a smooth wall produces a laminar sublayer separating the mass of the fluid from the rock wall. The wall roughness of rock conduits depends on the irregularities on the conduit wall and on the lithologic characteristics of the bedrock. The thickness of the laminar sublayer, δ is determined by the dimensionless equation:

$$\frac{\delta}{2r} = \frac{32.8}{N_R \sqrt{f}} \quad (21.10)$$

The relief of surface irregularities, e , is given in the same units as the conduit dimensions. A rough surface is one where the surface roughness is greater than the thickness of the laminar sublayer. For very rough

TABLE 21.2 Estimates of the Darcy–Weisbach Friction Factor

Location	From Discharge	From Roughness	Reference
Mendips (UK)	24–340	—	[1]
Castleguard (Canada)	0.87–2.31	0.33–0.90	[2]
Morecombe Bay (UK)	—	0.077	[3]
Glomdalsvatn (Norway)	0.116	0.039	[4]
Turnhole (KY)	27	—	[5]
Friars Hole (WV)	46–74	—	[5]
Holloch (Switzerland)	—	0.322	[6]
Maligne Basin (Canada)	130	—	[7]

References for Table 21.2: [1] Atkinson (1977); [2] Atkinson et al. (1983); [3] Gale (1984); [4] Lauritzen et al. (1985); [5] Worthington (1991); [6] Jeannin (2001); [7] Smart (1988).

surfaces the friction factor is given by:

$$\frac{1}{\sqrt{f}} = 2 \log \frac{2r}{e} + 1.14 \quad (21.11)$$

For the friction factor to be calculated from wall roughness, the conduit must be uniform. Most karst conduits are not uniform. An alternative approach is to monitor a karst drainage basin and measure all parameters in the Darcy–Weisbach equation except the friction factor. Then calculate the effective friction factor for the entire conduit from the measurements. Table 21.2 is a summary of Darcy–Weisbach friction factors determined either from wall roughness or from drainage basin calculations. The resulting numbers vary by more than two orders of magnitude, with the wall-roughness friction factor always smaller than those calculated from the entire basin.

Open channel flow is described by the Manning equation:

$$v = \frac{M}{n} R^{2/3} S^{1/2} \quad (21.12)$$

In the Manning equation, v is the mean channel velocity, R is a hydraulic radius, and S is the channel slope. The constant, n (known as “Manning’s n ”) is a roughness factor. It has been the custom to let n have the same numerical values regardless of units. The parameter, M , adjusts the units: $M = 1$ for velocity in m sec^{-1} and hydraulic radius in meters. $M = 1.486$ for velocity in ft/sec and hydraulic radius in feet. The hydraulic radius of an irregular channel is defined as the cross-sectional area divided by the wetted perimeter. Values for Manning’s n have been determined for a range of surface stream bed materials (Table 21.3). These are the best available approximations for underground stream channels since no direct experimental determinations of the roughness characteristics of karst conduits seem to have been made.

Water may flow in open channels in either laminar or turbulent regimes, although most underground free-surface streams have dimensions and velocities that place them in the turbulent regime. Open channel flows are also subject to a balance between inertial forces and gravity forces described by a dimensionless Froude number:

$$N_F = \frac{v}{\sqrt{gD}} \quad (21.13)$$

D is the hydraulic depth, equal to the actual water depth in a rectangular channel. Flows with $N_F < 1$ are termed subcritical; flows with $N_F > 1$ are termed supercritical. The transition from supercritical to subcritical flow is abrupt, with considerable release of energy producing a hydraulic jump.

Most underground streams flow in a subcritical regime, but supercritical flows occur in steep gradient vadose streams carrying water from allogenic catchments down to base level. Supercritical flows also occur in the shafts and drains that carry internal runoff through the unsaturated zone down to the water table.

TABLE 21.3 Manning's *n* for Open Channels

Bed Material	<i>n</i>
Smooth concrete	0.010
Good ashlar masonry	0.013
Rubble masonry	0.017
Firm gravel	0.020
Clay	0.026
Clay and silt	0.046
Sand and clay	0.030
Sand and gravel	0.042
Gravel over rock	0.027
Coarse sand with some gravel	0.049
Gravel, rock, and boulders	0.055
Large angular boulders	0.079

Source: Data from Vennard, J. K. (1962) *Elementary Fluid Mechanics*, Wiley, New York and Barnes, H. H. (1967) *Roughness Characteristics of Natural Channels*, U.S. Geological Survey Water Supply Paper 1849, Washington, D.C.

21.3.3 Sediment Fluxes and Transport Mechanisms

Consider the drainage basin illustrated in Figure 21.1. Various tributary surface streams sink into blind valleys. There is no overland flow path for these streams. All surface water moves into the groundwater system to appear eventually at the spring where it returns to a surface route. These tributary surface streams also carry a clastic load. Some clastic particles, mostly clay and some silt will be moved in suspension during storm flow. Other material, particularly the coarser material, will be dragged along the bed as bedload during periods of high flow. Both suspended and bedload sediments are transported into the underground system. They eventually reappear at the spring although there may be a substantial time delay. Clastic sediments are held in temporary storage in the conduit system until the occurrence of storm flows of sufficient magnitude to move the material through the system. If the clastic sediment did not move through the system, the conduit system would eventually clog and force the sinking streams back onto surface routes. However, the transport is episodic with sediments moving only during periods of high discharge. Sampling of the sediment load is possible by examining the sediments in caves which, in the active passages, can be taken as sediment in temporary storage, waiting for the next storm episode. Sediment transport is now considered an important part of karst hydrology and many studies have been undertaken (Sasowsky and Mylroie, 2004).

A flow sheet for the transport of clastic sediment through the conduit system is shown in Figure 21.4. In addition to the sediment carried into karst aquifers by sinking streams, there is also soil-washdown due to gradual flushing of soils from the epikarst, and plug injections of sediment from soil-piping failures in sinkholes. A fourth component is the insoluble residue from the dissolution of the bedrock to form the conduit system. The clastic sediments that are flushing into the conduit system are reworked considerably as they are transported through the system.

As sediments are transported in the conduit system mainly during storm events, direct investigation of the transport processes is difficult. Suspended sediment can be captured at the springs where the concentration can be related to the discharge hydrograph. Observations on springs give information mainly on the very small particle end of the sediment size scale. Many karst springs are perennially turbid to various degrees, reflecting the concentration of suspended material. Particles in the colloid size range and also larger particles of organic material remain in suspension indefinitely. Larger particles settle out and move only during storm flow when some springs become visually muddy. Transport of even larger particles, in the gravel to cobble range, only occurs during extreme storms which are infrequent and also when measurements are not likely to be possible.

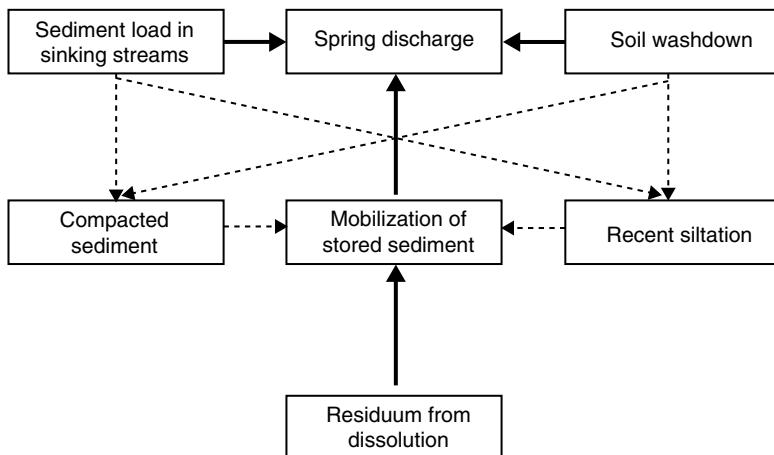


FIGURE 21.4 Flow chart for the sources of clastic sediment in a karst aquifer. A distinction is made between compacted sediments which tend to be immobile in the conduit system except during extreme storms and recent siltation which is easily remobilized and appears at the spring during very moderate storms.

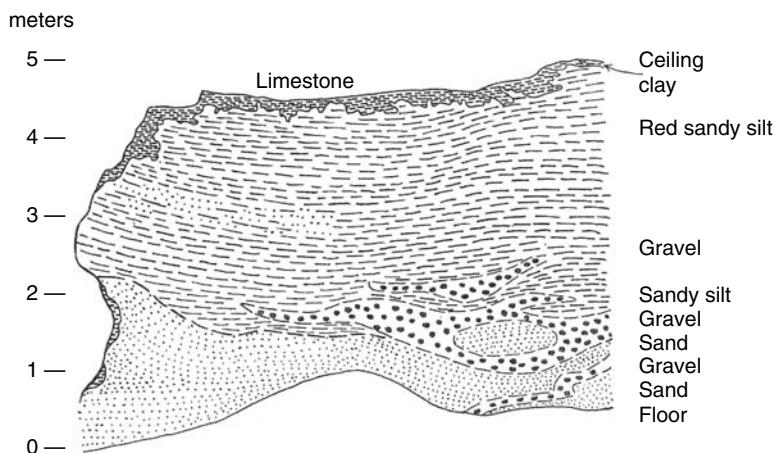


FIGURE 21.5 The clastic in-filling of Dismal Valley, an occluded master conduit in the Salts Cave section of Mammoth Cave, Kentucky. The bulk of the deposit is the channel facies with a layer of slackwater facies (clay) on top. Drawing from Davies, W.E. and Chao, E.C.T. (1958) Report on sediments in Mammoth Cave, Kentucky. *Administrative Report to National Park Service by U.S. Geological Survey*, 117 pp.

A more representative sampling of the total sediment flux can be obtained by examining sediment piles in caves. In caves that are part of active conduit systems, these deposits represent the sediment flux, although movement might be on very long time scales. Sediment piles in dry caves are samples of the material that was moving through the system when the cave passage was abandoned and which may represent flow conditions that are no longer present in the drainage basin. Sorting and reworking of the sediment in the conduit system results in recognizable lithofacies that can be distinguished by particle size and by sorting (Bosch and White, 2004). Two of these are illustrated (Figure 21.5) from Mammoth Cave, Kentucky.

Channel facies are those sediments likely to be moved as bedload during storms of moderate intensity. They typically consist of interbedded sand, silt and sometimes gravel, often with a rough sorting into distinct beds as shown in Figure 21.5. However, these beds cannot be traced continuously for distances

of more than a few tens of meters along the passageway. Bedload transport consists of material being dragged along the bed by the shearing action of water flowing above the bed. The shear is related to the velocity of the moving water by Newton's stress law:

$$\tau = \frac{1}{8} \rho f v^2 \quad (21.14)$$

The symbols are τ = boundary shear, N m^{-2} ; ρ = fluid density, kg m^{-3} ; f = dimensionless friction factor; v = fluid velocity at the boundary, m sec^{-1} . The boundary shear necessary for bedload movement has been extensively investigated for surface channels (Vanoni, 1975) and is related to sediment particle size by the regression equation:

$$\tau_c = 0.067 D_{50}^{1.08} \quad (21.15)$$

D_{50} is the mean particle diameter in mm. The presence of channel facies sediments in long horizontal conduits far from any possible sediment inlets is evidence that the flow velocity in these conduits must be in the range of tenths of a meter per second or the sediments would not be where they are observed.

The uppermost bed in many sediment deposits is a layer, typically a few centimeters thick, of fine-grained laminated clay. Note that this layer of clay, termed the slackwater facies in Figure 21.5 completely fills the passage between the top of the channel facies and the limestone ceiling. The slackwater facies provides a sample of the suspended load. During storm flow, base level conduits are frequently flooded to the ceiling. All recesses are filled with muddy water. This water is stagnant during the time in which the passage is flooded, and suspended sediment settles, forming a layer of clay. Successive floods add additional layers and thus is formed the bed of laminated clay.

The sediments in active stream channels are often well-winned so that the stream bed has a gravel or cobble armoring much like many surface streams. Sometimes, very coarse cobbles occur, tens of centimeters in diameter or larger. These can move as bedload only during exceptional floods whereas the smaller particles are stripped away during more frequent, but less intense storms. These stream channel deposits have been referred to as thalweg facies.

There is evidence in a few caves for a diamicton facies. These are masses of completely unsorted and unstratified material, apparently thrust into the conduit as a debris flow traveling in suspension at high velocity. The term and the first observations came from the high relief caves of the New Guinea Highlands (Gillieson, 1986). Since then similar deposits have been found in other caves, some with only modest relief.

21.3.4 Caves and Conduits: Hydrologic Interpretation of Cave Patterns

It is important to distinguish between caves and conduits. Conduits are solutionally formed pathways on many size scales that serve to transmit water from recharge points to discharge points. Some portions of the conduit system are below base level and are continuously water-filled. Some portions lie between the elevations of base flows and flood flows and are therefore used intermittently depending on the groundwater stage. Still other portions are abandoned as base levels are lowered and groundwater levels fall below the elevations of the conduits.

Caves are those fragments of the conduit system that are accessible to human exploration. Cave explorers traditionally include all accessible passages as part of the same cave. Thus a cave map may show fragments of several conduits interconnected by collapse or secondary dissolution. The cave map may also include conduit segments of greatly different ages ranging from abandoned high-level passages that may date from the Pliocene to vertical shafts that are apart of the contemporary groundwater system. Cave maps, in brief, are composites of all parts of the conduit system that are interconnected and which can be accessed by cave exploration. Although they provide valuable data on the conduit system, they must be interpreted with some care.

Although cave maps reveal a great diversity of patterns, some common features can be recognized. The important variables are the local geologic structure that determines many of the details of the cave

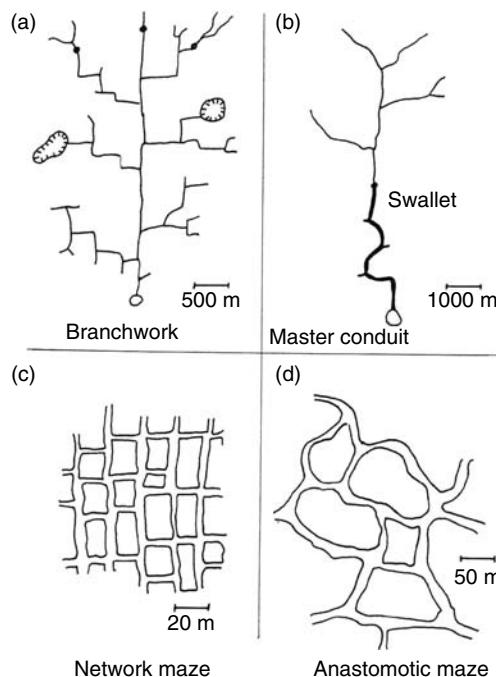


FIGURE 21.6 Patterns of conduit drainage systems: (a) Branchwork pattern with most tributaries originating as sinkhole drains. (b) Master conduit with surface streams making up most of the tributaries. (c) Network maze. (d) Anastomotic maze.

pattern, the overall hydraulic gradient of the drainage basin, and the relative contributions of alloegenic, internal runoff, and dispersed infiltration recharge (Palmer, 1991). Some of the possibilities are shown in Figure 21.6.

The conduit systems within aquifers dominated by recharge through internal runoff often exhibit branchwork patterns as small tributaries draining individual closed depressions merge to form larger drainage channels which ultimately lead to a spring. The smallest tributaries of the systems are often too small for human exploration. Branchwork patterns require sufficient gradient to allow a free-flowing drainage to develop. There may appear closed loops and spillover routes used during periods of high discharge. In detail, the branching tributaries are guided by joints or bedding plane partings.

If the aquifer is dominated by alloegenic recharge, the underground drainage may be through a single large conduit with few tributaries. In the limit, a large surface catchment could have a short segment of underground route that would be almost completely decoupled from the groundwater system. These systems are most like underground rivers and least like aquifers. Large conduits can be nearly linear, can be angulate following structural features in the bedrock or can be sinuous.

Caves frequently occur with maze patterns consisting of many closed loops along the same flow route. Maze caves may take the form of anastomotic mazes or network mazes depending of whether the fracture permeability is dominated by bedding plane partings or by joints. Both require that the dissolution process proceed along all possible pathways without the concentration of flow along a single pathway that leads to branchwork and single conduit caves. Anastomotic mazes generally have their greatest cross section developed along bedding planes. Freedom for water to select optimum pathways along the bedding plane parting leads to the gently curving loops on many size scales. Network mazes generally have their greatest cross section developed along joints so that the pattern is constrained by the joint pattern to produce the characteristic “city block” pattern. Anastomotic mazes generally occur in low gradient systems in nearly flat-bedded rock where many percolation paths would be possible. Network mazes are sometimes related

to dispersed infiltration through an overlying permeable caprock into a jointed limestone mass below the water table (Palmer, 1975). Other network mazes are the result of backflooding from surface streams, a form of bank storage. In a different category are floodwater mazes that occur in very high gradient conduit systems with large contributions of alloegenic recharge where high hydraulic heads from flood flows can force water along all possible joints and fractures.

To a certain extent, branchwork patterns are the subsurface analogs of surface tributary stream patterns. Single conduit patterns are the subsurface analogs of high-order surface streams. Anastomotic maze patterns may be taken as the subsurface equivalent of braided rivers, but most generally, maze patterns are the subsurface equivalent of swamps.

21.3.5 Paleohydrology

Unlike deepening surface valleys and their drainage systems which usually destroy their past as the valleys downcut, caves provide useful records of the history of the conduit drainage system. Vertical sequences of abandoned conduit fragments (tiered caves) may reveal changing drainage patterns and changing discharges. To be useful, the paleohydrologic record must be tied to the time sequence of the evolving surface drainage basin. U/Th isotope dating of calcite speleothems, useful in the age range back to 0.5 Myr, can place some limits on the ages of various passage levels (Richards and Dorale, 2003). Magnetic reversals preserved in cave sediments provide absolute time markers when they can be found and when the correct reversal can be identified (Schmidt, 1982). Most relevant to paleohydrology is the dating of clastic sediment by cosmogenic isotopes (Granger et al., 2001; Anthony and Granger, 2004). SiO_2 in the form of quartz sand and pebbles exposed on the land surface is subject to the bombardment by cosmic rays. The bombardment converts minute amounts of ^{28}Si to ^{26}Al and ^{16}O to ^{10}Be , both radioactive isotopes. When the irradiated quartz is swept into a cave and deposited there, the quartz is shielded from cosmic rays. The isotopes begin to decay but at different rates. Measurement, by means of an accelerator mass spectrometer, of the $^{26}\text{Al}/^{10}\text{Be}$ ratio, permits the calculation of the time at which the sediment was transported into the cave. This is only the minimum age for the cave passage but is likely to be close because most sediment deposition takes place when the conduit is an active part of the groundwater system.

Cave walls often exhibit a pattern of scallops (Figure 21.7) which record both the flow direction and the paleoflow velocity. The scallop's length is inversely related to velocity through a Reynolds number

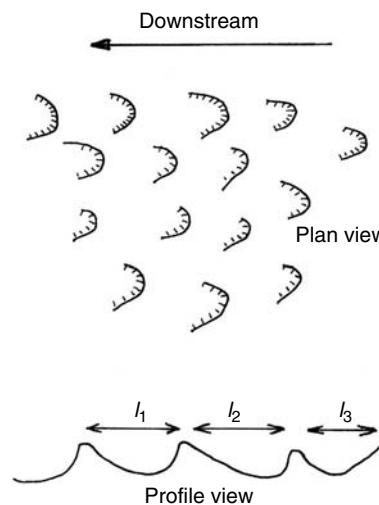


FIGURE 21.7 Plan and profile of a scalloped cave wall. Characteristic lengths of the scallops are measured along the flow direction.

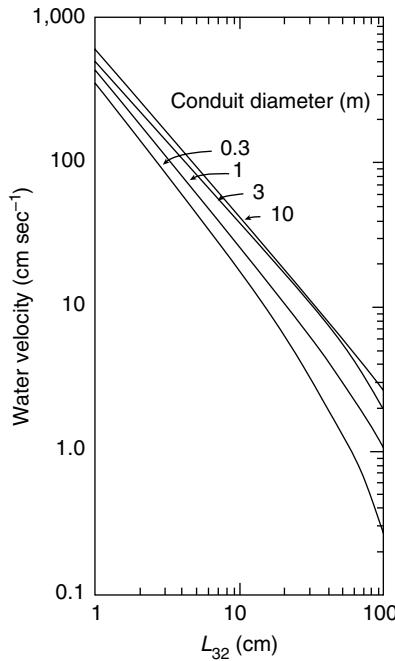


FIGURE 21.8 Calculated water velocity as a function of mean scallop length. L_{32} = Sauter mean. Curves based on the equations of Curl (1974).

(Curl, 1974):

$$\frac{\rho v L_{32}}{\eta} = N_R = 22,500 \quad (21.16)$$

The parameter L_{32} is the Sauter mean of a population of measured scallop lengths:

$$L_{32} = \frac{\sum_i \ell_i^3}{\sum_i \ell_i^2} \quad (21.17)$$

A detailed analysis of the transport processes shows a dependence on conduit size (Figure 21.8). Measurement of the Sauter mean on a population of scallops gives paleoflow velocity, which, combined with conduit cross-sectional area gives a paleodischarge. With some assumptions the paleodischarge allows an estimate of the paleocatchment areas. Calculations of hydraulic parameters from scallops in both paleoconduit and active conduits have produced reasonable agreement with the parameters determined by other measurements (Gale, 1984; Lauritzen et al., 1985; White and Deike, 1989).

21.4 Chemistry of Carbonate Rock Dissolution

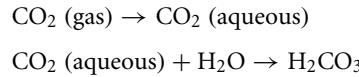
21.4.1 Chemical Equilibrium

The primary process in karstic weathering and in the development of fracture and conduit permeability in karst aquifers is the dissolution of carbonate rocks in weak solutions of carbonic acid. Carbonate chemistry has been extensively investigated, and more detail may be found in the aqueous geochemistry textbooks of Drever (1997) and Langmuir (1997).

Carbon dioxide occurs in the atmosphere at a volume fraction of 3.75×10^{-4} (2003 values). This reservoir provides a constant atmospheric background that CO₂ partial pressures in karstic waters rarely fall below. Much greater concentrations of CO₂ are generated in the soil due to respiration from roots,

from decay of organic matter, and from the action of microorganisms. CO₂ generation in the soil is a sensitive function of temperature and, in temperate and subpolar climates, exhibits a seasonal periodicity. The volume fraction of CO₂ in the soil varies from 0.01 to 0.1 with excursions to both greater and smaller values. Some of the soil CO₂ diffuses upward to be lost to the atmosphere, while some dissolves in infiltrating groundwater.

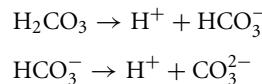
Carbonic acid is formed from gaseous CO₂ by a two-step process:



Molecular carbonic acid is a minor species and comprises only about 0.17% of the dissolved CO₂. Nonetheless, it is convenient to consider both reactions together and write the equilibrium constant as:

$$\frac{a_{\text{H}_2\text{CO}_3}}{P_{\text{CO}_2}} = K_{\text{CO}_2} \quad (21.18)$$

The symbol “*a*” in this and the following equations is the thermodynamic activity. H₂CO₃ ionizes:

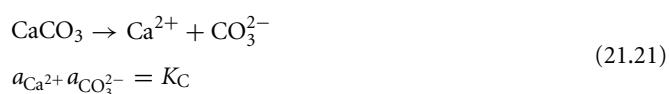


The equilibrium expressions are:

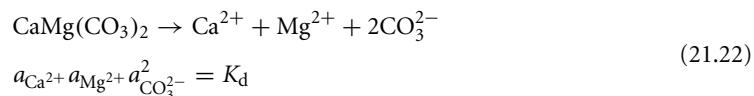
$$\frac{a_{\text{H}^+} a_{\text{HCO}_3^-}}{a_{\text{H}_2\text{CO}_3}} = K_1 \quad (21.19)$$

$$\frac{a_{\text{H}^+} a_{\text{CO}_3^{2-}}}{a_{\text{HCO}_3^-}} = K_2 \quad (21.20)$$

Calcite (limestone) is only slightly soluble in pure water:



Solving Equation 21.21 gives the concentration of dissolved calcite as 6 to 7 mg/L, less than the solubility of quartz. The difference is that the solubilities of carbonate minerals are highly sensitive to the presence of weak acids whereas the solubility of quartz is not. The dissolution of dolomite can be described by a similar reaction:



The equilibrium constants for Equations 21.18 through 21.22 have been refined and evaluated (Table 21.4).

In overall terms, the dissolution of limestone and dolomite in carbon dioxide-rich waters is described by the reactions:

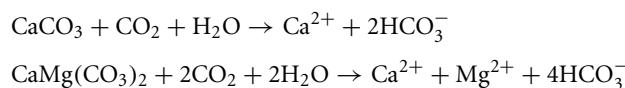


TABLE 21.4 Equilibrium Constants for Carbonate Reactions

T (°C)	pK _w	pK _{CO₂}	pK ₁	pK ₂	pK _{CaHCO₃⁺}	pK _{CaCO₃}	pK _c	pK _a	pK _d
0.0	14.938	1.108	6.579	10.629	-0.820	-3.134	8.381	8.218	16.56
5.0	14.727	1.192	6.516	10.554	-0.901	-3.127	8.394	8.235	16.63
10.0	14.528	1.269	6.463	10.488	-0.968	-3.134	8.410	8.255	16.71
15.0	14.340	1.341	6.419	10.428	-1.023	-3.154	8.430	8.279	16.79
20.0	14.163	1.407	6.382	10.376	-1.069	-3.185	8.453	8.306	16.89
25.0	13.995	1.468	6.352	10.329	-1.106	-3.224	8.480	8.336	17.00
30.0	13.836	1.524	6.328	10.288	-1.135	-3.271	8.510	8.370	17.12
35.0	13.685	1.577	6.310	10.252	-1.159	-3.325	8.543	8.406	17.25
40.0	13.542	1.625	6.297	10.222	-1.179	-3.383	8.580	8.446	17.39

pK = $-\log K$. K_w taken from *Handbook of Chemistry and Physics*, 84th edn., pp. 8–87 (2003). K_d taken from Langmuir (1971). All other K 's calculated from fitting equations of Plummer and Busenberg (1982).

These equations leave aside possible reactions with organic acids and also leave aside reactions involving sulfuric acid which is known to be an important mechanism in the development of some cave systems. A discussion of the sulfuric acid mechanism for the formation of Carlsbad Caverns and other caves of the Guadalupe Mountains of New Mexico may be found in Hill (1990).

To calculate the solubilities of the carbonate minerals requires that all of these reactions be simultaneously in equilibrium. Because carbonic acid and its ionization products are involved, one further reaction must be considered, the ionization of water:



Considering the case of pure calcite, the simplest case, crystalline calcite would be in equilibrium with an aqueous phase and with a gas phase containing a variable concentration of CO₂. The system would require the determination of seven variables — the concentrations of Ca²⁺, H₂CO₃, HCO₃⁻, CO₃²⁻, H⁺, and OH⁻ plus the partial pressure of CO₂. Seven unknowns require seven independent equations. The equilibrium expressions 21.18, 21.19, 21.20, 21.21, and 21.23 provide five of these. The requirement that positive and negative electric charges must balance gives a sixth relationship. For the determination of the seventh constraint on the system, three cases are of interest:

1. Available CO₂ is a fixed quantity (closed system)
2. The CO₂ pressure is fixed by the environment (open system)
3. The pH (hydrogen ion activity) is fixed by the environment

Closed system conditions occur only rarely in karst. Case 1 is relevant but complicated to describe. Langmuir (1997) discusses this case in some detail.

If pH is fixed by some external buffering system (Case 3) such as might be relevant to the analysis of limestone barriers for the remediation of acid mine drainage, the individual equations can be solved and no elaborate calculations are necessary.

Case 2 is of most interest in karst hydrology because the CO₂ partial pressure is usually fixed by the rate of CO₂ generation in the soil. The calculations involving the various equilibrium expressions, the charge balance requirement, and the constraint that $P_{\text{CO}_2} = \text{constant}$ yield the equilibrium concentration of dissolved Ca²⁺ (as molal concentration):

$$m_{\text{Ca}^{2+}} = P_{\text{CO}_2}^{1/3} \left(\frac{K_1 K_c K_{\text{CO}_2}}{4 K_2 \gamma_{\text{Ca}^{2+}} \gamma_{\text{HCO}_3^-}^2} \right)^{1/3} \quad (21.24)$$

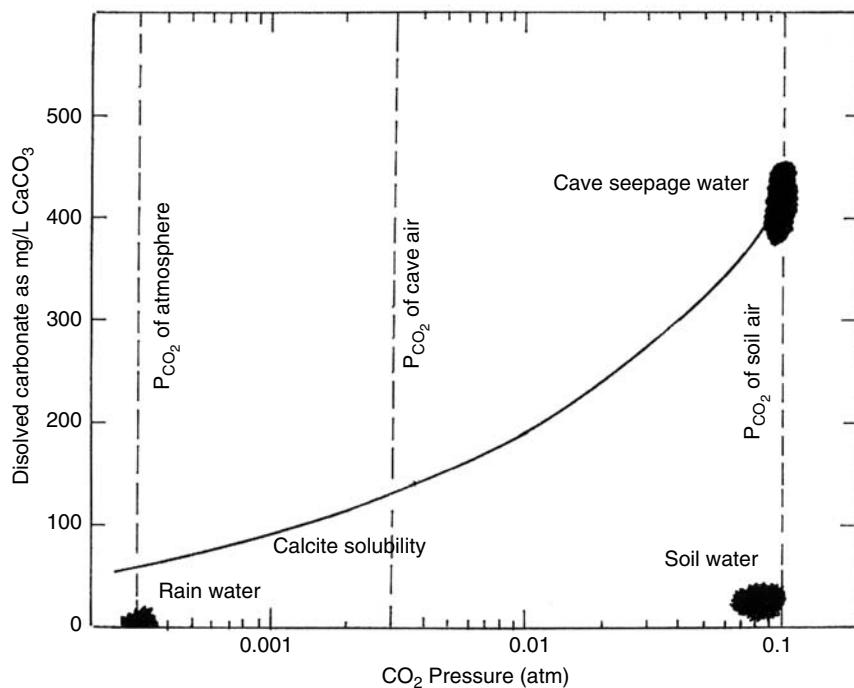


FIGURE 21.9 Solubility curve for calcite at 10°C as a function of CO₂ partial pressure. The chemical composition of dispersed infiltration water is shown at three locations along its flow path.

If the groundwater is at equilibrium with calcite at a given CO₂ partial pressure, the pH is also fixed.

$$\text{pH} = -\frac{2}{3} \log P_{\text{CO}_2} - \frac{1}{3} \log \left(\frac{K_1^2 K_2 K_{\text{CO}_2} \gamma_{\text{Ca}^{2+}}}{2 K_c \gamma_{\text{HCO}_3^-}} \right) \quad (21.25)$$

The symbol γ in these equations is the activity coefficient which links thermodynamic activity to molal concentration.

The solubility of calcite (and dolomite) is somewhat temperature dependent. The temperature dependence is incorporated in the temperature dependence of the equilibrium constants. The solubility of calcite, expressed as mass concentration of CaCO₃, is displayed in Figure 21.9. The solubility increases from 6 mg/L in the absence of CO₂ to 50 mg/L in equilibrium with atmospheric CO₂ to several hundred mg/L in equilibrium with the CO₂ pressures found in many soils.

21.4.2 Chemical Parameters for Karst Aquifers

As a practical tool in karst hydrology, neither the underlying physical chemistry as described above nor raw chemical analyses of groundwater are the most useful. Parameters, solidly based on theory, but using analytical data are the hardness, the Ca/Mg ratio, saturation indices for calcite and dolomite, and the calculated CO₂ partial pressure.

Hardness is one of the oldest water quality parameters. It is defined as:

$$Hd = M_{\text{CaCO}_3} (m_{\text{Ca}} + m_{\text{Mg}}) \quad (21.26)$$

M_{CaCO_3} is the molecular weight of calcium carbonate. Concentrations of dissolved Ca and Mg are given in molal units. The hardness describes the total dissolved calcium and magnesium but expresses the result

as mg/L CaCO₃. Mainly, it is a measure of the total dissolved carbonate rock although all other sources of Ca and Mg are included. The parameter has little theoretical significance but is very important for water supplies, particularly those in which precipitated carbonates in the plumbing and in boilers are important (Loewenthal and Marais, 1976).

The Ca/Mg ratio is simply the relative concentrations of Ca²⁺ and Mg²⁺ on an atomic scale (molal concentration):

$$\frac{\text{Ca}}{\text{Mg}} = \frac{m_{\text{Ca}}}{m_{\text{Mg}}} \quad (21.27)$$

The Ca/Mg ratio provides information on the rock units through which the groundwater has passed. The Ca/Mg ratio = 1 to 1.5 for dolomite aquifers and 6 to 8 for limestone aquifers. Intermediate values imply a dolomitic limestone or mixed limestone/dolomite sequence.

The degree of disequilibrium between groundwater and the carbonate wall rock is given by the saturation index, defined for calcite as:

$$SI_c = \log \frac{a_{\text{Ca}^{2+}} a_{\text{CO}_3^{2-}}}{K_c} = \log \frac{K_{\text{iap}}}{K_c} \quad (21.28)$$

K_{iap} is the ion activity product derived from actual analyses of the water. K_c , as above, is the solubility product constant for calcite. Saturation indices can be written for any mineral in a similar fashion. If the water is exactly at equilibrium, the experimental ion activity product will be exactly equal to the solubility product constant and $SI_c = 0$. If the water is undersaturated with respect to calcite, K_{iap} will be less than K_c and the saturation index will be negative. For a supersaturated water, SI_c will be positive.

Calcium ion activity can be determined directly from the analyzed calcium concentration. The carbonate ion, however, is a minority species in most carbonate groundwaters. Its activity must be calculated from the measured bicarbonate ion concentration and the pH. A final expression for saturation index as a function of experimentally determined quantities is:

$$SI_c = \log \frac{\gamma_{\text{Ca}^{2+}} m_{\text{Ca}^{2+}} \gamma_{\text{HCO}_3^-} m_{\text{HCO}_3^-} K_2}{K_1 10^{-\text{pH}}} \quad (21.29)$$

See White (1988), Drever (1997), or Langmuir (1997) for a full discussion of the thermodynamics.

The effective concentrations of carbonate species in the groundwater are related to the CO₂ partial pressure in the source area through Equations 21.18 and 21.19. These reactions may be combined to allow the calculation of carbon dioxide pressure from the water analysis

$$P_{\text{CO}_2} = \frac{\gamma_{\text{HCO}_3^-} m_{\text{HCO}_3^-} 10^{-\text{pH}}}{K_1 K_{\text{CO}_2}} \quad (21.30)$$

Because the observed CO₂ partial pressures in karst systems are rarely lower than the atmospheric background, it is often helpful to normalize the calculated CO₂ pressure by dividing by the background atmospheric CO₂ pressure. The result expresses the CO₂ in the groundwater sample as an enhancement over background.

21.4.3 Dissolution Kinetics of Limestone and Dolomite

The central role of dissolution kinetics is a unique characteristic of karst aquifers. In porous media aquifers and in most fracture aquifers, travel times and surface/volume ratios are both large. The groundwater is in intimate contact with the mineral grains of the aquifer rock and remains in contact for long periods of time. As a result, there is sufficient time for chemical reactions to proceed to completion, and it is generally safe to assume that chemical reactions are at equilibrium. Geochemical modeling of aquifer chemistry can make use of speciation and reaction path programs such as WATEQ4F, PHREEQC, and MINTEQA2.

In karst aquifers, travel times are measured in hours or days which are on the same order of magnitude as the reaction time of the water with the rock. As a result, karst groundwater can travel the entire width of the aquifer without coming into chemical equilibrium. The rates of dissolution reactions are an essential part of any attempt to model the evolution of a karst aquifer.

The characteristic time scales for the various carbonate reactions are:

1. Dissociation of carbonic acid into bicarbonate ion and bicarbonate ions into carbonate ions: milliseconds.
2. Hydration of aqueous CO_2 to form carbonic acid: 30 sec.
3. Dissolution of calcite: 2 to 3 days.
4. Dissolution of dolomite, 1st stage: 3 to 5 days.
5. Dissolution of dolomite, 2nd stage: months to years.

It can be assumed from (21.1) that the concentrations of H_2CO_3 , HCO_3^- , and CO_3^{2-} in solution are always in equilibrium. The 30 sec time scale for producing carbonic acid from aqueous CO_2 is important for sheet flow of thin films of rapidly moving water in shafts and chimneys in the vadose zone. The dissolution kinetics of calcite and dolomite are central to the interpretation of the evolution of karst aquifers.

The raw data for limestone and dolomite dissolution are deceptively simple (Figure 21.10). The rate is proportional to the slope of the curve. The rate is initially rapid as the rock is attacked by highly undersaturated water but then slows as the solution approaches saturation. Fitting the experimental curves to theoretical models has been challenging and there have been a tremendous number of investigations many of which are reviewed by Morse and Arvidson (2002).

A useful and widely adopted rate equation was developed by Plummer et al. (1978; 1979).

$$\text{Rate} = k_1 a_{\text{H}^+} + k_2 a_{\text{H}_2\text{CO}_3} + k_3 a_{\text{H}_2\text{O}} - k_4 a_{\text{Ca}^{2+}} a_{\text{HCO}_3^-} \quad (21.31)$$

In Equation 21.31, k_1 , k_2 , and k_3 are forward rate constants and k_4 is the backward rate constant. The rate is given in units of millimoles $\text{cm}^{-2} \text{ sec}^{-1}$.

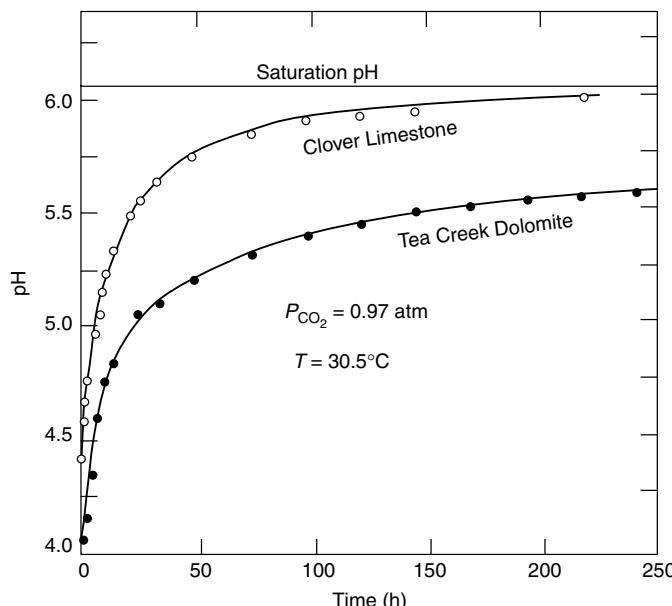


FIGURE 21.10 Rate curve for the dissolution of an Ordovician limestone and an Ordovician dolomite. These data are plotted in terms of the changing pH of a fixed volume of solution maintained at a fixed CO_2 pressure of one atmosphere. Based on unpublished data from author's laboratory.

The first term shows that the rate is directly proportional to the hydrogen ion activity and dominates in strongly acid solution. It describes, for example, the interaction of acid mine water with limestone. The first component of the forward rate is mass-transport controlled so that the rate depends on the rate of flow of water across the rock surface. The second term describes the dissolution process driven by carbonic acid. It is dependent on the available CO₂ but is surface-reaction controlled and does not depend strongly on fluid motion. The third term describes a constant background rate of limestone dissolving directly in water and is usually a minor contribution because the solubility of limestone in pure water is very low. All three forward rate constants are known functions of temperature:

$$\log k_1 = 0.198 - \frac{444}{T} \quad (21.32)$$

$$\log k_2 = 2.84 - \frac{2177}{T} \quad (21.33)$$

$$\log k_3 = -5.86 - \frac{317}{T} \quad (T < 298 \text{ K}) \quad (21.34)$$

$$\log k_3 = -1.10 - \frac{1737}{T} \quad (T > 298 \text{ K}) \quad (21.35)$$

The backward reaction rate constant can be calculated from the forward rate constants:

$$k_4 = \frac{K_2}{K_C} \left\{ k'_1 + \frac{1}{a_{\text{H}^+}} [k_2 a_{\text{H}_2\text{CO}_3} + k_3 a_{\text{H}_2\text{O}}] \right\} \quad (21.36)$$

K_2 and K_C are equilibrium constants (Table 21.4) and k'_1 refers to the hydrogen ion concentration at the dissolving interface. Experimental comparisons have been made by setting $k'_1 = k_1$.

When dissolution rate is plotted against some measure of saturation, the rate is found to be a function of CO₂ partial pressure but only slightly dependent on undersaturation until the saturation index reaches values near -0.3. Between this value and equilibrium the rate plummets by orders of magnitude (Berner and Morse, 1974; White, 1977). This abrupt decrease in dissolution rate in waters that remain somewhat undersaturated is a crucial feature in explaining the development of long conduits in karst aquifers.

Recently, researchers (e.g., Dreybrodt and Buhmann, 1991; Palmer, 1991) have found it convenient to use a genetic rate equation that does not attempt to separate the components of the dissolution reactions

$$\text{Rate} = \frac{A}{V} \frac{dC}{dt} = k \left(1 - \frac{C}{C_s} \right)^n \quad (21.37)$$

In this equation, A = surface area of limestone, V = volume of water in contact with the limestone, k = rate constant, C = instantaneous concentration, C_s = saturation concentration, and n = reaction order. For values of C/C_s less than 0.83, the reaction order, $n = 1$ and reaction rates are fast. As the solution approaches saturation and C/C_s becomes greater than 0.83 (approaching unity when the solution reaches equilibrium), the reaction order shifts abruptly to $n = 4$ and reaction rates greatly decrease.

Many factors influence the rate at which carbonate rocks dissolve in CO₂-containing groundwater. New factors appear when the flows are turbulent and when the groundwater is close to equilibrium. Among those identified (Dreybrodt and Buhmann, 1991; Svensson and Dreybrodt, 1992; Dreybrodt et al., 1996; Liu and Dreybrodt, 1997) are:

1. The hydration reaction of aqueous CO₂ to carbonic acid, which has a time constant of about 30 sec, is shown to be rate-controlling at high A/V ratios and under some conditions of turbulent flow.
2. Adsorption of impurity ions on the reactive surface becomes rate-controlling when the solution is near saturation.
3. A diffusion boundary layer becomes important under conditions of turbulent flow.

Karst development in dolomite is generally more subdued than karst development in limestone. It is not a matter of solubility. The equilibrium solubilities of limestone and dolomite are very similar. However, the kinetics of dolomite dissolution is more sluggish than the kinetics of limestone dissolution. The dissolution of dolomite in highly unsaturated water can be described by a rate equation of the form (Busenberg and Plummer, 1982; Herman and White, 1985)

$$\text{Rate} = k_1 a_{\text{H}^+}^n + k_2 a_{\text{H}_2\text{CO}_3}^n + k_3 a_{\text{H}_2\text{O}}^n - k_4 a_{\text{HCO}_3^-} \quad (21.38)$$

The exponent, $n = 0.5$ for temperatures below 45°C. However, unlike the calcite rate curve which approaches the equilibrium value on times on the order of a few days, the dolomite rate curve flattens out at a saturation index of $\text{SI}_d = -2$. A different and slower mechanism takes effect. The final approach to equilibrium requires times on the order of months to years. Groundwater in dolomite aquifers does eventually come into equilibrium with the wall rock because well waters from dolomite are found to be near saturation.

As a result of slower dissolution kinetics, conduit systems are generally less well developed in dolomite aquifers. Dissolution tends to follow pathways through limestones if these are available. If there are no alternative pathways available, large conduits do develop in dolomite but appear to do so at a much slower rate.

21.5 Evolution of Karst Aquifers

21.5.1 Life History of a Conduit

Although all aquifers evolve in the sense that the permeability distribution varies as minerals are dissolved or precipitated and the overlying land surface is lowered by erosion, these processes are generally very slow. Karst aquifers, in contrast, evolve on a time scale as short as 10^4 years as bedrock is dissolved and new conduits are developed. A particular pathway through the mass of soluble rock will have a “life history” as sketched in Figure 21.11. It is assumed that the initial state is obtained by mechanical fracturing of the bedrock. The initial pathway is along a sequence of joints, fractures, or bedding plane partings, more rarely a fault or sequence of primary vuggy porosity. The *initiation* phase is that time during which the primary mechanical openings are enlarged by dissolution from their initial apertures of tens to hundreds of micrometers to a width on the order of 10 mm. When the aperture reaches 10 mm, the flow regime crosses a set of critical thresholds in the dissolution and transport processes so that a 10-mm aperture is the practical boundary between openings that would be considered part of the fracture permeability and openings that would be part of the conduit permeability.

Once the critical thresholds have been passed, the conduit enters an *enlargement* phase and grows to dimensions typical of cave passages by the continued dissolutional retreat of its walls. As the enlargement phase comes to a close, several alternative scenarios are possible. If the conduit is so situated that it remains water-filled, solution will continue until the ceiling span becomes mechanically unstable and the passage will collapse. As the flooded passage enlarges, however, velocities decrease, the water will become more saturated, solution rates will slow, and so water-filled passages may persist for long periods of time. If the recharge source for the conduit remains constant as base levels are lowered, the passage will eventually grow to a size where the available recharge will not fill it and a free air surface will develop. If the recharge source is not diverted, there will be a transition from a tubular conduit to a canyon morphology and the canyon will continue deepening as base level continues to lower. If the recharge source is diverted to some lower route with lowering base levels, the conduit will be drained but will persist as an abandoned dry tube. Those tubes and canyons, which by chance circumstances develop entrances, are the population of explorable caves.

After the conduit has been drained, it enters a *stagnation* phase that may persist for millions of years. The clastic sediments in dry caves were mostly deposited at the transition from the enlargement to stagnation phases. Speleothems are deposited during the stagnation phase. These are the deposits that make caves an important paleoclimate archive. Eventually, however, erosional lowering will sweep the land surface past

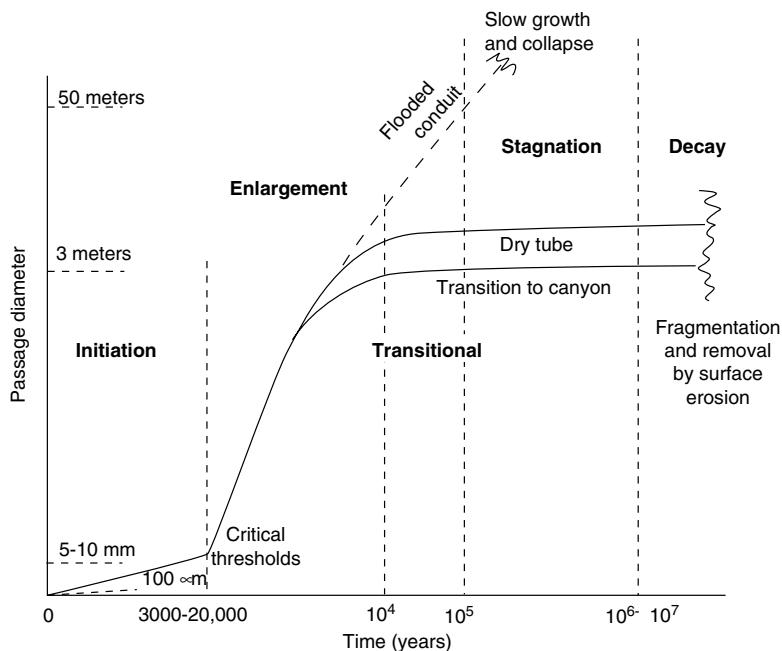


FIGURE 21.11 Schematic illustration of the “life history” of a conduit. The time scale is derived from geochemical modeling in the initiation and enlargement phases and from geologic and age-dating evidence in the stagnation and decay phases.

the cave passage. The conduit will be fragmented and ultimately destroyed. The later part of this sequence has been sketched separately as a final decay phase when cave passages are truncated by collapse and broken into short passage fragments, as many caves are observed to be, before they are destroyed completely.

One of the objectives of cave development theories is to assign a time scale to the various phases in the life history of the conduit, particularly the early ones where dissolution of the bedrock is the dominant process. The time scale of Figure 21.11 is based on calculations for the early stages and on geomorphic and age-dating evidence for the later stages. The stagnation and decay phases are determined by events elsewhere in the drainage basin or by regional tectonics. It is not possible to calculate what happens from a general model although events in specific drainage basins can sometimes be interpreted.

A question not yet addressed by any model is whether, given the calculations and analysis for single passages, the results could be put together to obtain the equivalent “life history” for an entire conduit system and its associated groundwater basin. If one could, it would allow us to say something about the historical development of the drainage, and still further extrapolated, to say something about events in the regional climatic and tectonic setting in which the drainage basin is embedded. Overall, a formidable task, but an important one.

21.5.2 Initiation and Critical Thresholds

In compact and impermeable limestones, the initial hydraulic pathway from recharge point to discharge point is along a sequence of joints, fractures and bedding plane partings with mechanical apertures in the range of tens to hundreds of micrometers. The discharge through the fracture system is described by the cubic law (Equation 21.4 and its recent modifications). The volume and velocity of water flowing through the fractures both increase rapidly as the fractures widen. When unsaturated water from the recharge source reaches the carbonate bedrock, much of the dissolutional capacity is consumed by reaction at the bedrock interface because of the rapid kinetics and high undersaturation. Water that enters the fracture system remains slightly undersaturated because of the very slow rate of reaction in water close

to equilibrium. In spite of the small fracture aperture and long travel times, slightly undersaturated water penetrates deep into the aquifer where the fractures are slowly enlarged. As the fractures enlarge, the volume of flow increases with the cube of the aperture and the saturation index becomes more negative. Eventually, the fracture system will become sufficiently enlarged that water undersaturated at the critical value of $SI_C = -0.3$ penetrates the entire length of the hydraulic pathway, a condition called “breakthrough” by some authors. The dissolution kinetics shift from fourth order to first order and dissolution rates increase rapidly. This transition occurs when the aperture reaches about ten mm.

Calculations based on a range of reasonable hydraulic gradients and carbon dioxide partial pressures gave a time scale for the initiation phase in the range of 5000 years (White, 1977). More elaborate calculations based on computer models give a range from 3000 to 20,000 years (Groves and Howard, 1994; Dreybrodt, 1996; Siemers and Dreybrodt, 1998). The elapsed time from initial fracture to breakthrough is a sensitive function of the initial fracture aperture

By coincidence, when the aperture of the enlarging fracture system reaches ten mm, not one, but three thresholds are crossed. These thresholds mark the transition from fracture permeability of conduit permeability:

1. The onset of turbulence
2. The penetration (breakthrough) of the aquifer by undersaturated water
3. The onset of clastic sediment transport

For a fixed input and fixed output defining the hydraulic gradient across an aquifer, the flow velocity within a widening fracture can be calculated. From the velocity and fracture aperture, the Reynolds number can be calculated and tracked as both velocity and aperture increase as dissolution proceeds. For a range of hydraulic gradients from 10^{-4} to 10^{-2} , the onset of turbulence occurs at an aperture of 10 to 5 mm, respectively. The onset of turbulence marks the beginning of a regime where a linear relation between head and flow velocity can no longer be used to construct models of groundwater flow.

When one of the possible hydraulic pathways through the aquifer achieves an aperture that permits highly undersaturated water to penetrate the aquifer, the dissolution kinetics shifts from the slow regime to the fast regime. Dissolution rates then increase rapidly so that the hydraulic pathway that first achieves breakthrough will undergo a runaway process and grow at the expense of alternate hydraulic pathways. It is this kinetic trigger that is responsible for the aquifer containing a small number of large conduits rather than an arrangement of small opening along all possible pathways. For the kinetic trigger and associated runaway process to function, the aquifer must have sufficient hydraulic gradient to permit velocity to increase in response to increasing aperture. A failure of the kinetic trigger to function leads to the development of maze caves.

Carbonate aquifers maintain a flux of clastic sediment. Sediment can be transported as bedload or as suspended load both of which move through the aquifer mainly during flood flow. Bedload transport requires a finite boundary shear between the bedded sediment and the flowing water. The velocity to initiate sediment movement is of the order of 0.1 m sec^{-1} . This will occur, depending on hydraulic gradient, when the aperture reaches the range of 10 mm.

21.5.3 Enlargement, Maturity, and Decay

The selection of hydraulic pathway during the initiation phase is determined by the arrangement of fractures and bedding plane partings, by changes in lithology that influence dissolution rates, and by the presence of confining layers and other geological barriers. The pattern of the conduit system is set at the initiation phase. Once the water in the growing conduit reaches the fast kinetics regime, dissolution rates are nearly constant and enlargement is uniform along the entire conduit. If Equation (21.37) is accepted as a description for the dissolution rate of limestone, the retreat of conduit walls can be described by Palmer (1991):

$$S = \frac{31.56k(1 - (C/C_S))^n}{\rho_r} \quad (21.39)$$

S is the rate of wall retreat in cm/year, k is the dissolution rate constant, C_s is the saturation concentration of dissolved carbonate, and ρ_r is the density of the wall rock. The effect of temperature and variations in the dissolution rates for rocks of different lithology are incorporated in the rate constant, k . The available CO₂ and also temperature determines the saturation concentration C_s . Palmer finds good agreement between observed dissolution rates and those predicted by Equation (21.39).

Typical calculated values for the retreat of a conduit wall are in the range of 40 cm/1000 years, a rate that would produce a conduit 4 m in radius in 10,000 years. These results represent an upper limit. Other factors such as walls coated with insoluble residue from the limestone or with mud carried into the cave from the land surface would slow down the dissolution process. Generally, the rapid enlargement converts the initial fracture geometry into a more nearly tubular cross section. The rate of enlargement depends on the path length and on the hydraulic gradient. Under the high gradients and short path lengths that would be found in such locations as dam abutments, significant dissolutional enlargement could take place within the design lifetime of the structure (Dreybrodt, 1992).

Because development of conduit permeability in karst aquifers takes place on very short time scales, continuous conduit development at successively lower elevations can easily keep pace with downcutting of surface river valleys. Geological evidence that the theoretical predictions are reasonable is provided by caves that are known to have developed in the 11,000 years since the retreat of the last Pleistocene ice sheet and also by caves in the Bahamas that have developed in limestones that themselves are only 85,000 years old (Mylroie and Carew, 1987).

In many karst regions, conduits are graded to local base levels. The conduits that can be readily explored are in the later stages of the enlargement phase or in the early stages of the stagnation phase. They are often occupied by free-surface streams and subject to flooding. Many conduits have alternating reaches of air-filled passage (at least during base flow) and permanently water-filled passage. The streams in these conduits and the springs where the water returns to the surface are usually undersaturated with respect to calcite, with saturation indices typically in the range of -0.5 to -0.3, exactly what is expected for continuous enlargement by dissolution in the fast kinetics mode.

Mature conduits lying close to regional base levels act as low gradient master drains for the matrix and fracture portions of the aquifer. Conduits may be blocked by rockfalls which pond the water and thus create sediment traps. Bypass routes and flood spillover routes evolve. Continually lowering base levels shift hydraulic gradients. The new gradients initiate new conduit development at lower elevations, and eventually the recharge sources are diverted into the new conduits. The original conduits may survive as dry cave passages for long time periods in the stagnation phase. These abandoned conduits provide useful records of past base levels and are markers for the geomorphic history of the drainage basin. Cosmogenic isotope dating of quartz sand and pebbles in the cave sediment can be used to establish an absolute chronology for the conduits. Such dating gives a sequence of tiered cave passages in Mammoth Cave extending back 3 million years (Granger et al., 2001), a sequence in the Cumberland Plateau of Tennessee extending to 5 million years (Anthony and Granger, 2004), and a sequence in the California Sierra extending to 1.2 million years (Despain and Stock, 2005). Shafts and chimneys carrying vadose water from overlying perched aquifers or from closed depressions sometimes intersect the older conduit system. Underground waterfalls and segments of very high gradient streams display contemporary drainage moving nearly vertically in the vadose zone as it passes through the abandoned conduits.

21.6 Quantitative Hydrology of Karst Aquifers

21.6.1 Water Budget

The concept of the groundwater basin allows the water budget to be used as a powerful tool in the evaluation of karst aquifers. The framework for an overall water balance as an input/output calculation is implied by Figure 21.3. Spring discharge and alloegenic recharge can be gauged directly. Diffuse infiltration and internal runoff can be estimated from catchment areas providing that they can be corrected for

evapotranspiration losses. High internal runoff tends to lower evapotranspiration losses because of the rapid transport of water underground.

The calculation requires that the basin area be known and for this the divides for groundwater basins must be drawn. As the divides for groundwater basins cannot be directly observed, they must be constructed from as much available evidence as possible:

1. Results of tracer tests linking sinking streams, cave streams, and other accessible inputs to the discharging spring.
2. Water table slopes determined from well inventories.
3. Geological constraints imposed by folds and faults, regional dips, fracture patterns, confining layers, and the upper and lower boundaries of the carbonate rocks.
4. Chemical fingerprinting — comparing spring water chemistry with water chemistry within the aquifer.

The groundwater basin as finally drawn must be consistent with the water budget. A seriously unbalanced budget is an indication that the boundaries of the groundwater basin have not been drawn correctly. Water budgets should be constructed under both high and low flow regimes. Spillover routes that transport water across basin boundaries are common and may become active only during high flow regimes.

21.6.2 Base Flow/Runoff Relations

Base flow in surface streams is maintained by groundwater discharge. If the region has a distinct dry season, the mean base flow for the surface streams gives a measure of the regional base flow. This is normalized to basin area to yield a unit base flow:

$$Q_N = \frac{\bar{Q}_B}{A} \quad (21.40)$$

Springs fed by conduits behave much like surface streams and exhibit a low storage and therefore a low base flow. For those aquifers for which the spring discharge reaches base flow conditions, Equation (21.40) can be used to estimate the area of the groundwater basin (E.L. White, 1977). Tests of the relationship by Quinlan and Ray (1995) suggest that the method is reasonably accurate provided that it is calibrated for the climate and geologic setting of the aquifer of interest.

21.6.3 Tracer Studies

Much of the advance in understanding karst aquifers in the past several decades have come from tracer studies (Jones, 1984; Mull et al., 1988; Käss, 1998). An identifiable substance is placed in a sinking stream, cave stream, well, or flushed into a closed depression. The substance is later detected at an outlet point, usually a spring but sometimes an underground stream or another well. Tracers that have been used include:

- Fluorescent dyes
- Spores
- Salt
- Rare elements (e.g., bromide)
- Noble gases
- Other substances

Although tracing using colored lycopodium spores received a great deal of attention in Europe, it has not proved useful in the conduit aquifers of eastern United States where springs are frequently muddy and the plankton nets used to capture the spores easily become clogged. Traces using NaBr combined with neutron activation analysis to detect low levels of Br⁻ have been used but are expensive because of the

analytical methods. Traces with massive injections of salt allow easy detection by measurement of specific conductance but are generally unacceptable from environmental considerations.

Injections of noble gases are useful for tracing fracture flows. A noble gas, typically helium or neon, is bled continuously into the groundwater through a well drilled along the fracture. Water collected from another well or spring is then analyzed for the noble gas by mass spectrometry. Unlike other tracers which use a single spike of the tracer material, noble gas traces use a continuous injection over a long period of time. It is thus suitable for traces through tight fractures where long travel times are expected.

By far the most important tracers for characterizing karst aquifers are fluorescent dyes. The technique is the same as with other tracer tests. Dye is injected into the aquifer at points of interest and detected at springs or other discharge points. With increasing levels of sophistication, these detection methods are:

- Visual observation of dye
- Visual test with dye receptor
- Fluorescence detection with dye receptor
- Quantitative test with automatic sampler
- Quantitative test with on-site spectrophotometer

Visual tests require the injection of sufficient dye to visibly color the water at the discharge point. Other than the inconvenience of positioning observers and the possibility of dye coming through at night, visual traces require such large quantities of dye that the test would be a source of groundwater contamination. Further, injection of large quantities of dye may be prohibitively expensive, particularly if multiple traces are needed. Direct visual observation has been used historically but is now rarely used.

A major advance made in the late 1950s was the introduction of packets of activated charcoal as dye receptors. A few grams of high-grade coconut charcoal contained in mesh bags are placed in all suspected resurgent points. Dye is strongly adsorbed onto the charcoal and is not desorbed by continuing flow of water. The packets are collected at convenient intervals and replaced by fresh packets. The dye is elutriated with an alcoholic solution of strong base. The detection of dye by visual observation of the elutriate can be improved by observing the fluorescence of the solution under ultraviolet light. Use of the charcoal receptors permits many possible discharge points to be tested at the same time.

Optical brighteners can also be used for qualitative traces. These are agents widely used in laundry detergents so background should be checked. The receptors are wads of unsized cotton onto which the brightening agent strongly adsorbs. The presence of the brightener on the cotton is determined by observing its bright blue luminescence under long wave ultraviolet radiation. Charcoal packets will also adsorb optical brightener.

A more quantitative trace can be made by installing automatic water samplers at suspected discharge points. These can be programmed to collect water samples at prescribed intervals. The water samples are analyzed by spectrophotometry. As the analysis is of the water itself rather than an elutriate, the analysis can be calibrated to determine fraction of dye recovered. A plot of dye concentration as a function of time gives a breakthrough curve (Figure 21.12) that shows the exact travel time. The shape of the dye pulse often gives other clues to the nature of the flow path.

To carry out tests on more than one injection point requires the use of multiple dyes. The fluorescence bands are broad and overlap so that the identification of mixed dyes is difficult by visual observation alone. However, fluorescence spectra can be deconvoluted with suitable software, allowing quantitative determination of the different dye combinations.

A recent development is the use of fiber optic probes and portable spectrophotometers directly at the detection site. The fluorescence signal is recorded on a data logger and transferred to a computer. This technique permits a much greater density of measurement points and eliminates the need for collecting large numbers of samples. As a guideline, charcoal packets are very inexpensive, automatic water samplers are expensive, and portable spectrophotometers more expensive.

The requirements for tracers is that they be nontoxic, stable in the groundwater environment, non adsorbent on clays, silts and the carbonate wall rock, and detectable at very low concentrations. Tracers that are most widely used are listed in Table 21.5. Details on the chemistry, properties and toxicity of these

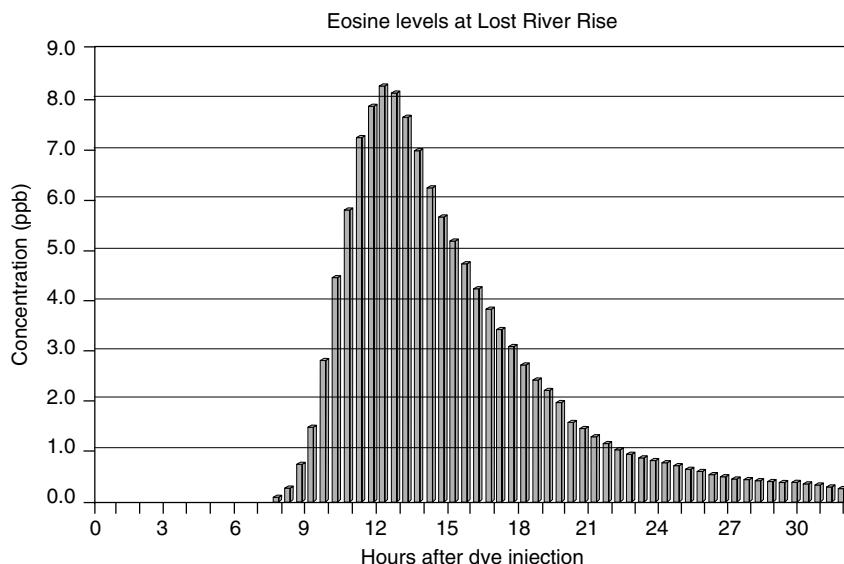


FIGURE 21.12 Dye breakthrough curve for the Lost River Rise, Warren County, Kentucky. Plot courtesy of the Crawford Hydrology Laboratory, Western Kentucky University.

TABLE 21.5 Commonly Used Tracer Dyes for Karst Hydrology

Common Name	Color Index	Fluorescence Wavelength (nm)	
		Elutriate	Water
Sodium fluorescein (uranine)	Acid yellow 73	515.5	508
Eosine	Acid red 87	542	535
Rhodamine WT	Acid red 388	568.5	576
Sulphorhodamine B	Acid red 52	576.5	585
Optical brightener	Tinopal CBS-X	398.0	397

Data courtesy of Crawford Hydrology Laboratory, Western Kentucky University.

and other tracer dyes may be found in Smart and Laidlaw (1977) and Smart (1984). Use of fluorescein as a water tracer has a 100-year history and remains the most popular tracer. It decomposes in sunlight and so does not persist after emerging from the groundwater system. The main drawback of fluorescein is that it is used in many common products such as antifreeze and toilet bowl cleaners, which often produce an unacceptable background.

If there are many potential discharge points to be tested, a qualitative investigation with dye receptors can be used to delineate the overall drainage pattern. These results can be followed up with quantitative studies to obtain travel times and flow characterization. When legal and regulatory issues are involved, the quantitative trace provides an objective record to be entered into evidence.

Tracer testing is by far the most powerful technique available for working out the conduit system of a carbonate aquifer. Combined with direct exploration and survey of such parts of the active conduits as may be accessible, it should be possible to delineate the groundwater basin, determine piracy routes, specify allogenic recharge and base level discharge points, and set down at least a rough plan of the conduit system.

21.6.4 Well Tests and Water Table Mapping

The literature is ambiguous concerning the properties of the karst water table and even its very existence. Many examples are known where air-filled cave passages lie below flowing streams. Certain alpine caves

ingest snow melt and runoff as turbulent frothing streams and crashing waterfalls. These have been cited as evidence for the absence of a water table in karst aquifers. Much of the confusion appears to arise from a failure to appreciate the dynamic response of the conduit system to storm runoff and failure to appreciate the mismatch in hydraulic conductivity between the conduit system and the associated fracture and matrix systems.

Because of their high flow capacity, conduits cannot support much hydraulic head and thus create a trough in the water table during periods between storms. Because of their rapid response to recharge, conduits quickly fill during storm flow and the groundwater trough may become a groundwater mound. Losing surface streams perched over conduits that are air-filled at base flow may become gaining streams when the conduit is flooded. If the hydraulic conductivities of the fracture and matrix systems are low, the conduit system, particularly one with a large component of alloegenic recharge, may act as a decoupled set of pipes with little influence on the water table.

The fracture and matrix permeability portions of the aquifer support a well-defined water table that can be mapped if a sufficient number of wells is available. Water tables that have been mapped in karst areas appear reasonable and the locations of conduits are often revealed by the groundwater troughs.

Drilling test wells is an appropriate way to probe the fracture and matrix component of the aquifer. However, account must be taken of the heterogeneous distribution of fracture permeability on the borehole scale. Wells drilled on fracture intersections may produce several orders of magnitude, more water than wells drilled in the same formation in rock between fractures. When the bedrock is not masked by thick soils or overlying caprock, and where the permeability is dominated by fractures rather than by bedding plane partings, air photographs can be used to map fracture traces at the land surface (Lattman and Parizek, 1964). These are then used to select drilling sites.

Standard pump tests must be used and interpreted with great care in carbonate aquifers. In the unlikely event of a well penetrating a master conduit, the yield will be high and the drawdown low. If the well happens to penetrate mostly unfractured rock, the yield will be low and the drawdown very large. It is possible that these two wells could be drilled within a few meters of each other. In either case, the pump test gives a very misleading interpretation of the overall properties of the aquifer. A better technique is to test the well in segments using packers to isolate the segments. Such tests carried out in the Silurian dolomite aquifer of the Door Peninsula of Wisconsin (Figure 21.13) shows the extreme variability in hydraulic conductivity found in a single well (Muldoon et al., 2001).

Downhole video cameras are a useful tool for locating fracture zones and directly examining solution features intersected by wells. If the well can be pumped dry, the camera can be used to observe which of the solution openings is actually discharging water into the well. Packers can then be used to isolate productive zones and these can be pump tested or injection tested independently of the remainder of the well bore.

21.6.5 Hydrographs of Karst Springs

While exploration and tracer studies reveal something of the conduit system and wells and pump tests characterize the fracture and matrix systems, the only hydraulic response that depends on both systems is that of the spring that drains the aquifer. Spring hydrographs and fluctuations in spring water chemistry provide very useful guides to aquifer hydraulics.

Spring hydrographs can be grouped into three types (Figure 21.14) depending on the response time of the aquifer compared with the mean spacing between storms. In fast response aquifers (Type I), storm pulses can pass completely through the aquifer and conditions return to base flow before the next storm pulse arrives. Each storm produces its own hydrograph at the spring which is not different in overall characteristics from hydrographs of surface streams. In slow response aquifers (Type III) the throughput time is sufficiently long to completely flatten the individual storm hydrographs and all that remains is a broad rise and fall relating to wet and dry seasons. In the limit, the slow response aquifer may produce only a horizontal straight line, a spring of constant flow. When the response time is comparable to the mean spacing between storms (Type II), some structure appears in the hydrograph but the individual storm pulses are not resolved.

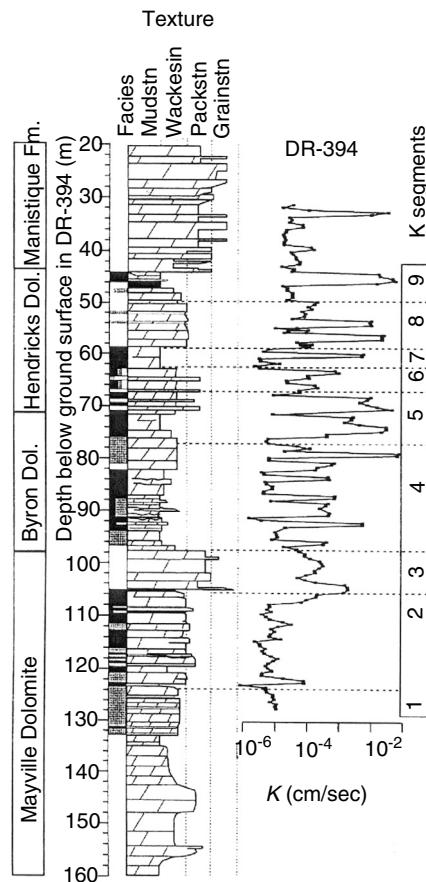


FIGURE 21.13 Distribution of hydraulic conductivity along the well bore for a well drilled in Silurian dolomite in Door County, Wisconsin. (Adapted from Muldoon, M.A., Simo, J.A. and Bradbury, K.R. (2001) *Hydrogeol. J.* 9, 570–583.)

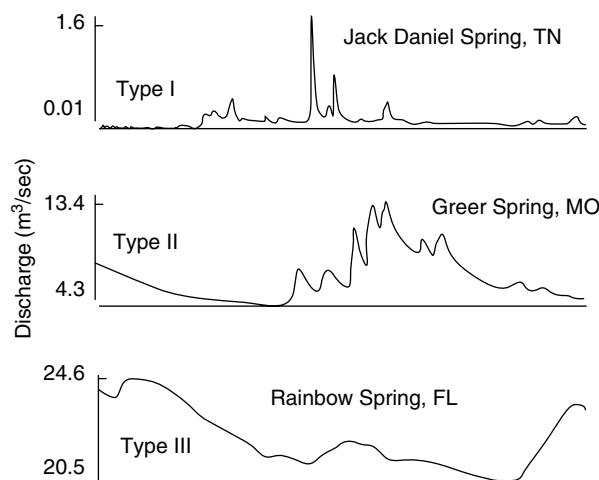


FIGURE 21.14 Types of spring hydrographs. Maximum and minimum discharges are given numerically. The x-axis represents one water year. Data from U.S. Geological Survey surface water records.

Factors that determine the type of spring hydrograph include degree of conduit development, the mix of matrix and fracture flow, area of groundwater basin, proportion of allogenic recharge, and hydraulic gradient from input to output. Small basins with well-developed conduit systems are more flashier than larger basins with the same degree of conduit development. Longer travel times in large basins tend to smooth out the hydrograph even if there are open conduits. Very low gradients reduce travel times and produce the same effect. Both factors seem to be operating in the large springs of Florida which are known from divers' explorations to be fed by large conduits yet have Type III hydrographs.

A useful parameter for describing the flashiness of an aquifer is the ratio of the maximum annual discharge, the quantity known in surface water hydrology as the annual flood, to the base flow, averaged over the number of years of record.

$$Q_f = \frac{1}{n} \sum_{i=1}^n \frac{(Q_{\max})_i}{(Q_{\text{base}})_i} \quad (21.41)$$

Type I responses have Q_f in the range of 70 to 100, Type II in the range of 5 to 10, and Type III in the range of 1 to 2, based on relatively sparse data.

Individual storm hydrographs can be further analyzed (Figure 21.15). There is a lag between the rainfall event and the arrival of the storm pulse at the spring. This lag, however, does not measure the time required for storm water to traverse the aquifer. The injection of water into the upstream reaches of the aquifer raises hydraulic gradients, floods conduit, and forces out by piston flow, water that has been stored in continuously flooded parts of the conduit system. The recession limb of the hydrograph can be described by one or more functions of the form

$$Q = Q_0 e^{-\varepsilon t} \quad (21.42)$$

The coefficient, ε , sometimes called an exhaustion coefficient, has units of reciprocal time and is inversely related to the response time of the aquifer, $t_R = 1/\varepsilon$.

Very characteristically, the recession limb of spring storm hydrographs will consist of a fast response segment and a slow response segment (Figure 21.16). Sometimes there are more than two distinct segments. The fast response segment can be assigned to the conduit portion of the aquifer which drains rapidly when the storm input ceases. The slow response segment can be assigned to the fracture system draining into the conduit after the conduit itself has emptied. During storm flow, flooded conduits have

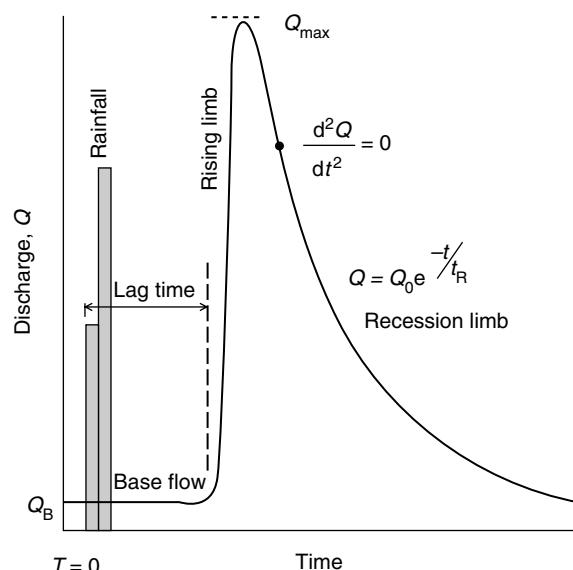


FIGURE 21.15 Sketch showing components of a storm hydrograph at a karst spring.

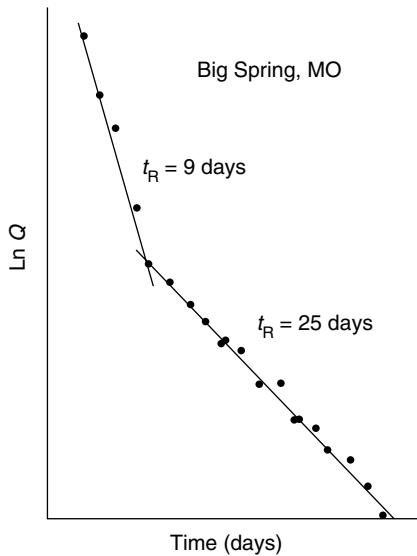


FIGURE 21.16 Semi-log plot of hydrograph recession curve showing fast response and slow response components. Data from U.S. Geological Survey surface water records.

gradients that force water into the surrounding aquifer. As the conduit drains after the storm, gradients reverse and the water stored in the fractures is released into the conduits.

Analysis of the recession limb of the spring hydrograph gives some information on the relative importance of the conduit and fracture systems and the coupling between them. The exhaustion coefficient gives an estimate of the volume of water in dynamic storage above regional base level (Burdon and Papakis, 1963):

$$V = \frac{86,400 Q_{\max}}{\varepsilon} \quad (21.43)$$

V is the volume of stored water in m^3 , 86,400 is the number of seconds in a day, Q_{\max} is the peak discharge in $\text{m}^3 \text{ sec}^{-1}$ and ε is in units of days^{-1} . The exhaustion coefficient is also related to the transmissivity, T , and storativity, S , through (Milanović, 1981):

$$\varepsilon = \frac{2T}{SL^2} \quad (21.44)$$

L is the horizontal distance from recharge point to discharge point. The storage derived from the recession curve is what may be termed “dynamic storage.” It is the water stored above the local base level defined by the elevation of the spring. The deep storage below local base level is not probed by these measurements and requires wells of sufficient depth to penetrate the bottom of the aquifer.

21.6.6 Chemographs of Karst Springs

The chemical composition of karst springs varies depending on the hydrogeology of the specific spring and for some springs, the chemistry also varies with season and with storm flow. Plots of various chemical parameters as a function of time are referred to as “chemographs” in analogy to the term hydrograph. Any chemical parameter can be used to construct a chemograph but hardness and specific conductance are most commonly used. Specific conductance is especially valuable in that it is linearly proportional to the concentration of dissolved ions in carbonate aquifers and conductance probes and data loggers are both reliable and relatively inexpensive.

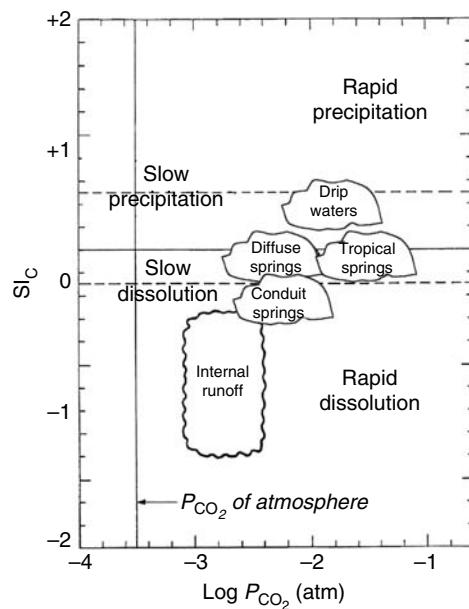


FIGURE 21.17 Schematic plot of karst water as a function of saturation index and calculated carbon dioxide partial pressure. The horizontal dashed lines indicate the saturation index at which carbonate dissolution shifts from fast kinetics to slow kinetics and the supersaturation at which nucleation barriers to calcite precipitation are exceeded.

The chemical compositions of ground and surface waters are commonly displayed as Piper or Stiff diagrams. These are useful when waters of widely varying composition are being compared but are less useful for karst waters, all of which are intrinsically carbonate type. The karst water parameters described in Section 21.4.2 are more useful as plotting variables. Figure 21.17 displays many of the chemical characteristics of karst waters. The enclosed areas labeled with various water types are generalized from a large body of data that have accumulated. Waters from various sources are often clearly distinguished by their chemistry.

When spring water chemistry is monitored over time periods of months to years, it is found that some spring waters have essentially no chemical variability while others undergo excursions of 25 to 50%, often in an erratic manner (Shuster and White, 1971). For drainage basins of modest size, in the range of tens of km^2 , the constant chemistry springs are mainly those discharging from fractures (called diffuse flow springs in the original report). Highly variable chemistry is associated with springs discharging from conduits and from drainage basins with large components of alloigenic recharge. Interpretation of data collected at weekly or longer intervals is limited. Continuous records, or point-by-point records that can map the chemistry over individual storm hydrographs are more useful.

Chemographs record events on many time scales (Figure 21.18). The chemograph of CO_2 pressure from a large basin reveals the seasonal oscillation due to the coming and going of the growing season. The concentration of Ca^{2+} ions measured on the Maramec Spring, Missouri, exhibits dips corresponding to individual storm events. Like the recession limb of the hydrograph, the chemograph has an exponential recovery tail. However, the recovery time for the chemistry is substantially longer than the hydrograph recession. Continuous chemographs recorded as specific conductance reveal details within individual storm events with time resolution as short as 15 min. This fine structure has been interpreted as the arrival of storm water from different tributaries of the conduit system.

Chemographs are most useful when they can be directly superimposed on hydrographs of individual storm events. The sudden decrease in specific conductance (hardness) is an indication of dilution of the spring water by storm water from the surface. If the decrease in conductance aligns with the rising limb

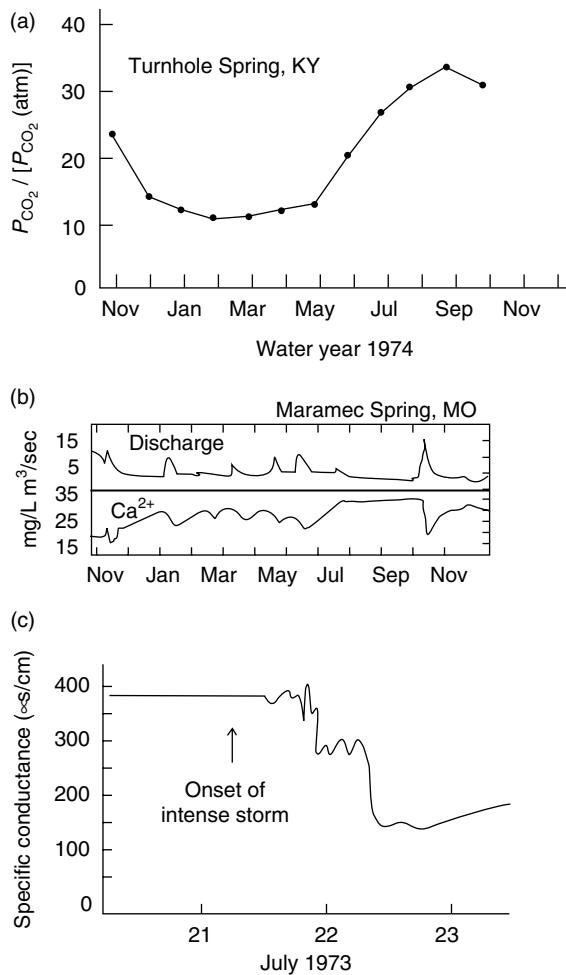


FIGURE 21.18 Chemographs for selected karst springs. (a) CO₂ partial pressure normalized to atmospheric CO₂, Turnhole Spring, Kentucky. (Data from Hess, J.W. and White, W.B. (1993) *Appl. Geochem.* 8, 189–204.) (b) Relation of Ca²⁺ ion concentration to discharge, Maramec Spring, Missouri. (Adapted from Dreiss, S.J. (1989) *Water Resour. Res.* 25, 117–125). (c) Response of specific conductance to a single intense storm, Turnhole Spring, Kentucky. (Adapted from Hess, J.W. and White, W.B. (1986) *J. Hydrol.* 99, 235–252.)

of the hydrograph, it is an indication that the system is open and the rise in the hydrograph is due to the arrival of storm water at the spring. If, as is often observed, there is a lag between the rising limb of the hydrograph and the falling limb of the chemograph, the indication is that water stored in continuously water-filled portions of the conduit system is being forced out by rising hydraulic head in the recharge region. Combining the lag time with the hydrograph allows calculation of the volume of water displaced (Ryan and Meiman, 1996). Comparison of chemographs and hydrographs have been used as a means of hydrograph separation to distinguish storm flow from phreatic storage and other sources of water appearing at the spring during storm discharge (Padilla et al., 1994; Grasso and Jeannin, 2002; Birk et al., 2004). Profiles of oxygen isotope ratios can also serve as a chemograph (Winston and Criss, 2004).

Chemographs of other parameters can be superimposed on storm hydrographs (Figure 21.19). In the example shown, the falling limb of the conductance curve is synchronous with the rising limb of the hydrograph. However, the CO₂ pressure, initially low and remaining low through the peak of the hydrograph, rises on the recession limb of the hydrograph. This can be interpreted as the arrival of soil water from the epikarst which brings in higher CO₂ concentrations to the receding storm flow.

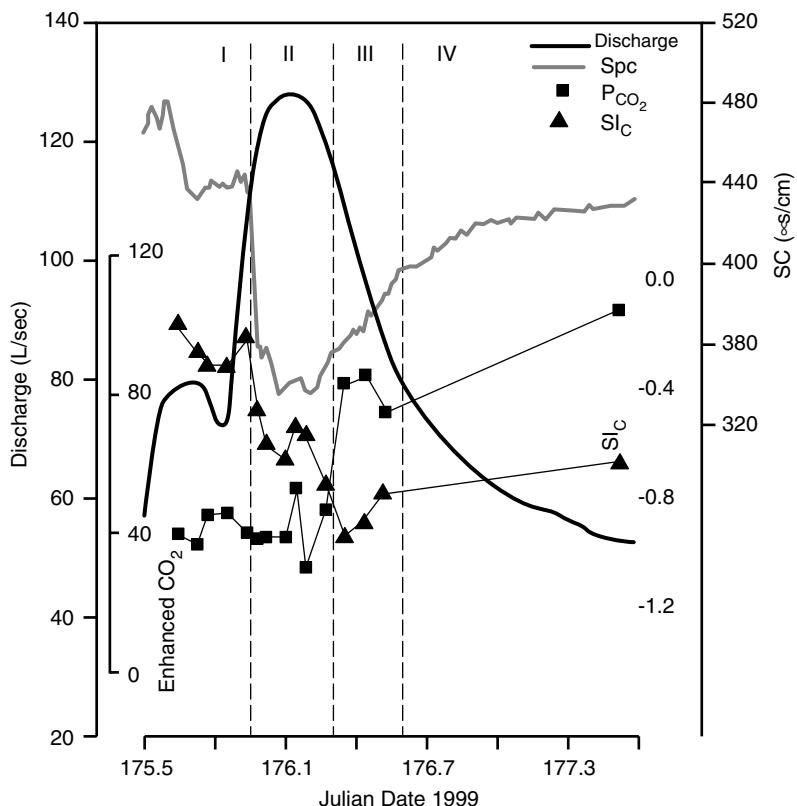


FIGURE 21.19 Hydrograph and selection of chemographs for a single storm at Beaver Spring, Fort Campbell Army Base, western Kentucky. (Adapted from Vesper, D.J. and White, W.B. (2004) *Hydrogeol. J.* 12, 135–143.)

21.7 Modeling of Karst Aquifers

The word “model” is used here in the sense of a quantitative description of some aspect of karst aquifers as distinguished from the various conceptual models presented earlier in this chapter. Karst aquifer models come in two varieties. There are models that attempt to describe the evolution of the aquifer or groundwater basin through time as dissolution of the bedrock progresses and conduit permeability is developed. The second variety of model is concerned with the recharge, flow, and discharge of groundwater through the aquifer, assuming a time scale sufficiently short that the permeability distribution may be taken as fixed. These models differ from models used for non-karstic aquifers in that they must take into account the breakdown of Darcy’s law in the conduit permeability and the extreme contrast in hydraulic conductivity in different parts of the aquifer. Karst aquifer modeling has been a very active field of research and the sections that follow only briefly mention some of the approaches that have been used. For a broader perspective, see Palmer et al. (1999) and Sasowsky and Wicks (2000).

21.7.1 Evolutionary Models

The qualitative concepts underlying evolutionary models were discussed earlier. The models that have been developed have been mainly a matter of reducing the qualitative ideas to computer programs. The usual starting point is a network of fractures for which the aperture distribution and connectivity are adjustable parameters. Some equation for the dissolution kinetics of carbonate rock is inserted and standard fluid mechanics is used to describe fracture flow and flow in conduits. These models show the evolution of

the fracture and conduit system as a function of time with the objective of replicating permeability patterns seen in real aquifers (Kaufmann and Braun, 1999; Kaufmann, 2002, 2004; Gabrovsek et al., 2004).

21.7.2 Equivalent Porous Media Models

Models adapted from standard porous media models such as the MODFLOW program developed by the United States Geological Survey assume that at large enough scales, the heterogeneities of the karst aquifer are smoothed out and can be represented by an average hydraulic conductivity. Scanlon et al. (2003) had some success in applying this type of model to the Edwards Aquifer in Texas. Equivalent porous media models work best for aquifers in which the karstic flow paths are dispersed and consist mainly of solutionally widened fractures. They work least well for aquifers with well-developed conduit systems, particularly those with large inputs of alloigenic recharge. Establishing a suitable value for the average hydraulic conductivity is also difficult because hydraulic conductivity is scale dependent. For some spring basins in Missouri, the hydraulic conductivity was found to increase with scale up to a certain limit after which it became scale independent (Rovey, 1994). For highly karstic aquifers, the hydraulic conductivity continues to increase with scale until the scale length becomes equal to the size of the groundwater basin (Sauter, 1991). In general, equivalent porous media models should be used with great care.

21.7.3 Pipe Flow Models

Equivalent porous media models ignore the conduit systems and their localized turbulent flows. Pipe Flow models focus entirely on the conduit system. Pipe flow models treat the conduit system as a network of pipes. Flow through the pipes is described by the Darcy–Weisbach equation (Equation 21.7). The drawback of the pipe flow models is that conduits are not exactly pipes. They have varying cross sections, complicated interconnections, and are often further modified by breakdown and sediment infillings. Much detail about conduit pattern is needed. However, calculations of travel times, head losses, and discharges for known conduit systems have been generally successful (Halihan and Wicks, 1998; Jeannin, 2001).

21.7.4 Coupled Continuum Pipe Flow Models

Realistic models for maturely developed karst aquifers must take account of both the conduit system which carries most of the flow and the fracture and matrix system which holds most of the storage (Worthington et al., 2000). The conduit system is strongly coupled to surface water through sinking streams and closed depressions and so has a very dynamic response to storms. The fracture and matrix system receives most of its recharge through the epikarst and has a slower response. A dominant component of the total flow system is the exchange flow between the conduits and the surrounding fractured matrix. During storm flow, heads in the conduit system rise rapidly and water is forced into the surrounding fractures. After the storm flow recedes, the conduit system drains rapidly, heads reverse, and the water stored in the fractures drains into the conduits.

Models have been constructed in which the fracture and matrix system is described as a continuum with Darcy flow. The conduits are put in “by hand” and described by pipe flow models. There is an exchange term that describes the influx and outflux of water between the fracture system and the conduits. These models have worked well in the sense that they reproduce storm hydrographs measured at the spring (Mohrlok and Sauter, 1997). The drawback is that independent information is required on the distribution of conduits. Tracer studies and cave exploration must supplement the strictly model calculation. Continued development of this approach to modeling has been very promising (Bauer et al., 2003; Liedl et al., 2003).

21.7.5 Input–Output Models

All of the modeling approaches described above are built on considerable knowledge of the aquifer being modeled — the geometry of the conduit system, the hydraulic conductivity of the fracture and matrix

system, and any geologic boundary conditions. In general, the more knowledge available, the more accurate the model. The diametrically opposite approach is to treat the aquifer as a black box and assume nothing about its internal properties. Instead, models are built from inputs and outputs, both of which can be measured directly. The relationship between input and output is described by linear systems theory as pioneered by Dreiss (1982, 1989) with more recent applications described by Wicks and Hoke (1999).

The relationship between input of water to the aquifer and output of water as discharge from springs is given by:

$$O(t) = \int_{-\infty}^{\infty} h(t - \tau) i(\tau) d\tau \quad (21.45)$$

The function $O(t)$ is the output — either a hydrograph or a chemograph. The function $i(\tau)$ is the input of the same variable but not necessarily on the same timescale. The quantity $h(t - \tau)$ is the kernal function that describes what goes on inside the black box. The kernal function must be calibrated by measuring both inputs and outputs for a range of discharges. Once the kernal function is known, it can then be used to relate output to input for future behavior of the aquifer.

21.8 Water Resource Issues in Karst Hydrology

21.8.1 Water Supplies in Karst

Water supplies in karst are obtained from springs and wells. Karst springs' draining fracture systems, as indicated by uniform discharge, lack of turbidity, and small chemical variation, provide reliable water supplies generally of good quality except for high hardness. Springs' draining conduit systems have a highly variable discharge, sometimes become turbid to muddy after-storms, and are at risk from pollutants introduced through sinking streams and sinkholes. The conduit systems offer very little filtration. Injected pollutants can be quickly transmitted to the spring with little purification except for dilution.

Location of production wells in karst is challenging except for wells in young limestones (e.g., the Floridan) with a high matrix permeability. Fractured limestone and dolomite aquifers provide the most reliable supplies of good quality water if the wells are drilled on fracture intersections. In favorable localities these can be determined by mapping lineaments seen on air photographs (Lattman and Parizek, 1964). Locations of wells in aquifers dominated by conduit permeability is difficult because conduits cannot generally be located from the surface. If the conduit is an accessible cave, wells can be located by use of a magnetic induction device (cave "radio") from within the cave. Wells drilled into conduits face the same uncertainty of discharge, turbidity, and pollution risk as conduit springs. Farm wells are sometimes placed in conduits; municipal water supply wells are usually not placed in conduits.

21.8.2 Flood Flows in Karst: Sinkhole Flooding

Karst aquifers serve to modify the flood flows in the associated surface water basins. Flood pulses of intermediate magnitude in tributary streams are injected into the conduit system at swallets. The conduit system holds the water in temporary storage and bleeds it out more slowly through the spring. The response time of even the most flashy conduit system is usually longer than the response time of flood pulses in surface channels. This has the effect of clipping the top of the hydrograph and spreading it out, thus reducing the flood crest downstream (E. L. White and Reich, 1970).

Extreme floods may override swallets and send flood waters down dry valleys where there is no stream channel. In developed areas, buildings constructed in dry swales and valleys may be damaged by the unexpected invasion of flood waters.

Sinkhole terrain is prone to flooding. One mechanism is that storm water ponds in closed depressions if the drain is plugged with clay. A second and more important mechanism is flooding from below. Storm flow into the conduit system quickly fills the groundwater trough and turns it into a groundwater mound. If the temporary water table created by flooding of the conduit system rises above the level of

the bottoms of closed depressions, water will be forced up sinkhole drains and create ponds. High level sinkhole flooding may occur only during exceptional storms. If the sinkholes are broad and shallow, urban development may spread across them. All is well until the exceptional storm causes water to rise from below (Quinlan, 1986).

21.8.3 Contaminants and Contaminant Transport in Karst Aquifers

Karst aquifers are highly vulnerable to groundwater contamination because of the negligible filtering of surface runoff into sinkholes and water from sinking streams. Contaminants can travel long distances constrained to the conduit systems rather than spreading to form plumes. The contaminants themselves are the same as those that impact other aquifers but some unique threats arise because of the flow paths in karst.

Fecal material from septic tanks, leaky sewer lines, barnyard, and manure spread on fields migrates downward to the epikarst and then through the fractures in the vadose zone to the aquifer below. Cave streams and conduit springs frequently contain coliform bacteria and high nitrate levels. Agricultural chemicals including insecticides and herbicides have been reported in karst groundwater in many farming areas in the United States (Katz et al., 2004; Panno and Kelly, 2004). Household waste water becomes a contaminant in karst area. Pharmaceuticals are now being found in karst waters (Wicks et al., 2003).

Sinkholes have long been used for solid waste disposal in rural areas. Household waste, junk, paint and chemical containers and dead, often diseased, farm animals end up in sinkholes. These wastes are leached by runoff into the sinkholes and carried down the sinkhole drain into the underlying conduit system. Solid waste can be flushed down open drains and dispersed along the conduit system where it continues to leach, but is impossible to clean up. Contaminants can be incorporated into the clastic sediments and flushed by flood flows deep into the aquifer. Likewise, trash discarded into sinking surface streams can be carried into the conduit system by flood flows.

Liquid wastes from improper disposal of solvents, from industrial outfalls directed into fractures and sinkholes, and from leaking pipelines and storage tanks can make their way into karst aquifers though any of the recharge routes. Soluble wastes dissolve in the karst groundwater and are carried along with it. Insoluble liquids with a density less than water (LNAPLs) float on underground streams or on the karst waters table. Insoluble liquids with a density greater than water (DNAPLs) sink to the bottom of the karst groundwater system and may continue downward into fractures and indeed travel in directions other than that of the main groundwater flow.

The most common LNAPLs are gasoline, diesel fuel, and heating oil from leaking underground storage tanks. Toxic substances — the BTEX components benzene, toluene, ethyl benzene, and xylene — are found in LNAPL hydrocarbons. Because LNAPLs float, they tend to pond behind sumps in the conduit system and tend to be trapped in pockets and other irregularities in the conduit system (Ewers et al., 1991). Rising flood waters can force the LNAPLs upward through fractures and sinkhole drains. They may invade basements and crawl spaces beneath buildings where they produce both toxic fumes and also an explosion hazard (Stroud et al., 1986).

DNAPLs consist mainly of chlorinated hydrocarbons used as solvents and degreasers. When such materials reach the conduit system, they may pool in low places in stream channels or be adsorbed into sediment piles. DNAPLs have a finite, although usually low, solubility so that the pooled liquid is eventually dissolved in the conduit water and carried away although this may take a long time. DNAPLs in sediment piles and in the fractures in the limestone may remain in place for very long times (Loop and White, 2001). DNAPLs spilled into karst aquifers often do not reappear at springs and wells. Storm flows may eject a slug of DNAPL that had remained sequestered in the aquifer.

The movement of suspended sediments through karst aquifers is an important mechanism for contaminant transport. Metals tend to adsorb onto clay particles and are carried through the system with the suspended particles. For the carbonate aquifer at Fort Campbell in western Kentucky, there was a close correlation between concentrations of nickel, lead, chromium, cadmium and arsenic with the concentration of aluminum in the analysis of unfiltered water samples (Vesper and White, 2003). Aluminum

was taken as a proxy for suspended clay minerals. Likewise, bacteria have been found to be transported through karst aquifers attached to suspended particles (Mahler et al., 2000).

21.8.4 Water Quality Monitoring

United States Environmental Protection Agency and state regulation require legally established landfills and other disposal sites to install a network of monitoring wells. Samples are drawn from these wells at prescribed intervals and analyzed for contaminants. A typical requirement is three down-gradient wells and one up-gradient well to be installed on the perimeter of the waste disposal site. The strategy is to detect any contamination before it moves offsite and to take early remedial action. The concept does not work in karst. Because of the heterogeneous distribution of conduit permeability and the general impossibility of locating subsurface conduits from the land surface, there is a very real possibility of contaminants moving offsite through the karst conduits while leaving no trace in the monitoring wells. As contaminants, once in the conduit system, move rapidly, detection and remedial action must be prompt.

To monitor waste sites in karst requires a different strategy. Trenches should be excavated to bedrock and tracers injected into the bedrock fractures with a water flush. The spring or springs to which the aquifer drains must be identified and used as a monitoring site regardless of the difficulties of offsite monitoring. Likewise, more frequent monitoring, especially following storm events, is necessary because of the rapid transport of contaminants through the conduit system (Quinlan, 1990).

21.9 Land Instability and Sinkhole Development

21.9.1 Sinkholes as Land Use Hazards

The term “sinkhole” is used for any closed depression in karst terrain. In this broad usage, sinkholes range in size from a few meters to kilometers in diameter and from fractions of a meter to kilometers in depth. Processes include bedrock dissolution, bedrock stoping and collapse of underlying solution cavities, and soil collapses resulting from piping. The sinkholes that are of main concern as land-use hazards are of the soil piping variety. Unlike processes of dissolution, which are slow on human time scales, soil piping failures are abrupt with time scales in the range of seconds to days.

Residents of karst may be dismayed to discover soil collapses — sinkholes — meters in diameter and meters deep appearing in their yards or, worse, under their houses after heavy rains or after the thawing of frozen ground. These sinkholes occur in two end-member varieties: (1) cover collapse sinkholes that form abruptly as a plug of soil falls into an underlying cavity and (2) raveling sinkholes that form gradually as loose, poorly cohesive soils drain away into the subsurface. The size and depth of sinkholes depend on the thickness of the regolith and also on the cohesive strength of soil and regolith. Damage to buildings, paved areas, and highways can be extensive with some sinkholes of sufficient size to swallow entire buildings.

21.9.2 Mechanisms of Sinkhole Development

The essential steps in the development of a cover collapse sinkhole are shown in Figure 21.20. In the pre-sinkhole stage there is an opening in the bedrock, which may be a solutionally widened fracture or a tubular throat that connects at the bottom with a conduit system that contains flowing water. Above is a cohesive soil of some specified thickness. The throat is assumed to be plugged with soil. The throat may be the drain for a larger area of the epikarst. As infiltrating water descends the throat, the clastic sediment is gradually flushed into the underlying conduit that must have sufficient flow velocity, at least during storm flow, to carry the material away. When the throat has been cleared, sediment from the base of the regolith column is also flushed down the throat, creating a cavity. As more and more of the regolith is carried into the subsurface, the cavity and its regolith arch roof continue to grow. The final step in the process is the collapse of the regolith arch. When the cavity reaches a critical size, the roof becomes mechanically unstable. Collapse is usually instantaneous with the fallen plug of roof material blocking the throat.

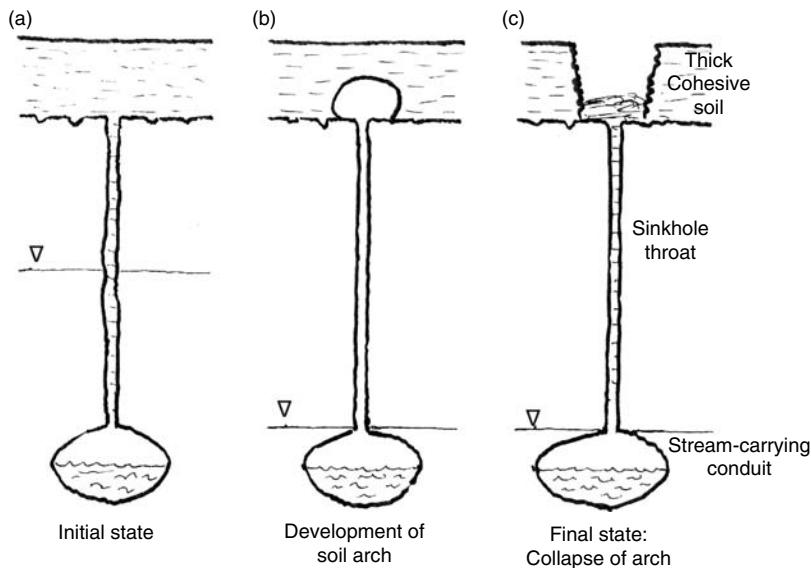


FIGURE 21.20 Components of a cover collapse sinkhole. (a) Initial conditions of thick, cohesive soil over karsted limestone. (b) Lowering water table increases hydraulic gradient so that storm runoff down the sinkhole throat flushes soil into the underlying cave passage and creates a soil arch. (c) The soil arch reaches a size where it becomes mechanically unstable. The remaining plug of soil near the surface collapses, forming a sinkhole.

21.9.3 Hydrologic Factors in Sinkhole Development

Sinkholes are a natural part of landscape evolution. Closed depressions are natural routes by which internal runoff reaches karst aquifers. Soils accumulate in the closed depressions and piping failures occur. When failures happen, a plug of soil is flushed into the underlying conduit system, the walls of the sinkhole slump and the throat reseals, more soil accumulates, and the cycle repeats. These sinkholes are found in rural areas, far from any human influence.

Excess runoff from roofs, parking lots, and other pavement concentrates the amount of water injected into the ground into a few locations. Increased hydrostatic heads and increased flow volumes focused on the pathways at the base of the epikarst accelerates the transport of regolith down sinkhole throats and thus the development of sinkholes. Detention basins for the management of storm runoff are particularly susceptible to the creation of sinkholes. Broken water mains or broken sewer lines are a common local cause of sinkholes.

Dewatering of karst aquifers also accelerates sinkhole development. If at least the lower portion of the throat is below the water table, sediment choking the throat will be water-saturated and less mobile. If the water table is lowered by excessive pumping from the aquifer, by quarry dewatering, or by other means, sediment plugs will drain and dry, hydraulic gradients will increase, and the system will be more susceptible to sinkhole development. Quarry dewatering has been a common source of sinkhole problems in surrounding areas.

21.9.4 Sinkhole Remediation

Sinkholes are part of hydrologic system and for this reason must be repaired with some care. Often sinkholes are repaired by simply filling them with soil or other material. This is seldom a long term solution. Plugging one sinkhole is likely to initiate another sinkhole somewhere in the vicinity as a drainage route for surface water must be maintained. A useful technique is to excavate the sinkhole to bedrock, expose the throat, then fill the bottom of the excavation with a stone too large to fit into the throat. Continue upward with progressively less coarse stone and gravel. The top can be capped with a geofabric and then

with soil. The objective is to provide a pathway that will allow free drainage of water but which will block the movement of soil and regolith. Many ingenious schemes have been devised for sinkhole remediation as discussed in detail by Sowers (1996) and Waltham et al. (2005).

References

- Anthony, D.M. and Granger, D.E. (2004) A late Tertiary origin for multilevel caves along the western escarpment of the Cumberland Plateau, Tennessee and Kentucky, established by cosmogenic ^{26}Al and ^{10}Be . *J. Cave Karst Stud.* 66, 46–55.
- Atkinson, T.C. (1977) Diffuse flow and conduit flow in limestone terrain in the Mendip Hills, Somerset (Great Britain). *J. Hydrol.* 35, 93–110.
- Atkinson, T.C., Smart, P.L. and Wigley, T.M.L. (1983) Climate and natural radon levels in Castleguard Cave, Columbia Icefields, Alberta, Canada. *Arc. Alpine Res.* 15, 487–502.
- Barnes, H.H., Jr. (1967) *Roughness Characteristics of Natural Channels*, U.S. Geological Survey Water Supply Paper 1849, Washington, D.C.
- Bauer, S., Liedl, R. and Sauter, M. (2003) Modeling of karst aquifer genesis: Influence of exchange flow. *Water Resour. Res.* 39, doi:10.1029/2003WR002218.
- Berner, R.A. and Morse, J.W. (1974) Dissolution kinetics of calcium carbonate in sea water IV. Theory of calcite dissolution. *Amer. J. Sci.* 274, 108–154.
- Birk, S., Liedl, R. and Sauter, M. (2004) Identification of localized recharge and conduit flow by combined analysis of hydraulic and physico-chemical spring responses (Urenbrunnen, SW-Germany). *J. Hydrol.* 286, 179–193.
- Bonaci, O. (1987) *Karst Hydrology*, Springer-Verlag, Berlin.
- Bosch, R.F. and White, W.B. (2004) Lithofacies and transport of clastic sediments in karstic aquifers. In *Studies of Cave Sediments*, Sasowsky, I.D. and Mylroie, J.E., Eds., Kluwer Academic/Plenum Publishers, New York, pp. 1–22.
- Brucker, R.W., Hess, J.W. and White, W.B. (1972) Role of vertical shafts in movement of groundwater in carbonate aquifers. *Ground Water* 10, 5–13.
- Brush, D.J. and Thomson, N.R. (2003) Fluid flow in synthetic rough-walled fractures: Navier–Stokes, Stokes, and local cubic law simulations. *Water Resour. Res.* 39, doi:10.1029/2002WR001346.
- Budd, D.A. and Vacher, H.L. (2002) Facies control on matrix permeability in the upper Floridan Aquifer, west-central Florida. *Karst Waters Inst. Spec. Pub.* 7, 14–24.
- Burdon, D.J. and Papakis, N. (1963) *Handbook of Karst Hydrogeology*, United Nations Special Fund, Institute for Geology and Subsurface Research, Athens, Greece.
- Busenberg, E. and Plummer, L.N. (1982) The kinetics of dissolution of dolomite in $\text{CO}_2\text{--H}_2\text{O}$ systems at 1.5 to 65°C and 0 to 1 atm P_{CO_2} . *Amer. J. Sci.* 282, 45–78.
- Curl, R.L. (1974) Deducing flow velocity in cave conduits from scallops. *Natl Speleol. Soc. Bull.* 36, 1–5.
- Davies, W.E. and Chao, E.C.T. (1958) Report on sediments in Mammoth Cave, Kentucky. *Administrative Report to National Park Service by U.S. Geological Survey*, 117 pp.
- Despain, J.D. and Stock, G.M. (2005) Geomorphic history of Crystal Cave, southern Sierra Nevada, California. *J. Cave Karst Studies* 67, 92–102.
- Dreiss, S.J. (1982) Linear kernels for karst aquifers. *Water Resour. Res.* 18, 865–876.
- Dreiss, S.J. (1989) Regional scale transport in a karst aquifer. 1. Component separation of spring flow hydrographs. *Water Resour. Res.* 25, 117–125.
- Drever, J.I. (1997) *The Geochemistry of Natural Waters*, 3rd ed. Prentice Hall, Englewood Cliffs, NJ.
- Dreybrodt, W. (1988) *Processes in Karst Systems*, Springer-Verlag, Berlin.
- Dreybrodt, W. and Buhmann, D. (1991) A mass transfer model for the dissolution and precipitation of calcite from solutions in turbulent motion. *Chem. Geol.* 90, 107–122.
- Dreybrodt, W. (1992) Dynamics of karstification: A model applied to hydraulic structures in karst terranes. *Appl. Hydrogeol.* 3, 20–32.

- Dreybrodt, W., Lauckner, J., Liu, Z., Svensson, U. and Buhman, D. (1996) The kinetics of the reaction $\text{CO}_2 + \text{H}_2\text{O} \rightarrow \text{H}^+ + \text{HCO}_3^-$ as one of the rate limiting steps of the dissolution of calcite in the system $\text{H}_2\text{O}-\text{CO}_2-\text{CaCO}_3$. *Geochim. Cosmochim. Acta* 60, 3375–3381.
- Dreybrodt, W. (1996) Principles of early development of karst conduits under natural and man-made conditions revealed by mathematical analysis of numerical models. *Water Resources Res.* 32, 2923–2935.
- Ewers, R.L., Estes, E.K., Idstein, P.J. and Johnson, K.M. (1991) The transmission of light hydrocarbon contaminants in limestone (karst) aquifers. *Proc. Third Conf. Hydrogeology, Ecology, Monitoring and Management of Ground Water in Karst Terranes*, Nashville, TN, pp. 287–306.
- Ford, D.C. and Williams, P.W. (1989) *Karst Geomorphology and Hydrology*, Unwin Hyman, London.
- Gabrovsek, F., Romanov, D. and Dreybrodt, W. (2004) Early karstification in a dual-fracture aquifer: The role of exchange flow between prominent fractures and a dense net of fissures. *J. Hydrol.* 299, 45–66.
- Gale, S.J. (1984) The hydraulics of conduit flow in carbonate aquifers. *J. Hydrol.* 70, 309–327.
- Gillieson, D. (1986) Cave sedimentation in the New Guinea Highlands. *Earth Surf. Proc. Landforms* 11, 533–543.
- Granger, D.E., Fabel, D. and Palmer, A.N. (2001) Pliocene-Pleistocene incision of the Green River, Kentucky, determined from radioactive decay of cosmogenic ^{26}Al and ^{10}Be in Mammoth Cave sediments. *Geol. Soc. Amer. Bull.* 113, 825–836.
- Grasso, D.A. and Jeannin, P.-Y. (2002) A global experimental system approach of karst springs' hydrographs and chemographs. *Ground Water* 40, 608–617.
- Groves, C.G. and Howard, A.D. (1994) Early development of karst systems. 1. Preferential flow path enlargement under laminar flow. *Water Resour. Res.* 30, 2837–2846.
- Halihan, T. and Wicks, C.M. (1998) Modeling of storm responses in conduit flow aquifers with reservoirs. *J. Hydrol.* 208, 82–91.
- Herman, J.S. and White, W.B. (1985) Dissolution kinetics of dolomite: Effects of lithology and fluid flow velocity. *Geochim. Cosmochim. Acta* 49, 2017–2026.
- Hess, J.W. and White, W.B. (1986) Storm response of the karstic carbonate aquifer in Southcentral Kentucky. *J. Hydrol.* 99, 235–252.
- Hess, J.W. and White, W.B. (1989) Water budget and physical hydrology. In *Karst Hydrology: Concepts from the Mammoth Cave Area*, White, W.B. and White, E.L., Eds, Van Nostrand Reinhold, New York, Chap. 4, pp. 105–126.
- Hess, J.W. and White, W.B. (1993) Groundwater geochemistry of the carbonate karst aquifer, Southcentral Kentucky, U.S.A. *Appl. Geochem.* 8, 189–204.
- Hill, C.A. (1990) Sulfuric acid speleogenesis of Carlsbad Caverns and its relationship to hydrocarbons, Delaware Basin, New Mexico and Texas. *Amer. Assoc. Petrol. Geol. Bull.* 74, 1685–1694.
- Jeannin, P.Y. (2001) Modeling flow in phreatic and epiphreatic karst conduits in the Höllöch Cave (Muotatal, Switzerland). *Water Resour. Res.* 37, 191–200.
- Jones, W.K. (1984) Dye tracer tests in karst areas. *Natl. Speleol. Soc. Bull.* 46, 3–9.
- Jones, W.K., Culver, D.C. and Herman, J.S. (2003) *EpiKarst*, Spec. Pub. 9, Karst Waters Institute, Charles Town, WV.
- Käss, W. (1998) *Tracing Technique in Geohydrology*, A.A. Balkema, Rotterdam.
- Katz, B.G., Chelette, A.R. and Pratt, T.R. (2004) Use of chemical and isotopic tracers to assess nitrate contamination and ground-water age, Woodville Karst Plain, USA. *J. Hydrol.* 289, 36–61.
- Kaufmann, G. and Braun, J. (1999) Karst aquifer evolution in fractured rocks. *Water Resour. Res.* 35, 3223–3238.
- Kaufmann, G. (2002) Karst aquifer evolution in a changing water table environment. *Water Resour. Res.* 38, doi:10.1029/2001WR000259.
- Kaufmann, G. (2003) Modelling unsaturated flow in an evolving karst aquifer. *J. Hydrol.* 276, 53–70.
- Konzuk, J.S. and Kueper, B.H. (2004) Evaluation of cubic law based models describing single-phase flow through a rough-walled fracture. *Water Resour. Res.* 40, doi:10.1029/2003WR002356.

- Langmuir, D. (1971) The geochemistry of some carbonate ground waters in central Pennsylvania. *Geochim. Cosmochim. Acta* 35, 1023–1045.
- Langmuir, D. (1997) *Aqueous Environmental Geochemistry*, Prentice Hall, Upper Saddle River, NJ.
- Lattman, L.H. and Parizek, R.R. (1964) Relationship between fracture traces and the occurrence of groundwater in carbonate rocks. *J. Hydrol.* 2, 73–91.
- Lauritzen, S.E., Abbott, J., Arnesen, R., Crossley, G., Grepperud, D., Ive, A. and Johnson, S. (1985) Morphology and hydraulics of an active phreatic conduit. *Cave Sci.* 12, 139–146.
- Liedl, R., Sauter, M., Hückinghaus, D., Clemens, T. and Teutsch, G. (2003) Simulation of the development of karst aquifers using a coupled continuum pipe model. *Water Resour. Res.* 39, doi:10.1029/2001WR001206.
- Liu, Z. and Dreybrodt, W. (1997) Dissolution kinetics of calcium carbonate minerals in $\text{H}_2\text{O}-\text{CO}_2$ solutions in turbulent flow: The role of the diffusion boundary layer and the slow reaction $\text{H}_2\text{O} + \text{CO}_2 \rightarrow \text{H}^+ + \text{HCO}_3^-$. *Geochim. Cosmochim. Acta* 61, 2879–2889.
- Loewenthal, R.E. and Marais, G.V.R. (1976) *Carbonate Chemistry of Aquatic Systems: Theory and Application*, Ann Arbor Science, Ann Arbor, MI.
- Loop, C.M. and White, W.B. (2001) A conceptual model for DNAPL transport in karst aquifers. *Ground Water* 39, 119–127.
- Mahler, B.J., Personné, J.-C., Lods, G.F. and Drogue, C. (2000) Transport of free and particulate-associated bacteria in karst. *J. Hydrol.* 238, 179–193.
- Milanović, P.T. (1981) *Karst Hydrogeology*, Water Resources Publications, Littleton, CO.
- Milanović, P.T. (2004) *Water Resources Engineering in Karst*, CRC Press, Boca Raton, FL.
- Mohrlok, U. and Sauter, M. (1997) Modelling groundwater flow in a karst terrane using discrete and double-continuum approaches — importance of spatial and temporal distribution of recharge. *Proc. 12th Internat. Congress Speleol.*, La Chaux de Fonds, Switzerland, Vol. 2, pp. 167–170.
- Morse, J.W. and Arvidson, R.S. (2002) The dissolution kinetics of major sedimentary carbonate minerals. *Earth-Sci. Rev.* 58, 51–84.
- Muldoon, M.A., Simo, J.A. and Bradbury, K.R. (2001) Correlation of hydraulic conductivity with stratigraphy in a fractured-dolomite aquifer, northeastern Wisconsin, U.S.A. *Hydrogeol. J.* 9, 570–583.
- Mull, D.S., Liebermann, T.D., Smoot, J.L. and Woosley, L.H. Jr. (1988) Application of dye-tracing techniques for determining solute-transport characteristics of groundwater in karst terranes. U.S. Environmental Protection Agency Report, EPA 9045/6-88-001.
- Mylroie, J.E. and Carew, J.L. (1987) Field evidence of the minimum time for speleogenesis. *Natl Speleol. Soc. Bull.* 49, 67–72.
- Padilla, A., Pulido-Bosch, A. and Mangin, A. (1994) Relative importance of baseflow and quickflow from hydrographs of karst springs. *Ground Water* 32, 267–277.
- Palmer, A.N. (1975) Origin of maze caves. *Natl Speleol. Soc. Bull.* 37, 57–76.
- Palmer, A.N. (1991) Origin and morphology of limestone caves. *Geol. Soc. Amer. Bull.* 103, 1–21.
- Palmer, A.N., Palmer, M.V. and Sasowsky, I.D. (1999) *Karst Modeling*, Spec. Pub. 5, Karst Waters Institute, Charles Town, WV.
- Panno, S.V. and Kelly, W.R. (2004) Nitrate and herbicide loading in two groundwater basins of Illinois' sinkhole plain. *J. Hydrol.* 290, 229–242.
- Plummer, L.N., Wigley, T.M.L. and Parkhurst, D.L. (1978) The kinetics of calcite dissolution in CO_2 -water systems at 5° to 60°C and 0.0 to 1.0 atm CO_2 . *Amer. J. Sci.* 278, 179–216.
- Plummer, L.N., Parkhurst, D.L. and Wigley, T.M.L. (1979) Critical review of the kinetics of calcite dissolution and precipitation. In *Chemical Modeling in Aqueous Systems*, Jenne, E.A., Ed., Amer. Chem. Soc. Symposium Ser. 93, Washington, DC, Chap. 25, pp. 537–573.
- Plummer, L.N. and Busenberg, E. (1982) The solubilities of calcite, aragonite and vaterite in $\text{CO}_2-\text{H}_2\text{O}$ solutions between 0 and 90°C, and an evaluation of the aqueous model for the system $\text{CaCO}_3-\text{CO}_2-\text{H}_2\text{O}$. *Geochim. Cosmochim. Acta* 46, 1011–1040.

- Quinlan, J.F. (1986) Legal aspects of sinkhole development and flooding in karst terranes: 1. Review and synthesis. *Environ. Geol. Water Sci.* 8, 41–61.
- Quinlan, J.F. (1990) Special problems of ground-water monitoring in karst terranes. In *Ground Water and Vadose Zone Monitoring*, Nielsen, D.M. and Johnson, A.I. Eds., ASTM STP 1053, American Society for Testing Materials, Philadelphia, PA, pp. 275–304.
- Quinlan, J.F. and Ray, J.A. (1995) Normalized base-flow discharge of groundwater basins: A useful parameter for estimating recharge area of springs and for recognizing drainage anomalies in karst terranes. In *Karst Geohazards*, Beck, B.F. Ed., A.A. Balkema, Rotterdam, pp. 149–164.
- Richards, D.A. and Dorale, J.A. (2003) Uranium-series chronology and environmental applications of speleothems. *Rev. Mineral.* 52, 407–460.
- Rovey, C.W. II. (1994) Assessing flow systems in carbonate aquifers using scale effects in hydraulic conductivity. *Environ. Geol.* 24, 244–253.
- Ryan, M. and Meiman, J. (1996) An examination of short-term variations in water quality at a karst spring in Kentucky. *Ground Water*. 34, 23–30.
- Sasowsky, I.D. and Wicks, C.M. (2000) *Groundwater Flow and Contaminant Transport in Carbonate Aquifers*, A.A. Balkema, Rotterdam.
- Sasowsky, I.D. and Mylroie, J. (2004) *Studies of Caves Sediments*, Kluwer Academic/Plenum Press, New York.
- Sauter, M. (1991) Assessment of hydraulic conductivity in a karst aquifer at local and region scale. *Proc. Third Conf. of Hydrogeology, Ecology, Monitoring, and Management of Ground Water in Karst Terranes*, Nashville, TN, pp. 39–57.
- Scanlon, B.R., Mace, R.E., Barrett, M.E. and Smith, B. (2003) Can we simulate regional groundwater flow in a karst system using equivalent porous media models? Case study, Barton Springs, Edwards Aquifer, USA. *J. Hydrol.* 276, 137–158.
- Schmidt, V.A. (1982) Magnetostratigraphy of sediments in Mammoth Cave, Kentucky. *Science* 217, 827–829.
- Shuster, E.T. and White, W.B. (1971) Seasonal fluctuations in the chemistry of limestone springs: A possible means for characterizing carbonate aquifers. *J. Hydrol.* 14, 93–128.
- Siemers, J. and Dreybrodt, W. (1998) Early development of karst aquifers on percolation networks of fractures in limestone. *Water Resour. Res.* 34, 409–419.
- Smart, C.C. (1988) Quantitative tracing of the Maligne Karst System, Alberta, Canada. *J. Hydrol.* 98, 185–204.
- Smart, P.L. and Laidlaw, I.M.S. (1977) An evaluation of some fluorescent dyes for water tracing. *Water Resour. Res.* 13, 15–33.
- Smart, P.L. (1984) A review of the toxicity of twelve fluorescent dyes used for water tracing. *Natl Speleol. Soc. Bull.* 46, 21–33.
- Sowers, G.F. (1996) *Building on Sinkholes*, American Society of Civil Engineers, New York, NY.
- Stroud, F.B., Gilbert, J., Powell, G.W., Crawford, N.C., Rigatti, M.J. and Johnson, P.C. (1986) Environmental Protection Agency emergency response to toxic fumes and contaminated ground-water in karst topography, Bowling Green, Kentucky. *Proc. Environmental Problems in Karst Terranes and Their Solutions Conference*, Bowling Green, KY, pp. 197–226.
- Svensson, U. and Dreybrodt, W. (1992) Dissolution kinetics of natural calcite minerals in CO₂–water systems approaching calcite equilibrium. *Chem. Geol.* 100, 129–145.
- Vanoni, V.A. (1975) *Sedimentation Engineering*, American Society of Civil Engineers, New York.
- Vennard, J.K. (1962) *Elementary Fluid Mechanics*, Wiley, New York.
- Vesper, D.J. and White, W.B. (2003) Metal transport to karst springs during storm flow: An example from Fort Campbell, Kentucky/Tennessee, U.S.A. *J. Hydrol.* 276, 20–36.
- Vesper, D.J. and White, W.B. (2004) Storm pulse chemographs of saturation index and carbon dioxide pressure: Implications for shifting recharge sources during storm events in the karst aquifer at Fort Campbell, Kentucky/Tennessee, USA. *Hydrogeol. J.* 12, 135–143.

- Waltham, T., Bell, F. and Culshaw, M. (2005) *Sinkholes and Subsidence: Karst and Cavernous Rocks in Engineering and Construction*, Springer-Verlag, Berlin.
- White, E.L. and Reich, B.M. (1970) Behavior of annual floods in limestone basins in Pennsylvania. *J. Hydrol.* 10, 193–198.
- White, E.L. (1977) Sustained flow in small Appalachian watersheds underlain by carbonate rocks. *J. Hydrol.* 32, 71–86.
- White, W.B. (1977) Role of solution kinetics in the development of karst aquifers. In *Karst Hydrogeology*, Tolson, J.S. and Doyle, F.L., Eds., Internat'l. Assoc. Hydrogeologists Memoir 12, pp. 503–517.
- White, W.B. (1988) *Geomorphology and Hydrology of Karst Terrains*, Oxford University Press, New York.
- White, W.B. and Deike, G.H. (1989) Hydraulic geometry of cave passages, In *Karst Hydrology: Concepts from the Mammoth Cave Area*, White, W.B. and White, E.L., Eds., Van Nostrand Reinhold, New York, Chap. 9, pp. 223–258.
- White, W.B. (1999) Conceptual models for karstic aquifers. *Karst Waters Inst. Spec. Pub.* 5, 10–16.
- Wicks, C.M. and Hoke, J.A. (1999) Linear systems approach to modeling groundwater flow and solute transport through karstic basins. *Karst Waters Inst. Spec. Pub.* 5, 97–101.
- Wicks, C.M., Kelley, C. and Peterson, E. (2003) Estrogen in a karstic aquifer. *Ground Water*. 42, 384–389.
- Winston, W.E. and Criss, R.E. (2004) Dynamic hydrologic and geochemical response in a perennial karst spring. *Water Resour. Res.* 40, doi:10.1029/2004WR003054.
- Witherspoon, P.A., Wang, J.S.Y., Iwai, K. and Gale, J.E. (1980) Validity of cubic law for fluid flow in a deformable rock fracture. *Water Resour. Res.*, 16, 1016–1024.
- Worthington, S.R.H. (1991) Karst hydrogeology in the Canadian Rocky Mountains. Ph.D. thesis, McMaster University, Hamilton, Ontario.
- Worthington, S.R.H. (1999) A comprehensive strategy for understanding flow in carbonate aquifers. *Karst Waters Inst. Spec. Pub.* 5, 30–37.
- Worthington, S.R.H., Ford, D.C. and Davies, G.J. (2000) Matrix, fracture and channel components of storage and flow in a Paleozoic limestone aquifer. In *Groundwater Flow and Contaminant Transport in Carbonate Aquifers*, Sasowsky, I.D. and Wicks, C.M., Eds. A.A. Balkema, Rotterdam, pp. 113–128.
- Worthington, S.R.H., Schindel, G.M. and Alexander, E.C. (2002) Techniques for investigating the extent of karstification in the Edwards Aquifer, Texas. *Karst Waters Inst. Spec. Pub.* 7, 173–175.

Further Information

It has been estimated that 15 to 20% of the Earth's land surface is karst. As a result, there is a very large literature in many languages. The study of karst has traditionally been the province of physical geographers and geomorphologists and most books approach the subject from this point of view. Useful books that emphasize the hydrogeologic aspects of karst include Bonacci (1987), White (1988), Dreybrodt (1988), Ford and Williams (1989), and Milanović (1981, 2004).

22

Contaminant Transport in the Unsaturated Zone: Theory and Modeling

22.1	Introduction.....	22-1
22.2	Variably Saturated Water Flow	22-3
	Mass Balance Equation • Uniform Flow • Preferential Flow	
22.3	Solute Transport	22-8
	Transport Processes • Advection–Dispersion Equations •	
	Nonequilibrium Transport • Stochastic Models •	
	Multicomponent Reactive Solute Transport • Multiphase	
	Flow and Transport • Initial and Boundary Conditions	
22.4	Analytical Models	22-22
	Analytical Approaches • Existing Models	
22.5	Numerical Models	22-27
	Numerical Approaches • Existing Models	
22.6	Concluding Remarks	22-36
	References	22-38

Jiří Šimůnek
University of California, Riverside
Martinus Th. van
Genuchten
*George E. Brown, Jr. Salinity
Laboratory, USDA-ARS*

22.1 Introduction

Human society during the past several centuries has created a large number of chemical substances that often find their way into the environment, either intentionally applied during agricultural practices, unintentionally released from leaking industrial and municipal waste disposal sites, or stemming from research or weapons production related activities. As many of these chemicals represent a significant health risk when they enter the food chain, contamination of both surface and subsurface water supplies has become a major issue. Modern agriculture uses an unprecedented number of chemicals, both in plant and animal production. A broad range of fertilizers, pesticides and fumigants are now routinely applied to agricultural lands, thus making agriculture one of the most important sources for non-point source pollution. The same is true for salts and toxic trace elements, which are often an unintended consequence of irrigation in arid and semiarid regions. While many agricultural chemicals are generally beneficial in surface soils, their leaching into the deeper vadose zone and groundwater may pose serious problems. Thus, management processes are being sought to keep fertilizers and pesticides in the root

zone and prevent their transport into underlying or down-gradient water resources. Agriculture also increasingly uses a variety of pharmaceuticals and hormones in animal production many of which, along with pathogenic microorganisms, are being released to the environment through animal waste. Animal waste and wash water effluent, in turn, is frequently applied to agricultural lands. Potential concerns about the presence of pharmaceuticals and hormones in the environment include: (1) abnormal physiological processes and reproductive impairment; (2) increased incidences of cancer; (3) development of antibiotic resistant bacteria; and (4) increased toxicity of chemical mixtures. While the emphasis above is mostly on non-point source pollution by agricultural chemicals, similar problems arise with point-source pollution from industrial and municipal waste disposal sites, leaking underground storage tanks, chemicals spills, nuclear waste repositories, and mine tailings, among other sources.

Mathematical models should be critical components of any effort to optimally understand and quantify site-specific subsurface water flow and solute transport processes. For example, models can be helpful tools for designing, testing and implementing soil, water, and crop management practices that minimize soil and water pollution. Models are equally needed for designing or remediating industrial waste disposal sites and landfills, or for long-term stewardship of nuclear waste repositories. A large number of specialized numerical models now exist to simulate the different processes at various levels of approximation and for different applications.

Increasing attention is being paid recently to the unsaturated or vadose zone where much of the subsurface contamination originates, passes through, or can be eliminated before it contaminates surface and subsurface water resources. Sources of contamination often can be more easily remediated in the vadose zone, before contaminants reach the underlying groundwater. Other chapters in this Handbook deal with water flow (Chapter 4) and solute transport (Chapter 18 and Chapter 19) in fully saturated (groundwater) systems and with water flow in the unsaturated zone (Chapter 6). The focus of this chapter thus will be on mathematical descriptions of transport processes in predominantly variably saturated media.

Soils are generally defined as the biologically active layer at the surface of the earth's crust that is made up of a heterogeneous mixture of solid, liquid, and gaseous material, as well as containing a diverse community of living organisms (Jury and Horton, 2004). The vadose (unsaturated) zone is defined as the layer between the land surface and the permanent (seasonal) groundwater table. While pores between solid grains are fully filled with water in the saturated zone (groundwater), pores in the unsaturated zone are only partially filled with water, with the remaining part of the pore space occupied by the gaseous phase. The vadose zone is usually only partially saturated, although saturated regions may exist, such as when perched water is present above a low-permeable fine-textured (clay) layer or a saturated zone behind the infiltration front during or after a high-intensity rainfall event.

As the transport of contaminants is closely linked with the water flux in soils and rocks making up the vadose zone, any quantitative analysis of contaminant transport must first evaluate water fluxes into and through the vadose zone. Water typically enters the vadose zone in the form of precipitation or irrigation (Figure 22.1), or by means of industrial and municipal spills. Some of the rainfall or irrigation water may be intercepted on the leaves of vegetation. If the rainfall or irrigation intensity is larger than the infiltration capacity of the soil, water will be removed by surface runoff, or will accumulate at the soil surface until it evaporates back to the atmosphere or infiltrates into the soil. Part of this water is returned to the atmosphere by evaporation. Some of the water that infiltrates into the soil profile may be taken up by plant roots and eventually returned to the atmosphere by plant transpiration. The processes of evaporation and transpiration are often combined into the single process of evapotranspiration. Only water that is not returned to the atmosphere by evapotranspiration may percolate to the deeper vadose zone and eventually reach the groundwater table. If the water table is close enough to the soil surface, the process of capillary rise may move water from the groundwater table through the capillary fringe toward the root zone and the soil surface.

Because of the close linkage between water flow and solute transport, we will first briefly focus on the physics and mathematical description of water flow in the vadose zone (Section 22.2). An overview is given of the governing equations for water flow in both uniform (Section 22.2.2) and structured (Section 22.2.3)

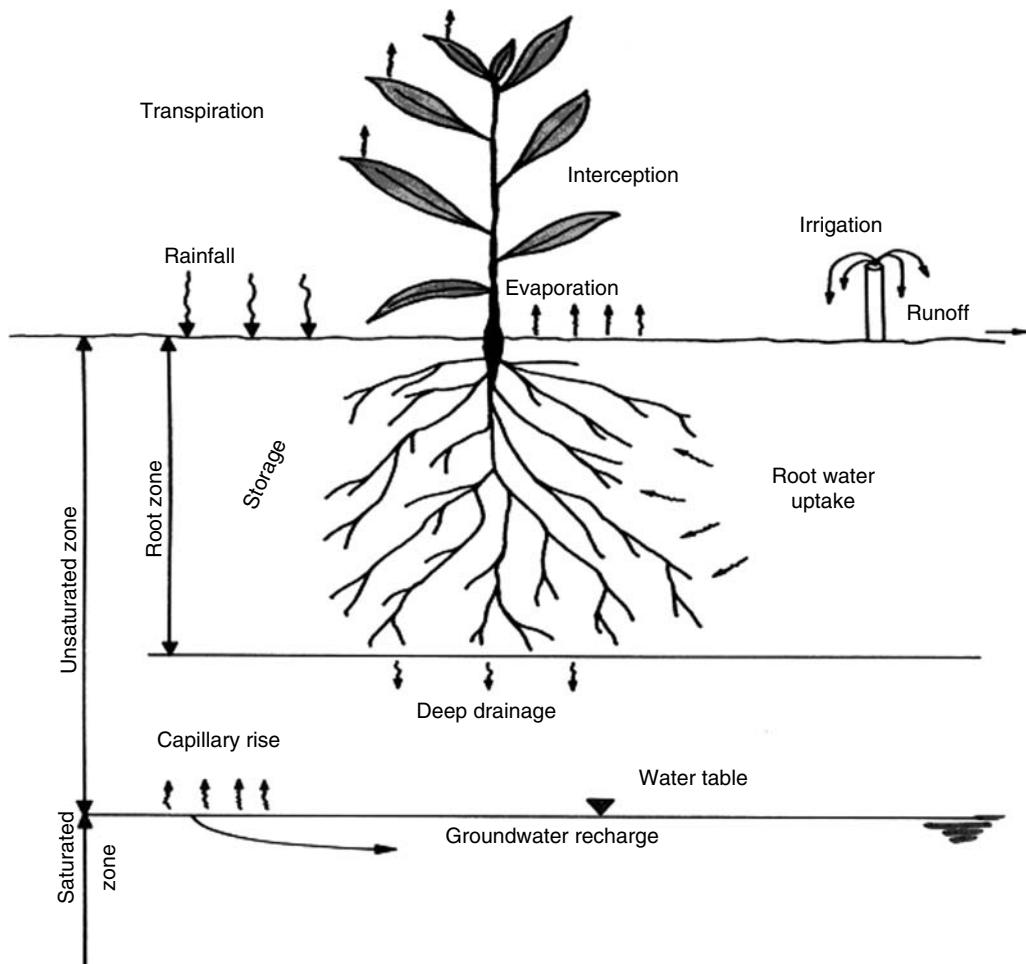


FIGURE 22.1 Schematic of water fluxes and various hydrologic components in the vadose zone.

media. This section is followed by a discussion of the governing solute transport equations (Section 22.3), again for both uniform (Section 22.3.2) and structured (fractured) (Section 22.3.3) media. We also briefly discuss alternative formulations for colloid (Section 22.3.3.2) and colloid-facilitated transport (Section 22.3.3.3), multicomponent geochemical transport (Section 22.3.5), and stochastic approaches for solute transport (Section 22.3.4). This is followed by a discussion of analytical (Section 22.4) and numerical (Section 22.5) approaches for solving the governing flow and transport equations, and an overview of computer models currently available for simulating vadose zone flow and transport processes (Sections 22.4.2 and 22.5.2).

22.2 Variably Saturated Water Flow

In this section, we briefly present the equations governing variably saturated water flow in the subsurface. More details about this topic, including the description of the soil hydraulic properties and their constitutive relationships, are given in Chapter 6. Traditionally, descriptions of variably saturated flow in soils are based on the Richards (1931) equation, which combines the Darcy–Buckingham equation for the fluid flux with a mass balance equation. The Richards equation typically predicts a uniform flow process in the vadose zone, although possibly modified macroscopically by spatially variable soil hydraulic

properties (e.g., as dictated by the presence of different soil horizons, but possibly also varying laterally). Unfortunately, the vadose zone can be extremely heterogeneous at a range of scales, from the microscopic (e.g., pore scale) to the macroscopic (e.g., field or larger scale). Some of these heterogeneities can lead to a preferential flow process that macroscopically is very difficult to capture with the standard Richards equation. One obvious example of preferential flow is the rapid movement of water and dissolved solutes through macropores (e.g., between soil aggregates, or created by earthworms or decayed root channels) or rock fractures, with much of the water bypassing (short-circuiting) the soil or rock matrix. However, many other causes of preferential flow exist, such as flow instabilities caused by soil textural changes or water repellency (Hendrickx and Flury, 2001; Šimůnek et al. 2003; Ritsema and Dekker 2005), and lateral funneling of water due to inclined or other textural boundaries (e.g., Kung 1990). Alternative ways of modeling preferential flow are discussed in a later section. Here we first focus on the traditional approach for uniform flow as described with the Richards equation.

22.2.1 Mass Balance Equation

Water flow in variably saturated rigid porous media (soils) is usually formulated in terms of a mass balance equation of the form:

$$\frac{\partial \theta}{\partial t} = -\frac{\partial q_i}{\partial x_i} - S \quad (22.1)$$

where θ is the volumetric water content [$L^3 L^{-3}$], t is time [T], x_i is the spatial coordinate [L], q_i is the volumetric flux density [LT^{-1}], and S is a general sink/source term [$L^3 L^{-3} T^{-1}$], for example, to account for root water uptake (transpiration). Equation 22.1 is often referred to as the mass conservation equation or the continuity equation. The mass balance equation in general states that the change in the water content (storage) in a given volume is due to spatial changes in the water flux (i.e., fluxes in and out of some unit volume of soil) and possible sinks or sources within that volume. The mass balance equation must be combined with one or several equations describing the volumetric flux density (q) to produce the governing equation for variably saturated flow. The formulations of the governing equations for different types of flow (uniform and preferential flow) are all based on this continuity equation.

22.2.2 Uniform Flow

Uniform flow in soils is described using the Darcy–Buckingham equation:

$$q_i = -K(h) \left(K_{ij}^A \frac{\partial h}{\partial x_j} + K_{iz}^A \right) \quad (22.2)$$

where K is the unsaturated hydraulic conductivity [LT^{-1}], and K_{ij}^A are components of a dimensionless anisotropy tensor \mathbf{K}^A (which reduces to the unit matrix when the medium is isotropic). The Darcy–Buckingham equation is formally similar to Darcy's equation, except that the proportionality constant (i.e., the unsaturated hydraulic conductivity) in the Darcy–Buckingham equation is a nonlinear function of the pressure head (or water content), while $K(h)$ in Darcy's equation is a constant equal to the saturated hydraulic conductivity, K_s (e.g., see discussion by Narasimhan [2005]).

Combining the mass balance Equation 22.1 with the Darcy–Buckingham Equation 22.2 leads to the general Richards equation (Richards, 1931)

$$\frac{\partial \theta(h)}{\partial t} = \frac{\partial}{\partial x_i} \left[K(h) \left(K_{ij}^A \frac{\partial h}{\partial x_j} + K_{iz}^A \right) \right] - S(h) \quad (22.3)$$

This partial differential equation is the equation governing variably saturated flow in the vadose zone. Because of its strongly nonlinear makeup, only a relatively few simplified analytical solutions can be derived. Most practical applications of Equation 22.3 require a numerical solution, which can be obtained

using a variety of numerical methods such as finite differences or finite elements (Section 22.5a). Equation 22.3 is generally referred to as the mixed form of the Richards equation as it contains two dependent variables, that is, the water content and the pressure head. Various other formulations of the Richards equation are possible.

22.2.3 Preferential Flow

Increasing evidence exists that variably saturated flow in many field soils is not consistent with the uniform flow pattern typically predicted with the Richards equations (Flury et al., 1994; Hendrickx and Flury, 2001). This is due to the presence of macropores, fractures, or other structural voids or biological channels through which water and solutes may move preferentially, while bypassing a large part of the matrix pore-space. Preferential flow and transport processes are probably the most frustrating in terms of hampering accurate predictions of contaminant transport in soils and fractured rocks. Contrary to uniform flow, preferential flow results in irregular wetting of the soil profile as a direct consequence of water moving faster in certain parts of the soil profile than in others. Hendrickx and Flury (2001) defined preferential flow as constituting all phenomena where water and solutes move along certain pathways, while bypassing a fraction of the porous matrix. Water and solutes for these reasons can propagate quickly to far greater depths, and much faster, than would be predicted with the Richards equation describing uniform flow.

The most important causes of preferential flow are the presence of macropores and other structural features, development of flow instabilities (i.e., fingering) caused by profile heterogeneities or water repellency (Hendrickx et al., 1993), and funneling of flow due to the presence of sloping soil layers that redirect downward water flow. While the latter two processes (i.e., flow instability and funneling) are usually caused by textural differences and other factors at scales significantly larger than the pore scale, macropore flow and transport are usually generated at the pore or slightly larger scales, including scales where soil structure first manifests itself (i.e., the pedon scale) (Šimůnek et al., 2003).

Uniform flow in granular soils and preferential flow in structured media (both macroporous soils and fractured rocks) can be described using a variety of single-porosity, dual-porosity, dual-permeability, multi-porosity, and multi-permeability models (Richards, 1931; Pruess and Wang, 1987; Gerke and van Genuchten, 1993a; Gwo et al., 1995; Jarvis, 1998; Šimůnek et al., 2003, 2005). While single-porosity models assume that a single pore system exists that is fully accessible to both water and solute, dual-porosity and dual-permeability models both assume that the porous medium consists of two interacting pore regions, one associated with the inter-aggregate, macropore, or fracture system, and one comprising the micropores (or intra-aggregate pores) inside soil aggregates or the rock matrix. Whereas dual-porosity models assume that water in the matrix is stagnant, dual-permeability models allow also for water flow within the soil or rock matrix.

Figure 22.2 illustrates a hierarchy of conceptual formulations that can be used to model variably saturated water flow and solute transport in soils. The simplest formulation (Figure 22.2a) is a single-porosity (equivalent porous medium) model applicable to uniform flow in soils. The other models apply in some form or another to preferential flow or transport. Of these, the dual-porosity model of Figure 22.2c assumes the presence of two pore regions, with water in one region being immobile and in the other region mobile. This model allows the exchange of both water and solute between the two regions (Šimůnek et al., 2003). Conceptually, this formulation views the soil as consisting of a soil matrix containing grains/aggregates with a certain internal microporosity (intra-aggregate porosity) and a macropore or fracture domain containing the larger pores (inter-aggregate porosity). While water and solutes are allowed to move through the larger pores and fractures, they can also flow in and out of aggregates. By comparison, the intra-aggregate pores represent immobile pockets that can exchange, retain, and store water and solutes, but do not contribute to advective (or convective) flow. Models that assume mobile-immobile flow regions (Figure 22.2b) are conceptually somewhere in between the single- and dual-porosity models. While these models assume that water will move similarly as in the uniform flow models, the liquid phase for purposes of modeling solute transport is divided in terms of mobile and

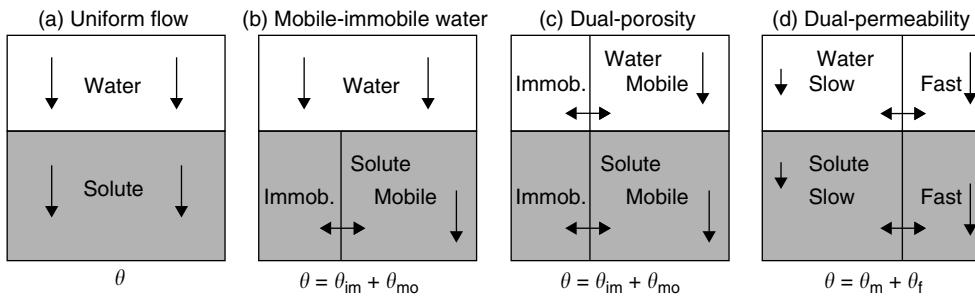


FIGURE 22.2 Conceptual models of water flow and solute transport (θ is the water content, θ_{mo} and θ_{im} in (b) and (c) are water contents in the mobile and immobile flow regions, respectively, and θ_{m} and θ_{f} in (d) are water contents in the matrix and macropore (fracture) regions, respectively).

immobile fractions, with solutes allowed to move by advection and dispersion only in the mobile fraction and between the two pore regions. This model has long been applied to solute transport studies (e.g., van Genuchten and Wierenga, 1976).

Finally, dual-permeability models (Figure 22.2d) are those in which water can move in both the inter- and intra-aggregate pore regions (and matrix and fracture domains). These models in various forms are now also becoming increasingly popular (Pruess and Wang, 1987; Gerke and van Genuchten, 1993a; Jarvis, 1994; Pruess, 2004). Available dual-permeability models differ mainly in how they implement water flow in and between the two pore regions (Šimůnek et al., 2003). Approaches to calculating water flow in the macropores or inter-aggregate pores range from those invoking Poiseuille's equation (Ahuja and Hebsom, 1992), the Green and Ampt or Philip infiltration models (Ahuja and Hebsom, 1992), the kinematic wave equation (Germann, 1985; Germann and Beven, 1985; Jarvis, 1994), and the Richards equation (Gerke and van Genuchten, 1993a). Multi-porosity and multi-permeability models (not shown in Figure 22.2) are based on the same concept as dual-porosity and dual-permeability models, but include additional interacting pore regions (e.g., Gwo et al., 1995; Hutson and Wagener, 1995). These models can be readily simplified to the dual-porosity/permeability approaches. Recent reviews of preferential flow processes and available mathematical models are provided by Hendrickx and Flury (2001) and Šimůnek et al. (2003), respectively.

22.2.3.1 Dual-Porosity Models

Dual-porosity models assume that water flow is restricted to macropores (or inter-aggregate pores and fractures), and that water in the matrix (intra-aggregate pores or the rock matrix) does not move at all. This conceptualization leads to two-region type flow and transport models (van Genuchten and Wierenga, 1976) that partition the liquid phase into mobile (flowing, inter-aggregate), θ_{mo} , and immobile (stagnant, intra-aggregate), θ_{im} , regions [$\text{L}^3 \text{L}^{-3}$]:

$$\theta = \theta_{\text{mo}} + \theta_{\text{im}} \quad (22.4)$$

The dual-porosity formulation for water flow can be based on a mixed formulation of the Richards Equation 22.3 to describe water flow in the macropores (the preferential flow pathways) and a mass balance equation to describe moisture dynamics in the matrix as follows (Šimůnek et al., 2003):

$$\begin{aligned} \frac{\partial \theta_{\text{mo}}(h_{\text{mo}})}{\partial t} &= \frac{\partial}{\partial x_i} \left[K(h_{\text{mo}}) \left(K_{ij}^A \frac{\partial h_{\text{mo}}}{\partial x_j} + K_{iz}^A \right) \right] - S_{\text{mo}}(h_{\text{mo}}) - \Gamma_w \\ \frac{\partial \theta_{\text{im}}(h_{\text{im}})}{\partial t} &= -S_{\text{im}}(h_{\text{im}}) + \Gamma_w \end{aligned} \quad (22.5)$$

where S_{im} and S_{mo} are sink terms for both regions [T^{-1}], and Γ_w is the transfer rate for water from the inter- to the intra-aggregate pores [T^{-1}].

Several of the above dual-porosity features were recently included in the HYDRUS software packages (Šimůnek et al., 2003, 2005). Examples of their application to a range of laboratory and field data involving transient flow and solute transport are given by Šimůnek et al. (2001), Zhang et al. (2004), Köhne et al. (2004a, 2005), Kodešová et al. (2005), and Haws et al. (2005).

22.2.3.2 Dual-Permeability Models

Different types of dual-permeability approaches may be used to describe flow and transport in structured media. While several models invoke similar governing equations for flow in the fracture and matrix regions, others use different formulations for the two regions. A typical example of the first approach is the work of Gerke and van Genuchten (1993a, 1996) who applied Richards equations to each of two pore regions. The flow equations for the macropore (fracture) (subscript f) and matrix (subscript m) pore systems in their approach are given by:

$$\frac{\partial \theta_f(h_f)}{\partial t} = \frac{\partial}{\partial x_i} \left[K_f(h_f) \left(K_{ij}^A \frac{\partial h_f}{\partial x_j} + K_{iz}^A \right) \right] - S_f(h_f) - \frac{\Gamma_w}{w} \quad (22.6)$$

and

$$\frac{\partial \theta_m(h_m)}{\partial t} = \frac{\partial}{\partial x_i} \left[K_m(h_m) \left(K_{ij}^A \frac{\partial h_m}{\partial x_j} + K_{iz}^A \right) \right] - S_m(h_m) + \frac{\Gamma_w}{1-w} \quad (22.7)$$

respectively, where w is the ratio of the volumes of the macropore (or fracture or inter-aggregate) domain and the total soil system [—]. This approach is relatively complicated in that the model requires characterization of water retention and hydraulic conductivity functions (potentially of different form) for both pore regions, as well as the hydraulic conductivity function of the fracture–matrix interface. Note that the water contents θ_f and θ_m in Equation 22.6 and Equation 22.7 have different meanings than in Equation 22.5 where they represented water contents of the total pore space (i.e., $\theta = \theta_{mo} + \theta_{im}$), while here they refer to water contents of the two separate (fracture or matrix) pore domains such that $\theta = w\theta_f + (1-w)\theta_m$.

22.2.3.3 Mass Transfer

The rate of exchange of water between the macropore and matrix regions, Γ_w , is a critical term in both the dual-porosity model (Equation 22.5) and the dual-permeability approach given by (Equation 22.6) and (Equation 22.7). Gerke and van Genuchten (1993a) assumed that the rate of exchange is proportional to the difference in pressure heads between the two pore regions:

$$\Gamma_w = \alpha_w (h_f - h_m) \quad (22.8)$$

in which α_w is a first-order mass transfer coefficient [T^{-1}]. For porous media with well-defined geometries, the first-order mass transfer coefficient, α_w , can be defined as follows (Gerke and van Genuchten, 1993b):

$$\alpha_w = \frac{\beta}{d^2} K_a \gamma_w \quad (22.9)$$

where d is an effective diffusion path length [L] (i.e., half the aggregate width or half the fracture spacing), β is a shape factor that depends on the geometry [—], and γ_w (= 0.4) is a scaling factor [—] obtained by matching the results of the first-order approach at the half-time level of the cumulative infiltration curve to the numerical solution of the horizontal infiltration equation (Gerke and van Genuchten, 1993b). Several other approaches based on water content of relative saturation differences have also been used (Šimůnek et al., 2003).

22.3 Solute Transport

Similarly, as shown in Equation 22.1 for water flow, mathematical formulations for solute transport are usually based on a mass balance equation of the form:

$$\frac{\partial C_T}{\partial t} = -\frac{\partial J_{Ti}}{\partial x_i} - \phi \quad (22.10)$$

where C_T is the total concentration of chemical in all forms [ML^{-3}], J_{Ti} is the total chemical mass flux density (mass flux per unit area per unit time) [$\text{ML}^{-2}\text{T}^{-1}$], and ϕ is the rate of change of mass per unit volume by reactions or other sources (negative) or sinks (positive) such as plant uptake [$\text{ML}^{-3}\text{T}^{-1}$]. In its most general interpretation, Equation 22.10 allows the chemical to reside in all three phases of the soil (i.e., gaseous, liquid, and solid), permits a broad range of transport mechanisms (including advective transport, diffusion, and hydrodynamic dispersion in both the liquid and gaseous phases), and facilitates any type of chemical reaction that leads to losses or gains in the total concentration.

While the majority of chemicals are present only in the liquid and solid phases, and as such are transported in the vadose zone mostly only in and by water, some chemicals such as many organic contaminants, ammonium, and all fumigants, can have a significant portion of their mass in the gaseous phase and are hence subject to transport in the gaseous phase as well. The total chemical concentration can thus be defined as:

$$C_T = \rho_b s + \theta c + a g \quad (22.11)$$

where ρ_b is the bulk density [ML^{-3}], θ is the volumetric water content [L^3L^{-3}], a is the volumetric air content [L^3L^{-3}], and $s[\text{MM}^{-1}]$, $c[\text{ML}^{-3}]$, and $g[\text{ML}^{-3}]$ are concentrations in the solid, liquid, and gaseous phases, respectively. The solid phase concentration represents solutes sorbed onto sorption sites of the solid phase, but can include solutes sorbed onto colloids attached to the solid phase or strained by the porous system, and solutes precipitated onto or into the solid phase.

The reaction term ϕ of Equation 22.10 may represent various chemical or biological reactions that lead to a loss or gain of chemical in the soil system, such as radionuclide decay, biological degradation, and dissolution. In analytical and numerical models these reactions are most commonly expressed using zero- and first-order reaction rates as follows:

$$\phi = \rho_b s \mu_s + \theta c \mu_w + a g \mu_g - \rho_b \gamma_s - \theta \gamma_w - a \gamma_g \quad (22.12)$$

where μ_s , μ_w , and μ_g are first-order degradation constants in the solid, liquid, and gaseous phases [T^{-1}], respectively, and γ_s [T^{-1}], γ_w [$\text{ML}^{-3}\text{T}^{-1}$], and γ_g [$\text{ML}^{-3}\text{T}^{-1}$] are zero-order production constants in the solid, liquid, and gaseous phases, respectively.

22.3.1 Transport Processes

When a solute is present in both the liquid and gaseous phase, then various transport processes in both of these phases may contribute to the total chemical mass flux:

$$J_T = J_l + J_g \quad (22.13)$$

where J_l and J_g represent solute fluxes in the liquid and gaseous phases [$\text{ML}^{-2}\text{T}^{-1}$], respectively. Note that in Equation 22.13, and further below, we omitted the subscript i accounting for the direction of flow. The three main processes that can be active in both the liquid and gaseous phase are molecular diffusion, hydrodynamic dispersion, and advection (often also called convection). The solute fluxes in the two phases

are then the sum of fluxes due to these different processes:

$$\begin{aligned} J_l &= J_{lc} + J_{ld} + J_{lh} \\ J_g &= J_{gc} + J_{gd} + J_{gh} \end{aligned} \quad (22.14)$$

where the additional subscripts c, d, and h denote convection (or advection), molecular diffusion, and hydrodynamic dispersion, respectively.

22.3.1.1 Diffusion

Diffusion is a result of the random motion of chemical molecules. This process causes a solute to move from a location with a higher concentration to a location with a lower concentration. Diffusive transport can be described using Fick's law:

$$\begin{aligned} J_{ld} &= -\theta \xi_l(\theta) D_l^w \frac{\partial c}{\partial z} = -\theta D_l^s \frac{\partial c}{\partial z} \\ J_{gd} &= -a \xi_g(\theta) D_g^w \frac{\partial g}{\partial z} = -a D_g^s \frac{\partial g}{\partial z} \end{aligned} \quad (22.15)$$

where D_l^w and D_g^w are binary diffusion coefficients of the solute in water and gas [L^2T^{-1}], respectively; D_l^s and D_g^s are the effective diffusion coefficients in soil water and soil gas [L^2T^{-1}], respectively; and ξ_l and ξ_g are tortuosity factors that account for the increased path lengths and decreased cross-sectional areas of the diffusing solute in both phases (Jury and Horton, 2004). As solute diffusion in soil water (air) is severely hampered by both air (water) and solid particles, the tortuosity factor increases strongly with water content (air content). Many empirical models have been suggested in the literature to account for the tortuosity (e.g., Moldrup et al., 1998). Among these, the most widely used model for the tortuosity factor is probably the equation of Millington and Quirk (1961) given by:

$$\xi_l(\theta) = \frac{\theta^{7/3}}{\theta_s^2} \quad (22.16)$$

where θ_s is the saturated water content (porosity) [L^3L^{-3}]. A similar equation may be used for the tortuosity factor of the gaseous phase by replacing the water content with the air content.

22.3.1.2 Dispersion

Dispersive transport of solutes results from the uneven distribution of water flow velocities within and between different soil pores (Figure 22.3). Dispersion can be derived from Newton's law of viscosity which states that velocities within a single capillary tube follow a parabolic distribution, with the largest velocity in the middle of the pore and zero velocities at the walls (Figure 22.3a). Solutes in the middle of a pore, for this reason, will travel faster than solutes that are farther from the center. As the distribution of solute ions within a pore depends on their charge, as well as on the charge of pore walls, some solutes may move significantly faster than others. In some situations (i.e., for negatively charged anions in fine-textured soils, leading to anion exclusion), the solute may even travel faster than the average velocity of water (e.g., Nielsen et al., 1986). Using Poiseuille's law, one can further show that velocities in a capillary tube depend strongly on the radius of the tube, and that the average velocity increases with the radius to the second power. As soils consist of pores of many different radii, solute fluxes in pores of different radii will be significantly different, with some solutes again traveling faster than others (Figure 22.3b).

The above pore-scale dispersion processes lead to an overall (macroscopic) hydrodynamic dispersion process that mathematically can be described using Fick's law in the same way as molecular diffusion, that is,

$$J_{lh} = -\theta D_{lh} \frac{\partial c}{\partial z} = -\theta \lambda v \frac{\partial c}{\partial z} = -\lambda q \frac{\partial c}{\partial z} \quad (22.17)$$

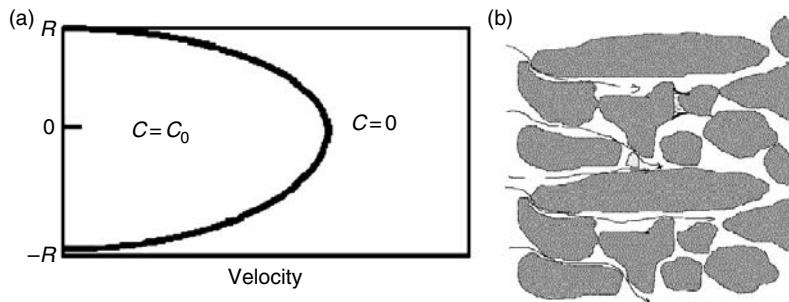


FIGURE 22.3 Distribution of velocities in a single capillary (a) and distribution of velocities in a more complex pore system (b).

where D_{lh} is the hydrodynamic dispersion coefficient [$L^2 T^{-1}$], v is the average pore-water velocity [LT^{-1}], and λ is the dispersivity [L]. The dispersion coefficient in one-dimensional systems has been found to be approximately proportional to the average pore-water velocity, with the proportionality constant generally referred as the (longitudinal) dispersivity (Biggar and Nielsen, 1967). The discussion above holds for one-dimensional transport; multi-dimensional applications require a more complicated dispersion tensor involving longitudinal and transverse dispersivities (e.g., Bear, 1972).

Dispersivity is a transport property that is relatively difficult to measure experimentally. Estimates are usually obtained by fitting measured breakthrough curves with analytical solutions of the advection-dispersion equation (discussed further below). The dispersivity often changes with the distance over which solute travels. Values of the longitudinal dispersivity typically range from about 1 cm for packed laboratory columns, to about 5 or 10 cm for field soils. Longitudinal dispersivities can be significantly larger (on the order of hundreds of meters) for regional groundwater transport problems (Gelhar et al., 1985). If no other information is available, a good first approximation is to use a value of one-tenth of the transport distance for the longitudinal dispersivity (e.g., Anderson, 1984), and a value of one-hundreds of the transport distance for the transverse dispersivity.

22.3.1.3 Advection

Advective transport refers to solute being transported with the moving fluid, either in the liquid phase (J_{lc}) or the gas phase (J_{gc}), that is,

$$\begin{aligned} J_{lc} &= qc \\ J_{gc} &= J_g g \end{aligned} \quad (22.18)$$

where J_g is the gaseous flux density [LT^{-1}]. **Advective transport in the gaseous phase is often neglected as its contribution in many applications is negligible compared to gaseous diffusion.**

The total solute flux density in both the liquid and gaseous phases is obtained by incorporating contributions from the various transport processes into Equation 22.14 to obtain

$$\begin{aligned} J_l &= qc - \theta D_l^s \frac{\partial c}{\partial z} - \theta D_{lh} \frac{\partial c}{\partial z} = qc - \theta D_e \frac{\partial c}{\partial z} \\ J_g &= -a D_g^s \frac{\partial g}{\partial z} \end{aligned} \quad (22.19)$$

where D_e is the effective dispersion coefficient [$L^2 T^{-1}$] that accounts for both diffusion and hydrodynamic dispersion. Dispersion in most subsurface transport problems dominates molecular diffusion in the liquid phase, except when the fluid velocity becomes relatively small or negligible. Notice that Equation 22.19 neglects advective and dispersive transport in the gaseous phase.

22.3.2 Advection–Dispersion Equations

22.3.2.1 Transport Equations

The equation governing transport of dissolved solutes in the vadose zone is obtained by combining the solute mass balance (Equation 22.10) with equations defining the total concentration of the chemical (Equation 22.11) and the solute flux density (Equation 22.19) to give

$$\frac{\partial(\rho_b s + \theta c + ag)}{\partial t} = \frac{\partial}{\partial x_i} \left(\theta D_{eij} \frac{\partial c}{\partial x_j} \right) + \frac{\partial}{\partial x_i} \left(a D_{gij}^s \frac{\partial g}{\partial x_j} \right) - \frac{\partial(q_i c)}{\partial x_i} - \phi \quad (22.20)$$

Notice that this equation is again written for multidimensional transport, and that D_{eij} and D_{gij}^s are thus components of the effective dispersion tensor in the liquid phase and a diffusion tensor in the gaseous phase [$L^2 T^{-1}$], respectively.

Many different variants of Equation 22.20 can be found in the literature. For example, for one-dimensional transport of nonvolatile solutes, the equation simplifies to

$$\frac{\partial(\rho_b s + \theta c)}{\partial t} = \frac{\partial(\theta R c)}{\partial z} = \frac{\partial}{\partial z} \left(\theta D_e \frac{\partial c}{\partial z} \right) - \frac{\partial(q c)}{\partial z} - \phi \quad (22.21)$$

where q is the vertical water flux density [LT^{-1}] and R is the retardation factor [–]

$$R = 1 + \frac{\rho_b}{\theta} \frac{ds(c)}{dc} \quad (22.22)$$

For transport of inert, nonadsorbing solutes during steady-state water flow we obtain

$$\frac{\partial c}{\partial t} = D_e \frac{\partial^2 c}{\partial z^2} - v \frac{\partial c}{\partial z} \quad (22.23)$$

The above equations are usually referred to as advection–dispersion equations (ADEs).

22.3.2.2 Linear and Nonlinear Sorption

The ADE given by Equation 22.20 contains three unknown concentrations (those for the liquid, solid, and gaseous phases), while Equation 22.21 contains two unknowns. To be able to solve these equations, additional information is needed that somehow relates these concentrations to each other. The most common way is to assume instantaneous sorption and to use adsorption isotherms to relate the liquid and adsorbed concentrations. The simplest form of the adsorption isotherm is the linear isotherm given by

$$s = K_d c \quad (22.24)$$

where K_d is the distribution coefficient [$L^3 M^{-1}$]. One may verify that substitution of this equation into Equation 22.22 leads to a constant value for the retardation factor (i.e., $R = 1 + \rho_b K_d / \theta$).

While the use of a linear isotherm greatly simplifies the mathematical description of solute transport, sorption and exchange are generally nonlinear and most often depend also on the presence of competing species in the soil solution. The solute retardation factor for nonlinear adsorption is not constant, as is the case for linear adsorption, but changes as a function of concentration. Many models have been used in the past to describe nonlinear sorption. The most commonly used nonlinear sorption models are those by Freundlich (1909) and Langmuir (1918) given by:

$$s = K_f c^\beta \quad (22.25)$$

$$s = \frac{K_d c}{1 + \eta c} \quad (22.26)$$

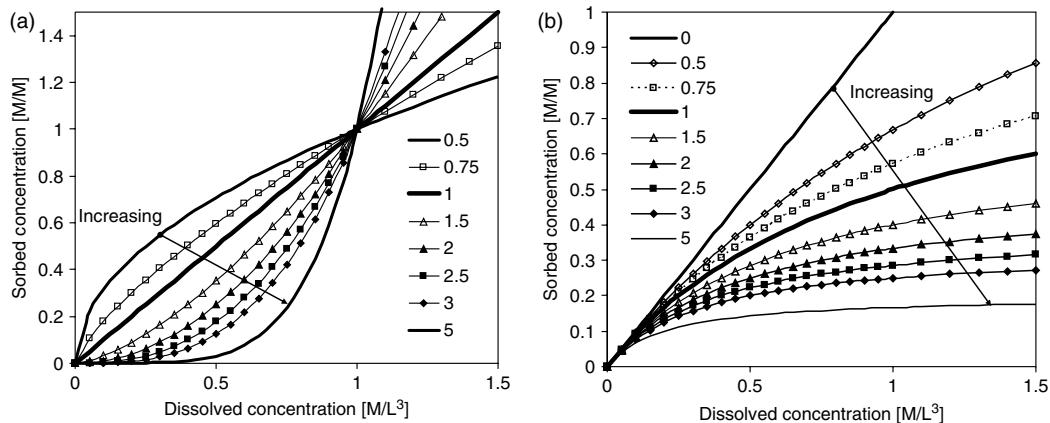


FIGURE 22.4 Plots of the Freundlich adsorption isotherm given by (22.25), with $K_d = 1$ and β given in the caption (a), and the Langmuir adsorption isotherm given by (22.26), with $K_d = 1$ and η given in the caption (b).

TABLE 22.1 Equilibrium Adsorption Equations adapted from van Genuchten and Šimunek (1996)

Equation	Model	Reference
$s = k_1 c + k_2$	Linear	Lapidus and Amundson (1952) Lindstrom et al. (1967)
$s = \frac{k_1 c^{k_3}}{1 + k_2 c^{k_3}}$	Freundlich–Langmuir	Sips (1950), Šimunek et al. (1994, 2005)
$s = \frac{k_1 c}{1 + k_2 c} + \frac{k_3 c}{1 + k_4 c}$	Double Langmuir	Shapiro and Fried (1959)
$s = k_1 c^{k_2/k_3}$	Extended Freundlich	Sibbesen (1981)
$s = \frac{k_1 c}{1 + k_2 c + k_3 \sqrt{c}}$	Gunary	Gunary (1970)
$s = k_1 c^{k_2} - k_3$	Fitter–Sutton	Fitter and Sutton (1975)
$s = k_1 \{1 - [1 + k_2 c^{k_3}]^{k_4}\}$	Barry	Barry (1992)
$s = \frac{RT}{k_1} \ln(k_2 c)$	Temkin	Bache and Williams (1971)
$s = k_1 c \exp(-2k_2 s)$		Lindstrom et al. (1971) van Genuchten et al. (1974)
$\frac{s}{s_T} = c[c + k_1(c_T - c) \exp\{k_2(c_T - 2c)\}]^{-1}$	Modified Kielland	Lai and Jurinak (1971)

k_1, k_2, k_3, k_4 : empirical constants; R : universal gas constant; T : absolute temperature; c_T : maximum solute concentration; s_T : maximum adsorbed concentration.

Source: Adapted from van Genuchten, M.Th. and Šimunek, J. In P.E. Rijtema and V. Eliáš (Eds.), *Regional Approaches to Water Pollution in the Environment*, NATO ASI Series: 2. Environment. Kluwer, Dordrecht, The Netherlands, pp. 139–172, 1996.

respectively, where $K_f [M^{-\beta} L^{3\beta}]$ and $\beta [-]$ are coefficients in the Freundlich isotherm, and $\eta [L^3 M^{-1}]$ is a coefficient in the Langmuir isotherm. Examples of linear, Freundlich and Langmuir adsorption isotherms are given in Figure 22.4. Table 22.1 lists a range of linear and other sorption models frequently used in solute transport studies.

22.3.2.3 Volatilization

Volatilization is increasingly recognized as an important process affecting the fate of many organic chemicals, including pesticides, fumigants, and explosives in field soils (Jury et al., 1983, 1984; Glotfelter and Schomburg, 1989). While many organic pollutants dissipate by means of chemical and microbiological degradation, volatilization may be equally important for volatile substances, such as certain pesticides. The volatility of pesticides is influenced by many factors, including the physicochemical properties of the chemical itself as well as such environmental variables as temperature and solar energy. Even though only a small fraction of a pesticide may exist in the gas phase, air-phase diffusion rates can sometimes be comparable to liquid-phase diffusion as gas-phase diffusion is about four orders of magnitude greater than liquid phase diffusion. The importance of gaseous diffusion relative to other transport processes depends also on the climate. For example, while transport of MTBE (gasoline oxygenate) is generally dominated by liquid advection in humid areas, gaseous diffusion may be equally or more important in arid climates; this even though only about 2% of MTBE may be in the gas phase.

The general transport equation given by Equation 22.20 can be simplified considerably when assuming linear equilibrium sorption and volatilization such that the adsorbed (*s*) and gaseous (*g*) concentrations are linearly related to the solution concentration (*c*) through the distribution coefficients, K_d in (22.24) and K_H , that is,

$$g = K_H c \quad (22.27)$$

respectively, where K_H is the dimensionless Henry's constant [−]. Equation 22.20 for one-dimensional transport then has the form:

$$\frac{\partial(\rho_b K_d + \theta + aK_H)c}{\partial t} = \frac{\partial}{\partial z} \left(\theta D_E \frac{\partial c}{\partial z} \right) + \frac{\partial}{\partial z} \left(aD_g^s K_H \frac{\partial c}{\partial z} \right) - \frac{\partial(qc)}{\partial x} - \phi \quad (22.28)$$

or

$$\frac{\partial \theta R c}{\partial t} = \frac{\partial}{\partial z} \left(\theta D_E \frac{\partial c}{\partial z} \right) - \frac{\partial(qc)}{\partial x} - \phi \quad (22.29)$$

where the retardation factor R [−] and the effective dispersion coefficient D_E [$L^2 T^{-1}$] are defined as follows:

$$R = 1 + \frac{\rho_b K_d + aK_H}{\theta} \quad (22.30)$$

$$D_E = D_E + \frac{aD_g^s K_H}{\theta}$$

Jury et al. (1983, 1984) provided for many organic chemicals their distribution coefficients K_d , Henry's constants K_H , and calculated percent mass present in each phase.

22.3.3 Nonequilibrium Transport

As equilibrium solute transport models often fail to describe experimental data, a large number of diffusion-controlled physical nonequilibrium and chemical-kinetic models have been proposed and used to describe the transport of both non-adsorbing and adsorbing chemicals. Attempts to model nonequilibrium transport usually involve relatively simple first-order rate equations. Nonequilibrium models have used the assumptions of two-region (dual-porosity) type transport involving solute exchange between mobile and immobile liquid transport regions, and one-, two- or multi-site sorption formulations (e.g., Nielsen et al., 1986; Brusseau, 1999). Models simulating the transport of particle-type pollutants, such as colloids, viruses, and bacteria, often also use first-order rate equations to describe such processes as attachment, detachment, and straining. Nonequilibrium models generally have resulted in better descriptions of observed laboratory and field transport data, in part by providing additional degrees of freedom for fitting observed concentration distributions.

22.3.3.1 Physical Nonequilibrium

22.3.3.1.1 Dual-Porosity and Mobile–Immobile Water Models

The two-region transport model (Figures 22.2b and Figure 22.2c) assumes that the liquid phase can be partitioned into distinct mobile (flowing) and immobile (stagnant) liquid pore regions, and that solute exchange between the two liquid regions can be modeled as a first-order exchange process. Using the same notation as before, the two-region solute transport model is given by (van Genuchten and Wagenet, 1989; Toride et al., 1993):

$$\begin{aligned}\frac{\partial \theta_{\text{mo}} c_{\text{mo}}}{\partial t} + \frac{\partial f \rho s_{\text{mo}}}{\partial t} &= \frac{\partial}{\partial z} \left(\theta_{\text{mo}} D_{\text{mo}} \frac{\partial c_{\text{mo}}}{\partial z} \right) - \frac{\partial q c_{\text{mo}}}{\partial z} - \phi_{\text{mo}} - \Gamma_s \\ \frac{\partial \theta_{\text{im}} c_{\text{im}}}{\partial t} + \frac{\partial (1-f) \rho s_{\text{im}}}{\partial t} &= -\phi_{\text{im}} + \Gamma_s\end{aligned}\quad (22.31)$$

for the mobile (macropores, subscript mo) and immobile (matrix, subscript im) domains, respectively, where f is the dimensionless fraction of sorption sites in contact with the mobile water [−], ϕ_{mo} and ϕ_{im} are reactions in the mobile and immobile domains [$\text{ML}^{-3}\text{T}^{-1}$], respectively, and Γ_s is the solute transfer rate between the two regions [$\text{ML}^{-3}\text{T}^{-1}$]. Notice that the same Equation 22.31 can be used to describe solute transport using both the mobile–immobile and dual-porosity models shown in Figures 22.2b and Figure 22.2c, respectively.

22.3.3.1.2 Dual-Permeability Model

Analogous to Equations 22.6 and Equation 22.7 for water flow, the dual-permeability formulation for solute transport can be based on advection–dispersion type equations for transport in both the fracture and matrix regions as follows (Gerke and van Genuchten, 1993a):

$$\frac{\partial \theta_f c_f}{\partial t} + \frac{\partial \rho s_f}{\partial t} = \frac{\partial}{\partial z} \left(\theta_f D_f \frac{\partial c_f}{\partial z} \right) - \frac{\partial q_f c_f}{\partial z} - \phi_f - \frac{\Gamma_s}{w} \quad (22.32)$$

$$\frac{\partial \theta_m c_m}{\partial t} + \frac{\partial \rho s_m}{\partial t} = \frac{\partial}{\partial z} \left(\theta_m D_m \frac{\partial c_m}{\partial z} \right) - \frac{\partial q_m c_m}{\partial z} - \phi_m + \frac{\Gamma_s}{1-w} \quad (22.33)$$

where the subscript f and m refer to the macroporous (fracture) and matrix pore systems, respectively; ϕ_f and ϕ_m represent sources or sinks in the macroporous and matrix domains [$\text{ML}^{-3}\text{T}^{-1}$], respectively; and w is the ratio of the volumes of the macropore-domain (inter-aggregate) and the total soil systems [−]. Equation 22.32 and Equation 22.33 assume complete advective–dispersive type transport descriptions for both the fractures and the matrix. Several authors simplified transport in the macropore domain, for example, by considering only piston displacement of solutes (Ahuja and Hebson, 1992; Jarvis, 1994).

22.3.3.1.3 Mass Transfer

The transfer rate, Γ_s , in Equation 22.31 for solutes between the mobile and immobile domains in the dual-porosity models can be given as the sum of diffusive and advective fluxes, and can be written as

$$\Gamma_s = \alpha_s (c_{\text{mo}} - c_{\text{im}}) + \Gamma_w c^* \quad (22.34)$$

where c^* is equal to c_{mo} for $\Gamma_w > 0$ and c_{im} for $\Gamma_w < 0$, and α_s is the first-order solute mass transfer coefficient [T^{-1}]. Notice that the advection term of Equation 22.34 is equal to zero for the mobile–immobile model (Figure 22.2b) as the immobile water content in this model is assumed to be constant. However, Γ_w may have a nonzero value in the dual-porosity model depicted in Figure 22.2c.

The transfer rate, Γ_s , in Equation 22.32 and Equation 22.33 for solutes between the fracture and matrix regions is also usually given as the sum of diffusive and advective fluxes as follows (e.g., Gerke and van Genuchten, 1996):

$$\Gamma_s = \alpha_s (1 - w_m) (c_f - c_m) + \Gamma_w c^* \quad (22.35)$$

in which the mass transfer coefficient, α_s [T^{-1}], is of the form:

$$\alpha_s = \frac{\beta}{d^2} D_a \quad (22.36)$$

where D_a is an effective diffusion coefficient [$L^2 T^{-1}$] representing the diffusion properties of the fracture–matrix interface.

Still more sophisticated models for physical nonequilibrium transport may be formulated. For example, Pot et al. (2005) and Köhne et al. (2006) considered a dual-permeability model that divides the matrix domain further into mobile and immobile subregions and used this model successfully to simulate bromide transport in laboratory soil columns at different flow rates or for transient flow conditions, respectively.

22.3.3.2 Chemical Nonequilibrium

22.3.3.2.1 Kinetic Sorption Models

An alternative to expressing sorption as an instantaneous process using algebraic equations (e.g., Equation 22.24, Equation 22.25 or Equation 22.26) is to describe the kinetics of the reaction using ordinary differential equations. The most popular and simplest formulation of a chemically controlled kinetic reaction arises when first-order linear kinetics is assumed:

$$\frac{\partial s}{\partial t} = \alpha_k (K_d c - s) \quad (22.37)$$

where α_k is a first-order kinetic rate coefficient [T^{-1}]. Several other nonequilibrium adsorption expressions were also used in the past (see Table 2 in van Genuchten and Šimůnek, 1996). Models based on this and other kinetic expressions are often referred to as one-site sorption models.

As transport models assuming chemically controlled nonequilibrium (one-site sorption) generally did not result in significant improvements in their predictive capabilities when used to analyze laboratory column experiments, the one-site first-order kinetic model was further expanded into a two-site sorption concept that divides the available sorption sites into two fractions (Selim et al., 1976; van Genuchten and Wagener, 1989). In this approach, sorption on one fraction (type-1 sites) is assumed to be instantaneous while sorption on the remaining (type-2) sites is considered to be time-dependent. Assuming a linear sorption process, the two-site transport model is given by (van Genuchten and Wagener, 1989)

$$\begin{aligned} \frac{\partial(f\rho_b K_d + \theta)c}{\partial t} + \rho_b \frac{\partial s_k}{\partial t} &= \frac{\partial}{\partial z} \left(\theta D_e \frac{\partial c}{\partial z} \right) - \frac{\partial(qc)}{\partial z} - \phi_e \\ \frac{\partial s_k}{\partial t} &= \alpha_k [(1-f)K_d c - s_k] - \phi_k \end{aligned} \quad (22.38)$$

where f is the fraction of exchange sites assumed to be at equilibrium [−], ϕ_e [$ML^{-3}T^{-1}$] and ϕ_k [$MM^{-1}T^{-1}$] are reactions in the equilibrium and nonequilibrium phases, respectively, and the subscript k refers to kinetic (type-2) sorption sites. Note that if $f = 0$, the two-site sorption model reduces to the one-site fully kinetic sorption model (i.e., when only type-2 kinetic sites are present). On the other hand, if $f = 1$, the two-site sorption model reduces to the equilibrium sorption model for which only type-1 equilibrium sites are present.

22.3.3.2.2 Attachment/Detachment Models

Additionally, transport equations may include provisions for kinetic attachment/detachment of solutes to the solid phase, thus permitting simulations of the transport of colloids, viruses, and bacteria. The transport of these constituents is generally more complex than that of other solutes in that they are affected by such additional processes as filtration, straining, sedimentation, adsorption and desorption, growth, and inactivation. Virus, colloid, and bacteria transport and fate models commonly employ a

modified form of the ADE, in which the kinetic sorption equations are replaced with equations describing kinetics of colloid attachment and detachment as follows:

$$\rho \frac{\partial s}{\partial t} = \theta k_a \psi c - k_d \rho s \quad (22.39)$$

where c is the (colloid, virus, bacteria) concentration in the aqueous phase [$N_c L^{-3}$], s is the solid phase (colloid, virus, bacteria) concentration [$N_c M^{-1}$], in which N_c is a number of (colloid) particles, k_a is the first-order deposition (attachment) coefficient [T^{-1}], k_d is the first-order entrainment (detachment) coefficient [T^{-1}], and ψ is a dimensionless colloid retention function [—]. The attachment and detachment coefficients in Equation 22.39 have been found to strongly depend upon water content, with attachment significantly increasing as the water content decreases.

To simulate reductions in the attachment coefficient due to filling of favorable sorption sites, ψ is sometimes assumed to decrease with increasing colloid mass retention. A Langmuirian dynamics (Adamczyk et al., 1994) equation has been proposed for ψ to describe this blocking phenomenon:

$$\psi = \frac{s_{\max} - s}{s_{\max}} = 1 - \frac{s}{s_{\max}} \quad (22.40)$$

in which s_{\max} is the maximum solid phase concentration [$N_c M^{-1}$].

A similar equation as Equation 22.39 was used by Bradford et al. (2003, 2004) to simulate the process of pore straining. Bradford et al. (2003, 2004) hypothesized that the influence of straining and attachment processes on colloid retention should be separated into two distinct components. They suggested the following depth-dependent blocking coefficient for the straining process:

$$\psi = \left(\frac{d_c + z - z_0}{d_c} \right)^{-\beta} \quad (22.41)$$

where d_c is the diameter of the sand grains [L], z_0 is the coordinate of the location where the straining process starts [L] (the surface of the soil profile, or interface between soil layers), and β is an empirical factor (Bradford et al., 2003) [—].

The attachment coefficient is often calculated using filtration theory (Logan et al., 1995), a quasi-empirical formulation in terms of the median grain diameter of the porous medium (often termed the collector), the pore-water velocity, and collector and collision (or sticking) efficiencies accounting for colloid removal due to diffusion, interception, and gravitational sedimentation (Rajagopalan and Tien, 1976; Logan et al., 1995):

$$k_a = \frac{3(1-\theta)}{2d_c} \eta \alpha v \quad (22.42)$$

where d_c is the diameter of the sand grains [L], α is the sticking efficiency (ratio of the rate of particles that stick to a collector to the rate they strike the collector) [—], v is the pore-water velocity [LT^{-1}], and η is the single-collector efficiency [—].

In related studies, Schijven and Hassanzadeh (2000) and Schijven and Šimůnek (2002) used a two-site sorption model based on two equations (22.39) to successfully describe virus transport at both the laboratory and field scale. Their model assumed that the sorption sites on the solid phase can be divided into two fractions with different properties and various attachment and detachment rate coefficients.

22.3.3.3 Colloid-Facilitated Solute Transport

There is an increasing evidence that many contaminants, including radionuclides (Von Gunten et al., 1988; Noell et al., 1998), pesticides (Vinten et al., 1983; Kan and Tomson, 1990, Lindqvist and Enfield, 1992), heavy metals (Grolimund et al., 1996), viruses, pharmaceuticals (Tolls, 2001; Thiele-Bruhn, 2003), hormones (Hanselman et al., 2003), and other contaminants (Magee et al., 1991; Mansfeldt et al., 2004) are transported in the subsurface not only with moving water, but also sorbed to mobile colloids. As many

colloids and microbes are negatively charged and thus electrostatically repelled by negatively charged solid surfaces, which may lead to an anion exclusion process, their transport may be slightly enhanced relative to fluid flow. Size exclusion may similarly enhance the advective transport of colloids by limiting their presence and mobility to the larger pores (e.g., Bradford et al., 2003). The transport of contaminants sorbed to mobile colloids can thus significantly accelerate their transport relative to more standard advection–transport descriptions.

Colloid-facilitated transport is a relatively complicated process that requires knowledge of water flow, colloid transport, dissolved contaminant transport, and colloid-contaminant interaction. Transport and mass balance equations, hence, must be formulated not only for water flow and colloid transport, but also for the total contaminant, for contaminant sorbed kinetically or instantaneously to the solid phase, and for contaminant sorbed to mobile colloids, to colloids attached to the soil solid phase, and to colloids accumulating at the air–water interface. Development of such a model is beyond the scope of this chapter. We refer interested readers to several manuscripts dealing with this topic: Mills et al. (1991), Corapcioglu and Jiang (1993), Corapcioglu and Kim (1995), Jiang and Corapcioglu (1993), Noell et al. (1998), Saiers et al. (1996), Saiers and Hornberger (1996), van de Weerd et al. (1998), and van Genuchten and Šimunek (2004).

22.3.4 Stochastic Models

Much evidence suggests that solutions of classical solute transport models, no matter how refined to include the most relevant chemical and microbiological processes and soil properties, often still fail to accurately describe transport processes in most natural field soils. A major reason for this failure is the fact that the subsurface environment is overwhelmingly heterogeneous. Heterogeneity occurs at a hierarchy of spatial and time scales (Wheatcraft and Cushman, 1991), ranging from microscopic scales involving time-dependent chemical sorption and precipitation/dissolution reactions, to intermediate scales involving the preferential movement of water and chemicals through macropores or fractures, and to much larger scales involving the spatial variability of soils across the landscape. Subsurface heterogeneity can be addressed in terms of process-based descriptions which attempt to consider the effects of heterogeneity at one or several scales. It can also be addressed using stochastic approaches that incorporate certain assumptions about the transport process in the heterogeneous system (e.g., Sposito and Barry, 1987; Dagan, 1989). In this section we briefly review several stochastic transport approaches, notably those using stream tube models and the transfer function approach.

22.3.4.1 Stream Tube Models

The downward movement of chemicals from the soil surface to an underlying aquifer may be described stochastically by viewing the field as a series of independent vertical columns, often referred to as “stream tubes” (Figure 22.5), while solute mixing between the stream tubes is assumed to be negligible. Transport

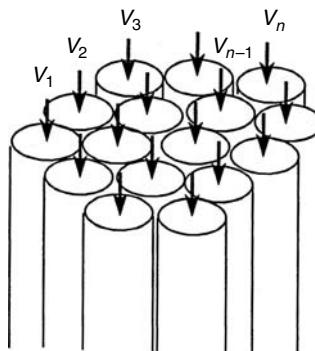


FIGURE 22.5 Schematic illustration of the stream tube model (Toride et al., 1995).

in each tube may be described deterministically with the standard ADE, or modifications thereof to include additional geochemical and microbiological processes. Transport at the field scale is then implemented by considering the column parameters as realizations of a stochastic process, having a random distribution (Toride et al., 1995). Early examples are by Dagan and Bresler (1979) and Bresler and Dagan (1979) who assumed that the saturated hydraulic conductivity had a lognormal distribution.

The stream tube model was implemented into the CXTFIT 2.0 program (Toride et al., 1995) for a variety of transport scenarios in which the pore-water velocity in combination with either the dispersion coefficient, D_e , the distribution coefficient for linear adsorption, K_d , or the first-order rate coefficient for nonequilibrium adsorption, α_k , are stochastic variables (Toride et al., 1995).

22.3.4.2 Transfer Function Models

Jury (1982) developed an alternative formulation for solute transport at the field scale, called the transfer function model. This model was developed based on two main assumptions about the soil system: (a) the solute transport is a linear process, and (b) the solute travel time probabilities do not change over time. These two assumptions lead to the following transfer function equation that relates the solute concentration at the outflow end of the system with the time-dependent solute input into the system:

$$c_{\text{out}}(t) = \int_0^t c_{\text{in}}(t-t')f(t') dt' \quad (22.43)$$

The outflow at time t , $c_{\text{out}}(t)$ [ML^{-3}], consists of the superposition of solute added at all times less than t , $c_{\text{in}}(t-t')$ [ML^{-3}], weighted by its travel-time probability density function (pdf) $f(t)$ [T^{-1}] (Jury and Horton, 2004). One important advantage of the transfer function approach is that it does not require knowledge of the various transport processes within the flow domain. Different model distribution functions can be used for the travel-time probability density function $f(t)$ in Equation 22.43. Most commonly used (Jury and Sposito, 1985) are the Fickian probability density function

$$f(t) = \frac{L}{2\sqrt{\pi Dt^3}} \exp\left[-\frac{(L-vt)^2}{4Dt}\right] \quad (22.44)$$

and the lognormal distribution

$$f(t) = \frac{1}{\sqrt{2\pi}\sigma t} \exp\left[-\frac{(\ln t - \mu)^2}{2\sigma^2}\right] \quad (22.45)$$

where $D[\text{L}^2\text{T}^{-1}]$, $v[\text{LT}^{-1}]$, μ , and σ are model parameters, and L is the distance from the inflow boundary to the outflow boundary [L]. To accommodate conditions when the water flux through the soil system is not constant, Jury (1982) expressed the travel-time pdf as a function of the cumulative net applied water I :

$$I = \int_0^t q(t') dt' \quad (22.46)$$

leading to the following transfer function equation:

$$c_{\text{out}}(I) = \int_0^{I(t)} c_{\text{in}}(I-I')f(I') dI' \quad (22.47)$$

22.3.5 Multicomponent Reactive Solute Transport

The various mathematical descriptions of solute transport presented thus far all considered solutes that would move independently of other solutes in the subsurface. In reality, the transport of reactive contaminants is more often than not affected by many often interactive physico-chemical and biological processes. Simulating these processes requires a more comprehensive approach that couples the physical processes of water flow and advective-dispersive transport with a range of biogeochemical processes. The soil solution is always a mixture of many ions that may be involved in mutually dependent chemical processes, such as complexation reactions, cation exchange, precipitation-dissolution, sorption-desorption, volatilization, redox reactions, and degradation, among other reactions (Šimunek and Valocchi, 2002). The transport and transformation of many chemical contaminants is further mediated by subsurface aerobic or anaerobic bacteria. Bacteria catalyze redox reactions in which organic compounds (e.g., hydrocarbons) act as the electron donor and inorganic substances (oxygen, nitrate, sulfate, or metal oxides) as the electron acceptor. By catalyzing such reactions, bacteria gain energy and organic carbon to produce new biomass. These and related processes can be simulated using integrated reactive transport codes that couple the physical processes of water flow and advective-dispersive solute transport with a range of biogeochemical processes. This section reviews various modeling approaches for such multicomponent transport systems.

22.3.5.1 Components and Reversible Chemical Reactions

Multi-species chemical equilibrium systems are generally defined in terms of components. Components may be defined as a set of linearly independent chemical entities such that every species in the system can be uniquely represented as the product of a reaction involving only these components (Westall et al., 1976). As a typical example, the chemical species CaCO_3^0



consists of the two components Ca^{2+} and CO_3^{2-} .

Reversible chemical reaction processes in equilibrium systems are most often represented using mass action laws that relate thermodynamic equilibrium constants to activities (the thermodynamic effective concentration) of the reactants and products (Mangold and Tsang, 1991; Appelo and Postma, 1993; Bethke, 1996). For example, the reaction



where b and c are the number of moles of substances B and C that react to yield d and e moles of products D and E , is represented at equilibrium by the law of mass action

$$K = \frac{a_D^d a_E^e}{a_B^b a_C^c} \quad (22.50)$$

where K is a temperature-dependent thermodynamic equilibrium constant, and a_i is the ion activity, being defined as the product of the activity coefficient (γ_i) and the ion molality (m_i), that is, $a_i = \gamma_i m_i$. Single-ion activity coefficients may be calculated using either the Davies equation, an extended version of the Debye-Hückel equation (Truesdell and Jones, 1974), or by means of Pitzer expressions (Pitzer, 1979). Equation 22.50 can be used to describe all of the major chemical processes, such as aqueous complexation, sorption, precipitation-dissolution, and acid-base and redox reactions, provided that the local chemical equilibrium assumption is valid (Šimunek and Valocchi, 2002).

22.3.5.2 Complexation

Equations for aqueous complexation reactions can be obtained using the law of mass action as follows (e.g., Yeh and Tripathi, 1990; Lichtner, 1996):

$$x_i = \frac{K_i^x}{\gamma_i^x} \prod_{k=1}^{N_a} (\gamma_k^a c_k)^{a_{ik}^x} \quad i = 1, 2, \dots, M_x \quad (22.51)$$

where x_i is the concentration of the i th complexed species, K_i^x is the thermodynamic equilibrium constant of the i th complexed species, γ_i^x is the activity coefficient of the i th complexed species, N_a is the number of aqueous components, c_k is the concentration of the k th aqueous component, γ_k^a is the activity coefficient of the k th aqueous component species, a_{ik}^x is the stoichiometric coefficient of the k th aqueous component in the i th complexed species, M_x is the number of complexed species, and subscripts and superscripts x and a refer to complexed species and aqueous components, respectively.

22.3.5.3 Precipitation and Dissolution

Equations describing precipitation–dissolution reactions are also obtained using the law of mass action, but contrary to the other processes are represented by inequalities rather than equalities, as follows (Šimůnek and Valocchi, 2002):

$$K_i^p \geq Q_i^p = \prod_{k=1}^{N_a} (\gamma_k^a c_k)^{a_{ik}^p} \quad i = 1, 2, \dots, M_p \quad (22.52)$$

where M_p is the number of precipitated species, K_i^p is the thermodynamic equilibrium constant of the i th precipitated species, that is, the solubility product equilibrium constant, Q_i^p is the ion activity product of the i th precipitated species, and a_{ik}^p is the stoichiometric coefficient of the k th aqueous component in the i th precipitated species. The inequality in Equation 22.52 indicates that a particular precipitate is formed only when the solution is supersaturated with respect to its aqueous components. If the solution is undersaturated then the precipitated species (if it exists) will dissolve to reach equilibrium conditions.

22.3.5.4 Cation Exchange

Partitioning between the solid exchange phase and the solution phase can be described with the general exchange equation (White and Zelazny, 1986):

$$z_j \cdot A^{z_i+} X_{z_i} + z_i \cdot B^{z_j+} \Leftrightarrow z_i \cdot B^{z_j+} X_{z_j} + z_j \cdot A^{z_i+} \quad (22.53)$$

where A and B are chemical formulas for particular cation (e.g., Ca^{2+} or Na^+), X refers to an “exchanger” site on the soil, and z_i is the valence of species. The mass action equation resulting from this exchange reaction is

$$K_{ij} = \left[\frac{\bar{c}_j^{z_j+}}{\gamma_j^a c_j^{z_j+}} \right]^{z_i} \left[\frac{\gamma_i^a c_i^{z_i+}}{\bar{c}_i^{z_i+}} \right]^{z_j} \quad (22.54)$$

where K_{ij} is the selectivity coefficient, and \bar{c}_i is the exchanger-phase concentration of the i th component (expressed in moles per mass of solid).

22.3.5.5 Coupled System of Equations

Once the various chemical reactions are defined, the final system of governing equations usually consists of several partial differential equations for solute transport (i.e., ADEs for each component) plus a set of nonlinear algebraic and ordinary differential equations describing the equilibrium and kinetic reactions, respectively. Each chemical and biological reaction must be represented by the corresponding algebraic or ordinary differential equations depending upon the rate of the reaction. Since the reaction of one

species depends upon the concentration of many other species, the final sets of equations typically are tightly coupled. For complex geochemical systems, consisting of many components and multidimensional transport, numerical solution of these coupled equations is challenging (Šimůnek and Valocchi, 2002). As an alternative, more general models have recently been developed that also more loosely couple transport and chemistry using a variety of sequential iterative or non-iterative operator-splitting approaches (e.g., Bell and Binning, 2004; Jacques and Šimůnek, 2005). Models based on these various approaches are further discussed in Section 22.5b.

22.3.6 Multiphase Flow and Transport

While the transport of solutes in variably saturated media generally involves two phases (i.e., the liquid and gaseous phases, with advection in the gaseous phase often being neglected), many contamination problems also increasingly involve nonaqueous phase liquids (NAPLs) that are often only slightly miscible with water. Nonaqueous phase liquids may consist of single organic compounds such as many industrial solvents, or of a mixture of organic compounds such as gasoline and diesel fuel. Some of these compounds can be denser than water (commonly referred to as DNAPLs) or lighter than water (LNAPLs). Their fate and dynamics in the subsurface is affected by a multitude of compound-specific flow and multicomponent transport processes, including interphase mass transfer and exchange (also with the solid phase).

Multiphase fluid flow models generally require flow equations for each fluid phase (water, air, NAPL). Two-phase air–water systems hence could be modeled also using separate equations for air and water. This shows that the standard Richards Equation 22.3 is a simplification of a more complete multiphase (air–water) approach in that the air phase is assumed to have a negligible effect on variably saturated flow, and that the air pressure varies only little in space and time. This assumption appears adequate for most variably saturated flow problems. Similar assumptions, however, are generally not possible when NAPLs are present. Hence mathematical descriptions of multiphase flow and transport in general require flow equations for each of the three fluid phases, mass transport equations for all organic components (including those associated with the solid phase), and appropriate equations to account for interphase mass transfer processes. We refer readers to reviews by Abriola et al. (1999) and Rathfelder et al. (2000) for discussions of the complexities involved in modeling such systems subject to multiphase flow, multicomponent transport, and interphase mass transfer. An excellent overview of a variety of experimental approaches for measuring the physical and hydraulic properties of multi-fluid systems is given by Lenhard et al. (2002).

22.3.7 Initial and Boundary Conditions

22.3.7.1 Initial Conditions

The governing equations for solute transport can be solved analytically or numerically provided that the initial and boundary conditions are specified. Initial conditions need to be specified for each equilibrium phase concentration, that is,

$$c(x, y, z, t) = c_i(x, y, z, 0) \quad (22.55)$$

where c_i is the initial concentration [ML^{-3}], as well as for all nonequilibrium phases such as concentrations in the immobile region, sorbed concentrations associated with kinetic sites, and initially attached or strained colloid concentrations.

22.3.7.2 Boundary Conditions

Complex interactions between the transport domain and its environment often must be considered for the water flow part of the problem being considered since these interactions determine the magnitude of water fluxes across the domain boundaries. By comparison, the solute transport part of most analytical and numerical models usually considers only three types of boundary conditions. When the concentration

at the boundary is known, one can use a first-type (or Dirichlet type) boundary condition:

$$c(x, y, z, t) = c_0(x, y, z, t) \quad \text{for } (x, y, z) \in \Gamma_D \quad (22.56)$$

where c_0 is a prescribed concentration [ML^{-3}] at or along the Γ_D Dirichlet boundary segments. This boundary condition is often referred to as a concentration boundary condition. A third-type (Cauchy type) boundary condition may be used to prescribe the concentration flux at the boundary as follows:

$$-\theta D_{ij} \frac{\partial c}{\partial x_j} n_i + q_i n_i c = q_i n_i c_0 \quad \text{for } (x, y, z) \in \Gamma_C \quad (22.57)$$

in which $q_i n_i$ represents the outward fluid flux [LT^{-1}], n_i is the outward unit normal vector and c_0 is the concentration of the incoming fluid [ML^{-3}]. In some cases, for example, when a boundary is impermeable ($q_0 = 0$) or when water flow is directed out of the region, the Cauchy boundary condition reduces to a second-type (Neumann type) boundary condition of the form:

$$\theta D_{ij} \frac{\partial c}{\partial x_j} n_i = 0 \quad \text{for } (x, y, z) \in \Gamma_N \quad (22.58)$$

Most applications require a Cauchy boundary condition rather than Dirichlet (or concentration) boundary condition. Since Cauchy boundary conditions define the solute flux across a boundary, the solute flux entering the transport domain will be known exactly (as specified). This specified solute flux is then in the transport domain divided into advective and dispersive components. On the other hand, Dirichlet boundary condition controls only the concentration on the boundary, and not the solute flux which, because of its advective and dispersive contributions, will be larger than for a Cauchy boundary condition. The incorrect use of Dirichlet rather than Cauchy boundary conditions may lead to significant mass balance errors at an earlier time, especially for relatively short transport domains (van Genuchten and Parker, 1984).

A different type of boundary condition is sometimes used for volatile solutes when they are present in both the liquid and gas phases. This situation requires a third-type boundary condition, but modified to include an additional volatilization term accounting for gaseous diffusion through a stagnant boundary layer of thickness d [L] above the soil surface. The additional solute flux is often assumed to be proportional to the difference in gas concentrations above and below this boundary layer (e.g., Jury et al., 1983). This modified boundary condition has the form:

$$-\theta D_{ij} \frac{\partial c}{\partial x_j} n_i + q_i n_i c = q_i n_i c_0 + \frac{D_g}{d} (k_g c - g_{\text{atm}}) \quad \text{for } (x, z) \in \Gamma_C \quad (22.59)$$

where D_g is the molecular diffusion coefficient in the gas phase [$\text{L}^2 \text{T}^{-1}$] and g_{atm} is the gas concentration above the stagnant boundary layer [ML^{-3}]. We note that Jury et al. (1983) assumed g_{atm} in Equation 22.59 to be zero.

Still other types of boundary conditions can be used. One example is the use of the Bateman equations (Bateman, 1910) to account for a finite rate of release of multiple solutes that are subject to a first-order sequential decay (e.g., radionuclides, nitrogen species). These solutes are typically released from a waste site into the environment as a consequence of decay reactions in the waste site (e.g., van Genuchten, 1985).

22.4 Analytical Models

A large number of computer models using both analytical and numerical solutions of the flow and solute transport equations are now available for a wide range of applications in research and management of natural subsurface systems (Šimůnek, 2005). Modeling approaches range from relatively simple analytical

and semi-analytical models, to more complex numerical codes that permit consideration of a large number of simultaneous nonlinear processes. While for certain conditions (e.g., for linear sorption, a concentration independent sink term ϕ , and a steady flow field) the solute transport equations (e.g., Equation 22.20, Equation 22.21, Equation 22.23, Equation 22.31, and Equation 22.38) become linear, the governing flow equations (e.g., Equation 22.3) are generally highly nonlinear because of the nonlinear dependency of the soil hydraulic properties on the pressure head or water content. Consequently, many analytical solutions have been derived in the past for solute transport equations, and are now widely used for analyzing solute transport during steady-state flow. Although a large number of analytical solutions also exist for the unsaturated flow equation, they generally can be applied only to relatively simple flow problems. The majority of applications for water flow in the vadose zone require a numerical solution of the Richards equation.

Analytical methods are representative of the classical mathematical approach for solving differential equations to produce an exact solution for a particular problem. Analytical models usually lead to an explicit equation for the concentration (or the pressure head, water content, or temperature) at a particular time and location. Hence, one can evaluate the concentration directly without the time stepping that is typical of numerical methods. While exceptions exist (e.g., Liu et al., 2000), analytical solutions usually can be derived only for simplified transport systems involving linearized governing equations, homogeneous soils, simplified geometries of the transport domain, and constant or highly simplified initial and boundary conditions. Unfortunately, analytical solutions for more complex situations, such as for transient water flow or nonequilibrium solute transport with nonlinear reactions, are generally not available and cannot be derived, in which case numerical models must be employed (Šimůnek, 2005).

22.4.1 Analytical Approaches

Analytical solutions are usually obtained by applying various transformations (e.g., Laplace, Fourier or other transforms) to the governing equations, invoking a separation of variables, and using the Green's function approach (e.g., Leij et al., 2000).

22.4.1.1 Laplace Transformation

The Laplace transform, \mathcal{L} , of the solute concentration with respect to time is defined as follows:

$$\bar{c}(x, s) = \mathcal{L}[c(x, t)] = \int_0^{\infty} c(x, t) \exp(-st) dt \quad (22.60)$$

where s is the transform variable [T^{-1}]. Laplace transforms can greatly simplify the governing solute transport equations by eliminating one independent variable, usually time. They also transform the original transport equation from a partial to an ordinary differential equation. Using Laplace transforms, one thus obtains a governing ADE in the Laplace domain that is much simpler to solve analytically than the original equation. The analytical solution in the Laplace space is subsequently inverted to the real space using either a table of Laplace transforms, by applying inversion theorems, or by using a numerical inversion program.

22.4.1.2 Fourier Transformation

Two- and three-dimensional problems often require not only Laplace transforms with respect to time (transformation from t to s) and one spatial coordinate (usually x being transformed to r), but often also a double Fourier transform with respect to two other spatial coordinates (y and z), which is given by (Leij and

Toride, 1997):

$$F_{YZ}[\tilde{c}(r, y, z, s)] = \tilde{c}(r, \gamma, \kappa, s) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} c(r, y, z, s) \exp(i\gamma y + i\kappa z) dy dz \quad (22.61)$$

where $i^2 = -1$ and γ and κ are transformation variables pertaining to the y and z coordinates. Similarly as for Laplace transforms, Fourier transforms lead to an equation that can be solved analytically more easily. One also needs to transform this analytical solution from the Laplace and Fourier domains back to the original time and space domains. Several analytical solutions, including the derivation process, for different initial and boundary conditions, and different transport domain geometries are given by Leij and Toride (1997) in the N3DADE manual.

22.4.1.3 Method of Moments

Statistical moments are often used to characterize the statistical distribution of the solute concentration versus time and space. For example, moments can be used to evaluate the average residence time or the width of the residence time distribution. These attributes can be quantified with the moments of the distribution (Skaggs and Leij, 2002). The p th time moment, m_p , is defined as:

$$m_p = \int_0^{\infty} t^p c(t) dt \quad (p = 0, 1, 2, \dots) \quad (22.62)$$

where $c(t)$ is the breakthrough curve measured at a certain location. The zeroth moment is related to the mass of solute contained in the breakthrough curve, while the first moment is related to the mean residence time. Normalized and central moments are defined as:

$$\begin{aligned} M_p &= \frac{m_p}{m_0} \\ M'_p &= \frac{1}{m_0} \int_0^{\infty} (t - M_1)^p c(t) dt \end{aligned} \quad (22.63)$$

The second central moment relates to the degree of solute spreading. Similar expressions can be written also for the spatial (depth) moments characterizing the spatial solute distribution. Moments for particular problems or processes can then be obtained by substituting appropriate analytical solutions in the above equations.

22.4.2 Existing Models

Numerous analytical solutions of the linear advection-dispersion solute transport equations (e.g., Equation 22.21, Equation 22.23, Equation 22.31, and Equation 22.38), or their two- and three-dimensional equivalents, have been developed during the last 40 years and are now widely used for predictive purposes and for analyzing laboratory and field observed concentration distributions. The majority of these solutions pertain to solute transport equations assuming constant water content, θ , and flux, q , values (i.e., for steady-state water flow conditions in a homogeneous medium).

22.4.2.1 One-Dimensional Models

Some of the more popular one-dimensional analytical transport models have been CFITM (van Genuchten, 1980b), CFITIM (van Genuchten, 1981), CXTFIT (Parker and van Genuchten, 1984), and CXTFIT2 (Toride et al., 1995). While CFITM considers only one-dimensional equilibrium solute transport in both finite and infinite solute columns, CFITIM additionally considers physical and chemical

nonequilibrium transport (i.e., the two-region mobile-immobile model for physical nonequilibrium and the two-site sorption model for chemical nonequilibrium). CXTFIT further expanded the capabilities of CFITIM by considering more general initial and boundary conditions, as well as degradation processes. CXTFIT2 (Toride et al., 1995), an updated version of CXTFIT, solves both direct and inverse problems for three different one-dimensional transport models (1) the conventional advection-dispersion equation, ADE; (2) the chemical and physical nonequilibrium ADEs; and (3) a stochastic stream tube model based upon the local-scale equilibrium or nonequilibrium ADE. These three types of models all consider linear adsorption, and include zero- and first-order decay/source terms. All of these models may be used also to solve the inverse problem by fitting the analytical ADE solutions to experimental results.

22.4.2.2 Multi-Dimensional Models

Some of the more popular multi-dimensional analytical transport models have been AT123D (Yeh, 1981), 3DADE (Leij and Bradford, 1994), N3DADE (Leij and Toride, 1997), and MYGRT (Ungs et al., 1998). These programs provide analytical solutions to transport problems in two- and three-dimensional domains. 3DADE also included parameter estimation capabilities.

A large number of analytical models for one-, two-, and three-dimensional solute transport problems were recently incorporated into the public domain software package STANMOD (STudio of ANalytical MODels) (Šimůnek et al., 1999a) (<http://www.hydrus2d.com>). Figure 22.6 shows a STANMOD dialog window where users can make a selection of the analytical program to be used. This Windows-based computer software package includes not only programs for equilibrium advective-dispersive transport (i.e., the CFITIM of van Genuchten 1980b) for one-dimensional transport and 3DADE (Leij and Bradford, 1994) for three-dimensional problems (Figure 22.7), but also programs for more complex problems. For example, STANMOD also incorporates the CFITIM (van Genuchten, 1981) and N3DADE (Leij and Toride, 1997) programs for nonequilibrium transport (i.e., the two-region mobile-immobile model for physical nonequilibrium and the two-site sorption model for chemical nonequilibrium) in one and multiple dimensions, respectively.

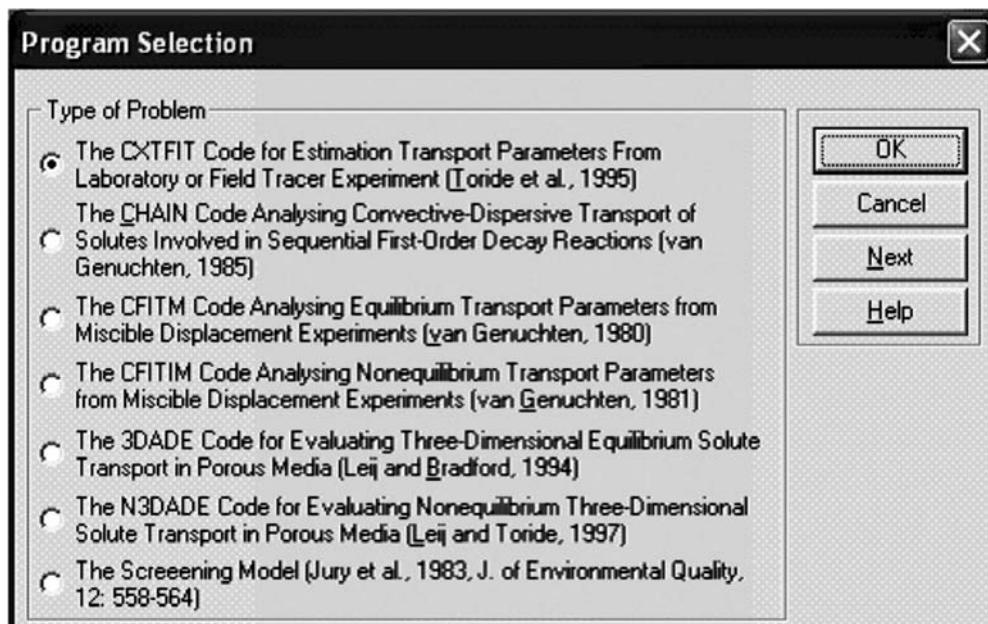


FIGURE 22.6 Selection of an analytical model in the STANMOD software package (Šimůnek et al., 1999a).

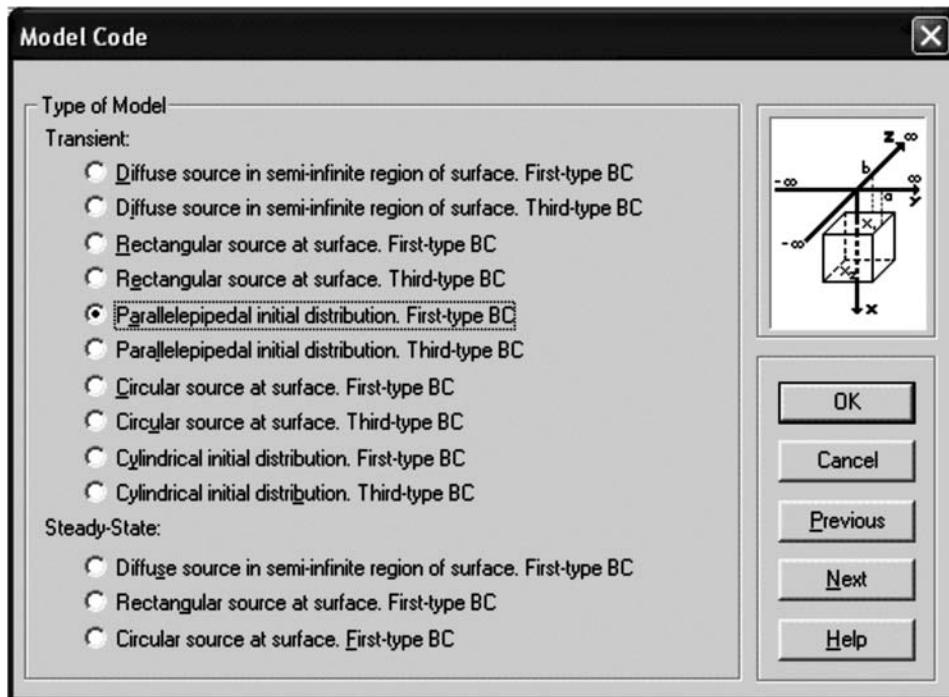


FIGURE 22.7 Selection of the geometry of the transport domain for the 3DADE code (Leij and Bradford, 1994) within the STANMOD software package (Šimůnek et al., 1999a).

Figure 22.7 shows a dialog window of STANMOD in which users can select the geometry of the transport domain for use in the 3DADE program. STANMOD also includes CXTFIT2 (Toride et al., 1995) as discussed earlier, and the CHAIN code of van Genuchten (1985) for analyzing the advective-dispersive transport of up to four solutes involved in sequential first-order decay reactions. Examples of the latter are the migration of radionuclides in which the chain members form first-order decay reactions, and the simultaneous movement of various interacting nitrogen or organic chemicals. The latest version of STANMOD also includes the screening model of Jury et al. (1983) for describing transport and volatilization of soil-applied volatile organic chemicals.

An application of CFITIM within of the STANMOD program is demonstrated in Figure 22.8, which shows measured and fitted breakthrough curves of a nonreactive (${}^3\text{H}_2\text{O}$) solute for transport through a saturated Glendale clay loam soil. In this example, a tritiated water pulse of 3.102 pore volumes was applied to a 30 cm long column, with the breakthrough curve being determined from the effluent concentrations. An analytical solution for two-region (mobile-immobile) physical nonequilibrium transport (van Genuchten, 1981) was used for the analysis. With the pore-water velocity known ($v = 37.54 \text{ cm d}^{-1}$) and assuming that the retardation factor R for ${}^3\text{H}_2\text{O}$ was equal to one (no sorption or anion exclusion), only three parameters were optimized against the breakthrough curve, that is, the dispersion coefficient D ($= 15.6 \text{ cm}^2 \text{ d}^{-1}$), a dimensionless variable β ($= 0.823$) for partitioning the liquid phase in mobile and immobile regions, and a dimensionless mass transfer coefficient ω ($= 0.870$).

An example application of the 3DADE program (Leij and Bradford, 1994) in the STANMOD package is demonstrated in Figure 22.9, which shows calculated concentration contours for a transport problem in which solute is applied at a concentration $C_0 = 1$, assuming a third-type boundary condition, to a rectangular area of the soil surface (or within groundwater) of $15 \times 15 \text{ cm}^2$. Other transport parameters were as follows: $v = 10 \text{ cm d}^{-1}$, $D_x = 100$ and $D_y = D_z = 10 \text{ cm}^2 \text{ d}^{-1}$, and $R = 1$. Figure 22.9 shows concentration contours in the xy -plane at $t = 1$ days, and $z = 0$.

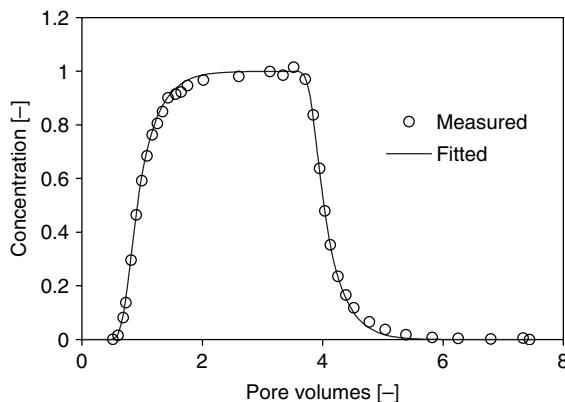


FIGURE 22.8 Breakthrough curve for nonreactive tritium transport as analyzed with the two-region physical nonequilibrium program CXTFIT2 (Toride et al., 1995) within STANMOD (Šimunek et al., 1999a).

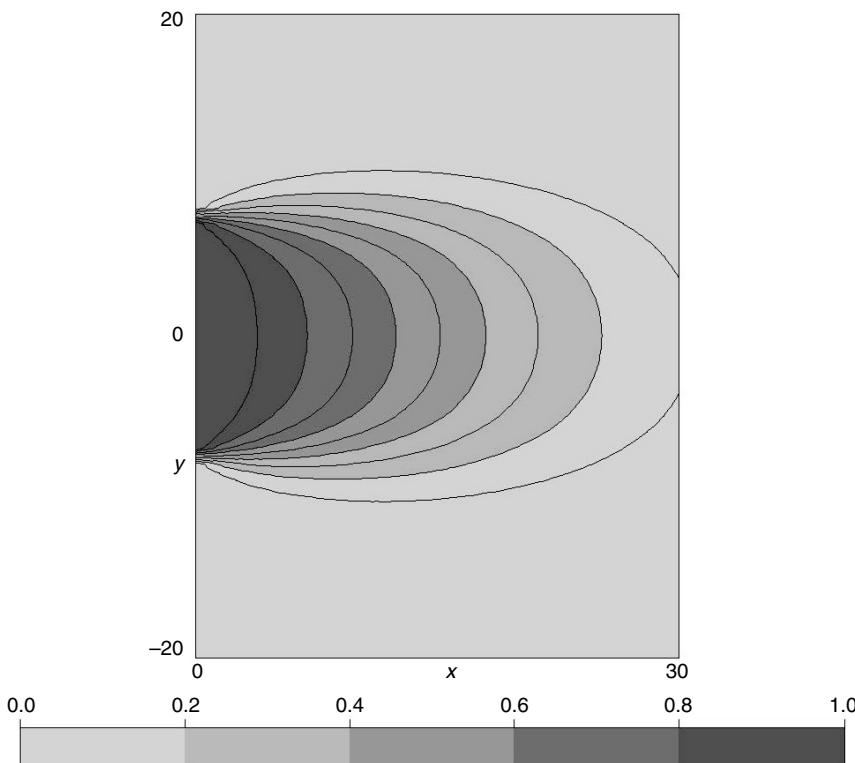


FIGURE 22.9 Solute distribution in the xy -plane for continuous application at $x = 0$ and $-7.5 < y < 7.5$ cm after 1 days, with $v = 10 \text{ cm d}^{-1}$, $D_x = 100$ and $D_y = D_z = 10 \text{ cm}^2 \text{ d}^{-1}$ as calculated with 3DADE (Leij and Bradford, 1994) within STANMOD (Šimunek et al., 1999a).

22.5 Numerical Models

22.5.1 Numerical Approaches

While analytical and semi-analytical solutions are still popularly used for many relatively simple applications, the ever-increasing power of personal computers, and the development of more accurate and

stable numerical solution techniques, have led to the much wider use of numerical models in recent years. Numerical methods are in general superior to analytical methods in terms of being able to solve much more practical problems (Šimůnek, 2005). They allow users to design complicated geometries that reflect complex natural geologic and hydrologic conditions, control parameters in space and time, prescribe more realistic initial and boundary conditions, and permit the implementation of nonlinear constitutive relationships. Numerical methods usually subdivide the time and spatial coordinates into smaller pieces, such as finite differences, finite elements, and finite volumes, and reformulate the continuous form of governing partial differential equations in terms of a system of algebraic equations. To obtain solutions at certain times, numerical methods generally require intermediate simulations (time-stepping) between the initial condition and the points in time for which the solution is needed.

Reviews of the history of development of various numerical techniques used in vadose zone flow and transport models are given by van Genuchten and Šimůnek (1996) and Šimůnek (2005). Here we will summarize only the main principles behind some of the more popular techniques.

22.5.1.1 Finite Differences

The method of finite differences is generally very intuitive and relatively easy to implement. Time and space are both divided into small increments Δt and Δz (or Δx and Δz) (Figure 22.10). Temporal and spatial derivatives in the governing equation are then replaced with finite differences (formally using Taylor series expansions). For example, the standard advection-dispersion equation for steady-state water flow:

$$\frac{\partial c}{\partial t} = -\frac{\partial J}{\partial z} = D \frac{\partial^2 c}{\partial z^2} - v \frac{\partial c}{\partial z} \quad (22.64)$$

can be approximated as follows using an explicit (forward-in-time) finite difference scheme:

$$\frac{c_i^{j+1} - c_i^j}{\Delta t} = -\frac{J_{i+1/2}^j - J_{i-1/2}^j}{\Delta z} = D \frac{c_{i+1}^j - 2c_i^j + c_{i-1}^j}{(\Delta z)^2} - v \frac{c_{i+1}^j - c_{i-1}^j}{2\Delta z} \quad (22.65)$$

where subscripts refer to spatial discretization and superscript to temporal discretization (j and $j+1$ are for the previous and actual time levels, respectively; see Figure 22.10), Δt is the time step, and Δz is the spatial step (assumed to be constant). Notice that this equation contains only one unknown variable (i.e., the concentration c_i^{j+1} at the new time level), which hence can be evaluated directly (explicitly) by solving the equation.

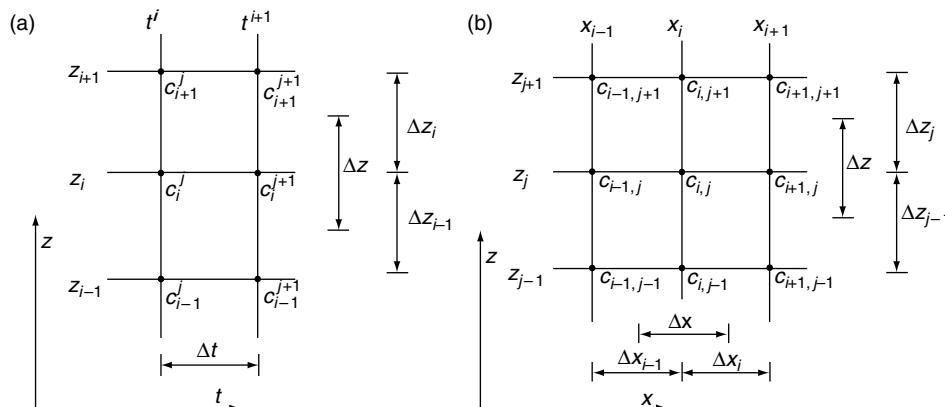


FIGURE 22.10 Examples of the spatial and temporal finite difference discretization of a one-dimensional problem (a), and the finite difference discretization of a two-dimensional domain (b).

By comparison, a fully implicit (backward-in-time) finite difference scheme can be written as follows:

$$\frac{c_i^{j+1} - c_i^j}{\Delta t} = -\frac{J_{i+1/2}^{j+1} - J_{i-1/2}^{j+1}}{\Delta z} = D \frac{c_{i+1}^{j+1} - 2c_i^{j+1} + c_{i-1}^{j+1}}{(\Delta z)^2} - \nu \frac{c_{i+1}^{j+1} - c_{i-1}^{j+1}}{2\Delta z} \quad (22.66)$$

and an implicit (weighted) finite difference scheme as:

$$\frac{c_i^{j+1} - c_i^j}{\Delta t} = D \frac{\varepsilon(c_{i+1}^{j+1} - 2c_i^{j+1} + c_{i-1}^{j+1}) + (1-\varepsilon)(c_{i+1}^j - 2c_i^j + c_{i-1}^j)}{(\Delta z)^2} - \nu \frac{\varepsilon(c_{i+1}^{j+1} - c_{i-1}^{j+1}) + (1-\varepsilon)(c_{i+1}^j - c_{i-1}^j)}{2\Delta z} \quad (22.67)$$

where ε is a temporal weighting coefficient. Different finite-difference schemes results depending upon the value of ε , that is, an explicit scheme when $\varepsilon = 0$, a Crank–Nicholson time-centered scheme when $\varepsilon = 0.5$, and a fully implicit scheme when $\varepsilon = 1$. Of these, the latter two cases lead to several unknown concentrations in each equation at the new time level. Consequently, equations for all nodes of the transport domain need to be assembled into an algebraic system of linear equations to produce a matrix equation of the form

$$[P]^{j+1} \{c\}^{j+1} = \{F\} \quad (22.68)$$

in which $[P]$ is a tridiagonal matrix given by:

$$[P] = \begin{vmatrix} d_1 & e_1 & 0 & & & 0 \\ b_2 & d_2 & e_2 & 0 & & 0 \\ 0 & b_3 & d_3 & e_3 & 0 & 0 \\ & \vdots & \vdots & \vdots & & \\ 0 & 0 & b_{N-2} & d_{N-2} & e_{N-2} & 0 \\ 0 & & 0 & b_{N-1} & d_{N-1} & e_{N-1} \\ 0 & & & 0 & b_N & d_N \end{vmatrix} \quad (22.69)$$

Equation 22.68 is subsequently solved using a matrix equation solver to yield the concentrations at the new time level.

The problem becomes more complex for problems involving transient flow. In that case one must first obtain estimates of the water contents and fluxes at the new time level using a numerical solution of the Richards equation. These water contents and fluxes are subsequently used for assembling the matrix equations for solute transport. While finite differences are relatively straightforward to implement for many transport problems, one major disadvantage is that this method cannot be easily applied to relatively complicated (irregular) transport domains in two and three dimensions.

22.5.1.2 Finite Elements

Finite element methods can be implemented in very much the same way as finite differences for one-, two-, and three-dimensional problems. A major advantage of the finite elements is that they are much easier used to discretize complex two- and three-dimensional transport domains (Figure 22.11). As an example, Figure 22.11 shows triangular unstructured finite element grids for a regular rectangular and an irregular domain as generated with the automated MeshGen2D mesh generator of HYDRUS-2D (Šimůnek et al., 1999b). Note that even though the figure on the right (Figure 22.11) has an irregular soil surface, as well as a tile drain within the transport domain, MeshGen2D could easily discretize/accommodate this transport domain using an unstructured triangular finite element mesh.

The finite element method assumes that the dependent variable in the advection–dispersion equation (e.g., the concentration function $c(x, t)$ in (22.64)) can be approximated by a finite series $c'(x, t)$ of

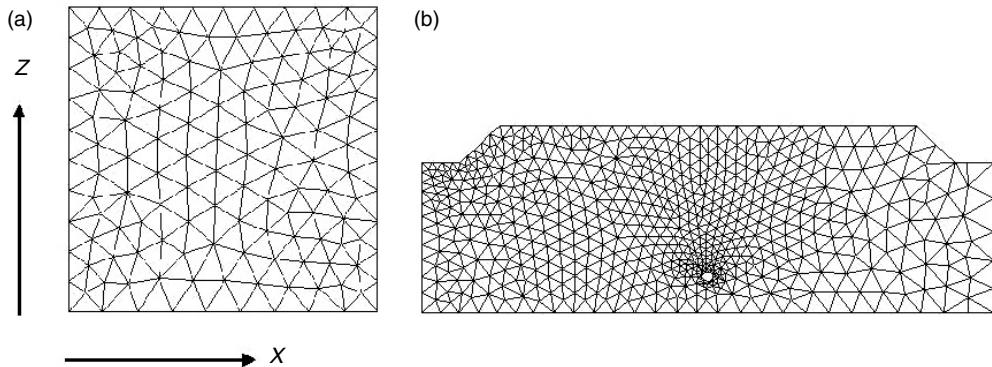


FIGURE 22.11 Examples of the triangular finite element grids for regular (a) and irregular (b) two-dimensional transport domains.

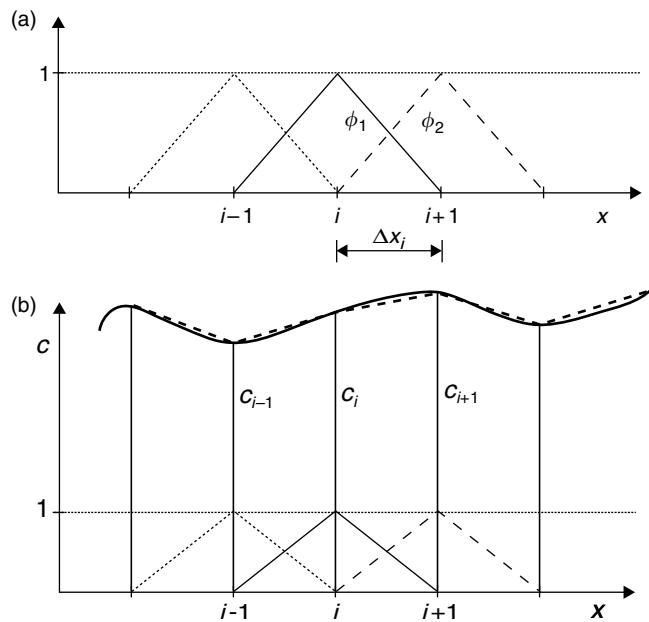


FIGURE 22.12 Example of one-dimensional linear basis functions (a) and their use to approximate solution $c(x, t)$ of the advection-dispersion equation (b).

the form (Figure 22.12b):

$$c'(x, t) = \sum_{m=1}^N \phi_m(x) c_m(t) \quad (22.70)$$

where ϕ_m are the selected linear basis functions that fulfill the condition $\phi_m(x_n) = \delta_{nm}$, δ_{nm} is Kronecker delta ($\delta_{nm} = 1$ for $m = n$, and $\delta_{nm} = 0$ for $m \neq n$), c_m are the unknown time-dependent coefficients which represent solutions of (22.64) at the finite element nodal points, and N is the total number of nodal points. For example, for the one-dimensional finite element $x_i < x < x_{i+1}$, linear basis functions have

the following form (Figure 22.12a):

$$\left. \begin{array}{l} \phi_1 = 1 - \frac{x - x_i}{\Delta x} \\ \phi_2 = \frac{x - x_i}{\Delta x} \end{array} \right\} \quad x_i \leq x \leq x_{i+1} \quad (22.71)$$

where $\Delta x (= x_{i+1} - x_i)$ is the size of a finite element [L], that is, the distance between two neighboring nodal points. The approximate solution $c'(x, t)$ converges to the correct solution $c(x, t)$ as the number of basis functions N increases.

Application of the Galerkin method, which postulates that the differential operator associated with the transport equation is orthogonal to each of the N basis functions (e.g., Pinder and Gray, 1977), leads to the following system of N time-dependent differential equations with N unknown values $c_n(t)$:

$$\int_0^L \left(-\frac{\partial c}{\partial t} + D \frac{\partial^2 c}{\partial x^2} - v \frac{\partial c}{\partial x} \right) \phi_n \, dx = 0 \quad (22.72)$$

where ϕ_n are the selected linear weighting functions that can be the same as the basis functions ϕ_m . Integrating by parts the terms containing the spatial derivatives to remove the second-order derivatives, and substituting (22.70) for $c(x, t)$ leads to the following equation:

$$\sum_e \int_0^{L_e} \left(-\frac{\partial c_m}{\partial t} \phi_m \right) \phi_n \, dx - \sum_e \int_0^{L_e} \left(D c_m \frac{\partial \phi_m}{\partial x} - v c_m \phi_m \right) \frac{\partial \phi_n}{\partial x} \, dx - q_{sL} \phi_n(L) + q_{s0} \phi_n(0) = 0 \quad (22.73)$$

where q_{s0} and q_{sL} are solute fluxes across the left and right boundaries, respectively; e is the element index, L_e is the size of an element e , and L is the size of the transport domain (length of the soil profile). The integrals in Equation 22.73 can be evaluated either numerically or analytically, depending upon the type of invoked basis and weighting functions (ϕ_m and ϕ_n , respectively). The equation can be rewritten in matrix form similarly as Equation 22.68 and then solved using a matrix equation solver. For a detailed derivation of the matrix equation we refer to standard finite element textbooks (e.g., Pinder and Gray, 1977; Zienkiewicz, 1977; Huyakorn and Pinder, 1983) or manuals of some of the models (e.g., the HYDRUS models by Šimůnek et al., 1998a, 1999b).

22.5.1.3 Stability and Oscillations

Numerical solutions of the transport equation often exhibit oscillatory behavior and/or excessive numerical dispersion near relatively sharp concentration fronts (Huyakorn and Pinder, 1983; Šimůnek et al., 1998a). These problems can be especially serious for advection-dominated transport characterized by small dispersivities. One way to partially circumvent numerical oscillations is to use upstream weighing in which case the flux term of Equation 22.64 is not weighted using regular linear basis functions ϕ_n , but instead with nonlinear quadratic functions ϕ_n^u . Undesired oscillations can often be prevented also by selecting an appropriate combination of the space and time discretizations. Two dimensionless numbers may be used to characterize the discretizations in space and time. One of these is the grid Peclet number, Pe^e , which defines the predominant type of solute transport (notably the ratio of the advective and dispersive transport terms) in relation to coarseness of the finite element grid:

$$Pe^e = \frac{v \Delta x}{D} \quad (22.74)$$

where Δx is the characteristic length of a finite element [L]. The Peclet number increases when the advective part of the transport equation dominates the dispersive part (i.e., when a relatively steep concentration front is present). To achieve acceptable numerical results, the spatial discretization must be kept relatively fine to maintain a low Peclet number. Numerical oscillations can be virtually eliminated when the local Peclet numbers do not exceed about 5. However, acceptably small oscillations may be obtained with local Peclet numbers as high as 10 (Huyakorn and Pinder, 1983). Undesired oscillations for higher Peclet numbers can be effectively eliminated by using upstream weighing.

A second dimensionless number, the Courant number, Cr^e , may be used to characterize the relative extent of numerical oscillations in the numerical solution. The Courant number is associated with the time discretization, Δt [T], as follows:

$$Cr^e = \frac{v\Delta t}{\Delta x} \quad (22.75)$$

Given a certain spatial discretization, the time step must be selected such that the Courant number remains less than or equal to 1.

Perrochet and Berod (1993) developed a criterion for minimizing or eliminating numerical oscillations that is based on the product of the Peclet and Courant numbers:

$$Pe \cdot Cr \leq \omega_s \quad (=2) \quad (22.76)$$

where ω_s is the performance index [—]. This criterion indicates that advection-dominated transport problems having large Pe numbers can be safely simulated provided Cr is reduced according to (22.76) (Perrochet and Berod, 1993). When small oscillations in the solution can be tolerated, ω_s can be increased to about 5 or 10.

22.5.2 Existing Models

22.5.2.1 Single-Species Solute Transport Models

A large number of numerical models are now available for evaluating variably saturated water flow and solute transport processes in the subsurface. Some of these models are in the public domain, such as MACRO (Jarvis, 1994), SWAP (van Dam et al., 1997), UNSATH (Fayer, 2000), VS2DI (Healy, 1990), and HYDRUS-1D (Šimůnek et al., 1998a, 2005), while others are in the commercial domain, such as HYDRUS-2D (Šimůnek et al., 1999b) and MODFLOW-SURFACT (HydroGeoLogic, 1996). These models vary widely in terms of their complexity, sophistication, and ease of use. Although some models are still being run under the DOS operating system, with associated difficulties of preparing input files and interpreting tabulated outputs, many others, especially those in the commercial domain, are supported by sophisticated graphics-based interfaces that greatly simplify their use (Šimůnek et al., 1998a; 1999b). Several studies have recently reviewed and compared various numerical models for vadose zone applications (e.g., Wilson et al., 1999; Scanlon et al., 2002; MDH Engineered Solutions Corp., 2003; Vanderborgh et al., 2005). These studies typically evaluated comparable precision, speed, and ease of use of the codes involved.

While many of the early numerical models were developed mostly for agricultural applications (e.g., optimizing soil moisture conditions for crop production), this focus has increasingly shifted to environmental research, with the primary concern now being the subsurface fate and transport of various agricultural and other contaminants, such as radionuclides, hydrocarbons, fumigants, pesticides, nutrients, pathogens, pharmaceuticals, viruses, bacteria, colloids, and/or toxic trace elements. The earlier models solved the governing flow and transport equations for relatively simplified system-independent boundary conditions (i.e., specified pressure heads or fluxes, and free drainage). Models developed recently can cope with much more complex system-dependent boundary conditions evaluating surface flow and energy balances and accounting for the simultaneous movement of water, vapor, and heat. Examples are DAISY (Hansen et al., 1990), TOUGH2 (Pruess, 1991), SHAW (Flerchinger et al., 1996), SWAP (van Dam et al., 1997), HYDRUS-1D (Šimůnek et al., 1998a, 2005), UNSATH (Fayer, 2000), and COUP (Jansson

and Karlberg, 2001). Several models now also account for the extremely nonlinear processes associated with the freezing and thawing cycle (e.g., DAISY, SHAW, and COUP).

Models have recently also become increasingly sophisticated in terms of the type and complexity of solute transport processes that can be simulated. Transport models are no longer being limited to solutes undergoing relatively simple chemical reactions such as linear sorption and first-order decay, but now consider also a variety of nonlinear sorption and exchange processes, physical and chemical nonequilibrium transport, volatilization, gas diffusion, colloid attachment/ detachment, decay chain reactions, and many other processes (e.g., the HYDRUS-1D and -2D codes of Šimůnek et al., 1999b, 2005, or MODFLOW-SURFACT of HydroGeoLogic, Inc., 1996). For example, the general formulation of the transport equations in the HYDRUS codes permit simulations of not only non-adsorbing or linearly sorbing chemicals, but also of a variety of other contaminants, such as viruses (Schijven and Šimůnek, 2001), colloids (Bradford et al., 2002), cadmium (Seuntjens et al., 2001), and hormones (Casey et al., 2003, 2004), or chemicals involved in the sequential biodegradation of chlorinated aliphatic hydrocarbons (Schaerlaekens et al., 1999; Casey and Šimůnek, 2001).

Much effort has recently been directed also toward improving models for purposes of simulating nonequilibrium and/or preferential flow. Examples of this are the TOUGH codes (Pruess, 1991, 2004), MACRO (Jarvis, 1994), and HYDRUS-1D (Šimůnek et al., 2003, 2005). These models typically assume the presence of dual-porosity and dual-permeability regions, with different fluxes possible in the two regions. As discussed in detail in Section 22.2b, dual-porosity, mobile–immobile water flow models result when the Richards or kinematic wave equations are used for flow in the fractures, and immobile water is assumed to exist in the matrix, whereas dual-permeability models assume that water is mobile in both the matrix and fracture domains, while invoking terms accounting for the exchange of water and solutes between the two regions. Example applications of these dual-porosity and dual-permeability models are given by Šimůnek et al. (2001), Haws et al. (2005), Köhne et al. (2004b, 2005), and Pot et al. (2005), among many others.

As an example of available vadose zone flow and transport models, we discuss here in somewhat greater detail the HYDRUS software packages of Šimůnek et al. (1999b, 2005) and the MODFLOW-SURFACT model developed by HydroGeoLogic (1996).

22.5.2.1.1 *The HYDRUS Software Packages*

HYDRUS-1D (Šimůnek et al., 2005) and HYDRUS-2D (Šimůnek et al., 1999b) are software packages (www.hydrus2d.com) that simulate the one- and two-dimensional movement of water, heat, and multiple solutes in variably saturated media, respectively. Both programs use finite elements to numerically solve the Richards equation for saturated–unsaturated water flow and Fickian-based advection–dispersion equations for both heat and solute transport. The flow equation includes a sink term to account for water uptake by plant roots as a function of both water and salinity stress. The flow equation may also consider dual-porosity-type flow with a fraction of water content being mobile, and fraction immobile. The heat transport equation considers conduction as well as advection with flowing water. The solute transport equations assume advective-dispersive transport in the liquid phase, and diffusion in the gaseous phase. The transport equations also include provisions for nonlinear and/or nonequilibrium reactions between the solid and liquid phases, linear equilibrium reactions between the liquid and gaseous phases, zero-order production, and two first-order degradation reactions: one which is independent of other solutes, and one which provides the coupling between solutes involved in the sequential first-order decay reactions. In addition, physical nonequilibrium solute transport can be accounted for by assuming a two-region, dual-porosity type formulation which partitions the liquid phase into mobile and immobile regions. Alternatively, the transport equations include provisions for kinetic attachment/detachment of solute to the solid phase and thus can be used to simulate the transport of viruses, colloids, or bacteria (Schijven and Šimůnek, 2002; Bradford et al., 2003).

Both programs may be used to analyze water and solute movement in unsaturated, partially saturated, or fully saturated media. The flow region itself may consist of nonuniform (layered) soils. HYDRUS-2D additionally can handle flow regions delineated by irregular boundaries. The flow region itself may be

composed of nonuniform soils having an arbitrary degree of local anisotropy. Flow and transport can occur in the vertical plane, the horizontal plane, or in a three-dimensional region exhibiting radial symmetry about the vertical axis. The unsaturated soil hydraulic properties (the constitutive relationships) can be described using van Genuchten (1980a), Brooks and Corey (1964), Kosugi (1996), and Durner (1994) type analytical functions, or modified van Genuchten type functions that produce a better description of the hydraulic properties near saturation. HYDRUS-1D incorporates hysteresis by assuming that drying scanning curves are scaled from the main drying curve, and wetting scanning curves from the main wetting curve. Root growth is simulated by means of a logistic growth function, while root water uptake can be simulated as a function of both water and salinity stress.

The HYDRUS-1D software package additionally includes modules for simulating carbon dioxide and major ion solute movement (Šimůnek et al., 1996). Diffusion in both the liquid and gas phases and advection in the liquid phase are considered as the main CO₂ transport mechanisms (Šimůnek and Suarez, 1993). The major variables of the chemical system are Ca, Mg, Na, K, SO₄, Cl, NO₃, H₄SiO₄, alkalinity, and CO₂ (Šimůnek and Suarez, 1994; Šimůnek et al., 1996). The model accounts for equilibrium chemical reactions between these components such as complexation, cation exchange and precipitation-dissolution. For the precipitation-dissolution of calcite and dissolution of dolomite, either equilibrium or multicomponent kinetic expressions are used, involving both forward and back reactions. Other dissolution-precipitation reactions considered include gypsum, hydromagnesite, nesquehonite, and sepiolite. Since the ionic strength of soil solutions can vary considerably with time and space and often reach high values, both modified Debye-Hückel and Pitzer expressions were incorporated into the model as options to calculate single ion activities.

In addition, both HYDRUS programs implement a Marquardt-Levenberg type parameter estimation technique (Marquardt, 1963; Šimůnek and Hopmans, 2002) for inverse estimation of soil hydraulic (Šimůnek et al., 1998b,c; Hopmans et al., 2002) and/or solute transport and reaction (Šimůnek et al., 2002) parameters from measured transient or steady-state flow and/or transport data.

The HYDRUS packages use a Microsoft Windows based Graphics User Interface (GUI) to manage the input data required to run the program, as well as for nodal discretization and editing, parameter allocation, problem execution, and visualization of results. All spatially distributed parameters, such as those for various soil horizons, the root water uptake distribution, and the initial conditions for water, heat and solute movement, are specified in a graphical environment. HYDRUS-2D includes the MeshGen2D mesh generator, specifically designed for variably-saturated subsurface flow transport problems, for defining more general domain geometries, and for discretizing the transport domain into an unstructured finite element mesh. The program offers graphs of the distributions of the pressure head, water content, water and solute fluxes, root water uptake, temperature, and solute concentrations in the soil profile at pre-selected times. Also included is a small catalog of unsaturated soil hydraulic properties (Carsel and Parish, 1988), as well as pedotransfer functions based on neural networks (Schaap et al., 2001). Many different applications of both software packages were referenced throughout this book chapter.

22.5.2.1.2 MODFLOW-SURFACT

MODFLOW-SURFACT (HydroGeoLogic, 1996) is a fully integrated flow and transport code, based on the MODFLOW groundwater modeling software package of the U.S. Geological Survey (McDonald and Harbaugh, 1988). While MODFLOW (see detailed discussion in Chapter 23) deals only with fully saturated groundwater flow, MODFLOW-SURFACT expands the applicability of the code to unsaturated domains. MODFLOW-SURFACT includes new flow packages that enhance schemes for performing unconfined simulations to rigorously model desaturation and resaturation of aquifers. It provides an option for discretizing the domain using an axi-symmetric geometry for efficient simulation of pumping tests, recovery tests, etc. In addition to normally allowed external stresses in MODFLOW (i.e., constant head, constant flux, areal recharge, evapotranspiration, drains, and streams), MODFLOW-SURFACT provides a rigorous well withdrawal package, unconfined recharge boundary conditions, and seepage face boundary conditions. Additionally, MODFLOW-SURFACT also includes options for adaptive time-stepping and

output control procedures, and a new solution package based on the preconditioned conjugate gradient method.

MODFLOW-SURFACT includes new Analysis of Contaminant Transport (ACT) modules that allow simulations of single-species and multi-component contaminant transport. The program can handle up to 5 chemical species, which may undergo linear or nonlinear retardation, first-order decay and biochemical degradation, as well as generate transformation products. MODFLOW-SURFACT also permits mass partitioning of a single or multicomponent contaminant with diffusive mass movement in the inactive phase, and mass partitioning of a single or multicomponent contaminant from a depleting immobile NAPL phase with advective and dispersive transport in the active phase and diffusive transport in the inactive phase.

22.5.2.2 Geochemical Transport Models

Significant efforts have recently been carried out also in coupling physical flow and transport models with geochemical models to simulate ever more complex reactions, such as surface complexation, precipitation/dissolution, cation exchange, and biological reactions. Recent reviews of the development of hydrogeochemical transport models involving reactive multiple components are given by Mangold and Tsang (1991), Lichtner (1996) and Steefel and MacQuarrie (1996), Šimůnek and Valocchi (2002), and Bell and Binning (2004). Most modeling efforts involving multicomponent transport have thus far focused on the saturated zone, where changes in the water flow velocity, temperature and pH are often much more gradual and hence less important than in the unsaturated zone. Consequently, most multicomponent transport models assumed one- or two-dimensional steady-state saturated water flow with a fixed value of the flow velocity, temperature and pH. Only recently have several multicomponent transport models been published which also consider variably saturated flow; these include DYNAMIX (Liu and Narasimhan, 1989), HYDROGEOCHEM (Yeh and Tripathi, 1990), TOUGH-REACT (Pruess, 1991), UNSATCHEMA (Šimůnek and Suarez, 1994; Šimůnek et al., 1996, 1997), FEHM (Zyvoloski et al., 1997), MULTIFLO (Lichtner and Seth, 1996), OS3D/GIMRT (Steefel and Yabusaki, 1996), HYDROBIOGEOCHEM (Yeh et al., 1998), FLOTTRAN (Lichtner, 2000), MIN3P (Mayer et al., 2002), HP1 (Jacques et al., 2002; Jacques and Šimůnek, 2005), and HYDRUS-1D (Šimůnek et al., 2005).

Geochemical models can be divided into two major groups: those with specific chemistry and general models (Šimůnek and Valocchi, 2002). Models with specific chemistry are limited in the number of species they can handle, while their application is restricted to problems having a prescribed chemical system. They are, however, much easier to use and computationally can be much more efficient than general models. Typical examples of models with specified chemistry are those simulating the transport of major ions, such as LEACHM (Wagenet and Hutson, 1987), UNSATCHEMA (Šimůnek and Suarez, 1994; Šimůnek et al., 1996), and HYDRUS-1D (Šimůnek et al., 2005). Models with generalized chemistry (DYNAMIX, HYDROGEOCHEM, MULTIFLO, FLOTTRAN, OS3D/GIMRT, and HP1, all referenced above) provide users with much more freedom in designing a particular chemical system; possible applications of these models are also much wider.

22.5.2.2.1 HP1

HYDRUS-1D was recently coupled with the PHREEQC geochemical code (Parkhurst and Appelo, 1999) to create a new comprehensive simulation tool, HP1 (acronym for HYDRUS1D-PHREEQC) (Jacques et al., 2003; Jacques and Šimůnek, 2005). The combined code contains modules simulating (1) transient water flow in variably saturated media, (2) the transport of multiple components, (3) mixed equilibrium/kinetic biogeochemical reactions, and (4) heat transport. HP1 is a significant expansion of the individual HYDRUS-1D and PHREEQC programs by preserving most of their original features and capabilities. The code still uses the Richards equation for simulating variably saturated water flow and advection-dispersion type equations for heat and solute transport. However, the program can now simulate also a broad range of low-temperature biogeochemical reactions in water, the vadose zone and in ground water systems, including interactions with minerals, gases, exchangers, and sorption surfaces, based on thermodynamic equilibrium, kinetics, or mixed equilibrium-kinetic reactions.

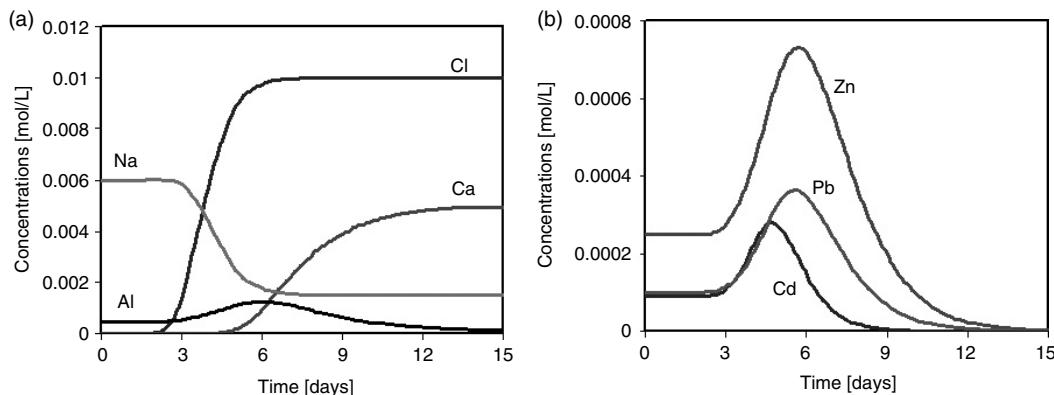


FIGURE 22.13 Major ions (a) and heavy metals (Zn, Pb, and Cd) (b) concentrations in the effluent of an 8-cm soil long column.

Jacques et al. (2003) and Jacques and Šimůnek (2005) demonstrated the versatility of the HP1 model on several examples such as (a) the transport of heavy metals (Zn^{2+} , Pb^{2+} , and Cd^{2+}) subject to multiple cation exchange reactions, (b) transport with mineral dissolution of amorphous SiO_2 and gibbsite ($Al(OH)_3$), (c) heavy metal transport in a medium with a pH-dependent cation exchange complex, (d) infiltration of a hyperalkaline solution in a clay sample (this example considers kinetic precipitation-dissolution of kaolinite, illite, quartz, calcite, dolomite, gypsum, hydrotalcite, and sepiolite), (e) long-term transient flow and transport of major cations (Na^+ , K^+ , Ca^{2+} , and Mg^{2+}) and heavy metals (Cd^{2+} , Zn^{2+} , and Pb^{2+}) in a soil profile, (f) cadmium leaching in acid sandy soils, (g) radionuclide transport (U and its aqueous complexes), and (h) the fate and subsurface transport of explosives (TNT and its daughter products 2ADNT, 4ADNT, and TAT) (Šimůnek et al., 2006).

22.5.2.2 Leaching of Heavy Metals from a Soil Column

As an example application of HP1, Figure 22.13 shows calculated effluent concentrations of heavy metals (Zn^{2+} , Pb^{2+} , and Cd^{2+}) leached from an 8-cm long soil column having an initial solution as defined in Table 22.2, and with its ion-exchange complex in equilibrium with this solution. Heavy metals initially present in the soil column are leached from the exchange complex by major ions (Ca^{2+} , Mg^{2+} , and Al^{3+}). Water was applied to the top of the column at a steady flow rate of 2 cm day^{-1} and having a chemical composition as given in Table 22.2. Details are given by Jacques and Šimůnek (2005).

22.6 Concluding Remarks

This chapter demonstrates the abundance of models and modeling approaches that are currently available for simulating variably saturated water flow and contaminant transport at various levels of approximation and for different applications. Models range from relatively simple analytical approaches for analyzing solute transport problems during one-dimensional steady-state flow, to sophisticated numerical models for addressing multi-dimensional variably saturated flow and contaminant transport problems at the field scale.

One may expect that unsaturated zone flow and transport models will be used increasingly as tools for developing cost-effective, yet technically sound strategies for resource management and pollution remediation and prevention. Improved understanding of the key underlying processes, continued advances in numerical methods, and the introduction of more and more powerful computers make such simulations increasingly practical for many field-scale problems. For example, models can be helpful tools for designing, testing and implementing soil, water and crop management practices that minimize soil and water pollution. Models are equally needed for designing or remediating industrial waste disposal

TABLE 22.2 Chemical Components (mmol/L) and Species Considered, and Elemental Compositions of Initial and Boundary Solutions Used in the Column Simulation. X Refers to Ion Exchanger

Components	Species	Solutions	
		Boundary	Initial
Al	Al ³⁺ , Al(OH) ²⁺ , Al(OH) ₂ ⁺ , Al(OH) ₃ ⁰ , Al(OH) ₄ ⁻	0.1	0.5
Br	Br ⁻	3.7	11.9
Cl	Cl ⁻ (and Cd, Pb, and Zn-species, see below)	10	0
Ca	Ca ²⁺ , Ca(OH) ⁺	5	0
K	K ⁺ , KOH ⁰	0	2
Na	Na ⁺ , NaOH ⁰	0	6
Mg	Mg ²⁺ , Mg(OH) ⁺	1	0.75
Cd	Cd ²⁺ , Cd(OH) ⁺ , Cd(OH) ₂ ⁰ , Cd(OH) ₃ ⁻ , Cd(OH) ₄ ²⁻ , CdCl ⁺ , CdCl ₂ , CdCl ₃ ⁻	0	0.9
Pb	Pb ²⁺ , Pb(OH) ⁺ , Pb(OH) ₂ ⁰ , Pb(OH) ₃ ⁻ , Pb(OH) ₄ ²⁻ , PbCl ⁺ , PbCl ₂ ⁰ , PbCl ₃ ⁻ , PbCl ₄ ²⁻	0	0.1
Zn	Zn ²⁺ , Zn(OH) ⁺ , Zn(OH) ₂ ⁰ , Zn(OH) ₃ ⁻ , Zn(OH) ₄ ²⁻ , ZnCl ⁺ , ZnCl ₂ ⁰ , ZnCl ₃ ⁻ , ZnCl ₄ ²⁻	0	0.25
X	AlX ₃ , AlOHX ₂ , CaX ₂ , CdX ₂ , KX, NaX, MgX ₂ , PbX ₂ , ZnX ₂ (mmol)	NA	11

sites and landfills, for predicting contaminant transport from mining wastes, or for long-term stewardship of nuclear waste repositories. The main challenge is to make the models as realistic as possible for the various applications.

Continued progress in subsurface flow and transport modeling requires equal advances in both numerical techniques as well as the underlying science. Addressing preferential flow phenomena, and the related problems of subsurface heterogeneity, including the stochastic nature of boundary conditions (precipitation and/or evapotranspiration), will continue to pose formidable challenges. The same is true for improving multicomponent geochemical transport modeling for the vadose zone. For example, numerical algorithms and databases for multicomponent transport models must be extended to higher temperatures and ionic strengths, complex contaminant mixtures (including especially mixed organic and inorganic wastes), multiphase flow, redox disequilibria for low-temperature systems, and coupled physico-chemical systems to account for possible changes in the soil water retention and hydraulic conductivity functions. Better integration is also needed between variably saturated subsurface and existing larger-scale surface numerical models, which in turn requires further research on such issues as spatial and temporal scaling of hydrological, chemical and biological processes and properties, linking constitutive (soil hydraulic) relationships to measurements scales, preferential flow, and issues of parameter and model uncertainty.

Many science questions related to colloid and colloid-facilitated transport are also still largely unresolved. This is an area of research where our understanding lags far behind current numerical capabilities. Much work is needed to better understand the processes of filtration, straining, size exclusion, colloid–colloid interactions, mobilization of colloids and microorganisms; accumulation at air–water interfaces, interactions between microorganisms and contaminants (including biodegradation), the effects of both physical factors (water content, flow velocity, textural interfaces) and chemical processes (ionic strength, solution composition, pH) on colloid retention and mobilization, and modeling colloid-facilitated transport during conditions of transient flow.

Also, to the best of our knowledge, no models are currently available that consider all the various processes simultaneously and in their full complexity, including their mutual interactions. That is, no models exist that consider transient preferential flow and transport in structured soils or fractured rocks, while simultaneously considering complex biogeochemical reactions between contaminants, organic and inorganic colloids and/or organic complexes, and solid and air phases of a soil, including widely

varying rates of these various reactions. Further integration of the different types of numerical models is needed to address practical problems of contaminant transport (trace elements, radionuclides, organic contaminants) in complex vadose zone environments.

References

- Abriola, L.M., Pennell, K.D., Weber, Jr., W.J. Lang, J.R., and Wilkins, M.D., Persistence and interphase mass transfer of organic contaminants in the unsaturated zone: Experimental observations and mathematical modeling. In J.Y. Parlange and J.W. Hopmans (Eds.) *Vadose Zone Hydrology: Cutting Across Disciplines*. Oxford University Press, New York, pp. 210–234, 1999.
- Adamczyk, Z., Siwek, B., Zembala, M. and Belouschek, P., Kinetics of localized adsorption of colloid particles. *Adv. Coll. Interf. Sci.* 48, 151–280, 1994.
- Ahuja, L.R. and Hebson, C., *Root Zone Water Quality Model*. GPSR Tech. Rep. No. 2, USDA, ARS, Fort Collins, CO, 1992.
- Anderson, M.P., Movement of contaminants in groundwater: Groundwater transport–advection and dispersion. In *Groundwater Contamination, Studies in Geophysics*, National Academy Press, Washington, D.C. pp. 37–45, 1984.
- Appelo, C.A.J. and Postma, D., *Geochemistry, Groundwater and Pollution*. A. A. Balkema. Rotterdam, 1993.
- Bache, B.W. and Williams, E.G., A phosphate sorption index for soils. *J. Soil Sci.* 27, 289–296, 1971.
- Barry, D.A., Modelling contaminant transport in the subsurface: theory and computer programs. In H. Ghadiri and C.W. Rose (Eds.), *Modelling Chemical Transport in Soil: Natural and Applied Contaminants*, Lewis Publishers, Chelsea, MI, pp. 105–144, 1992.
- Bateman, H., The solution of a system of differential equations occurring in the theory of radioactive transformation. *Proc. Cambridge Philos. Soc.* 15 (pt V) 423–427, 1910.
- Bear, J., *Dynamics of Fluids in Porous Media*, Elsevier Science, New York, 1972.
- Bell, L.S.J. and Binning, P.J., A split operator approach to reactive transport with the forward particle tracking Eulerian–Lagrangian localized adjoint method. *Adv. Water Resour.* 27, 323–334, 2004.
- Bethke, C. M., *Geochemical Reaction Modeling — Concepts and Applications*. Oxford University Press, New York, p. 397, 1996.
- Biggar, J.W. and Nielsen, D.R., Miscible displacement and leaching phenomena. *Agronomy* 11, 254–274, 1967.
- Bradford, S.A., Yates, S.R., Bettahar, M. and Šimůnek, J., Physical factors affecting the transport and fate of colloids in saturated porous media. *Water Resour. Res.* 38, 1327, doi:10.1029/2002WR001340, 2002.
- Bradford, S.A., Šimůnek, J., Bettahar, M., van Genuchten, M.Th. and Yates, S.R., Modeling colloid attachment, straining, and exclusion in saturated porous media. *Environ. Sci. Technol.* 37, 2242–2250, 2003.
- Bradford, S.A., Bettahar, M., Šimůnek, J. and van Genuchten, M.Th., Straining and attachment of colloids in physically heterogeneous porous media. *Vadose Zone J.* 3, 384–394, 2004.
- Bresler, E. and Dagan G., Solute dispersion in unsaturated heterogeneous soil at field scale: I Applications. *Soil Sci. Soc. Am. J.* 43, 467–472, 1979.
- Brooks, R.H. and Corey, A. T., Hydraulic properties of porous media. *Hydrol. Paper No. 3*, Colorado State University, Fort Collins, CO, 1964.
- Brusseau, M.L., Nonideal transport of reactive solutes in porous media: Cutting across history and disciplines. In Parlange, M.B. and Hopmans, J.W. (Eds.), *Vadose Zone Hydrology: Cutting Across Disciplines*, Oxford University Press, New York, pp. 130–154, 1999.
- Carsel, R.F. and Parrish, R.S., Developing joint probability distributions of soil water retention characteristics. *Water Resour. Res.* 24, 755–769, 1988.
- Casey, F.X.M. and Šimůnek, J., Inverse analyses of the transport of chlorinated hydrocarbons subject to sequential transformation reactions. *J. Environ. Qual.* 30, 1354–1360, 2001.
- Casey, F.X.M., Larsen, G.L., Hakk, H. and Šimůnek, J., Fate and transport of 17β -Estradiol in soil-water systems. *Environ. Sci. Technol.* 37, 2400–2409, 2003.

- Casey, F.X.M., Larsen, G.L., Hakk, H. and Šimůnek, J., Fate and transport of testosterone in agriculturally significant soils. *Environ. Sci. Technol.* 38, 790–798, 2004.
- Corapcioglu, M.Y. and Jiang, S., Colloid-facilitated ground-water contaminant transport. *Water Resour. Res.* 29, 2215–2226, 1993.
- Corapcioglu, M.Y. and Kim, S., Modeling facilitated contaminant transport by mobile bacteria. *Water Resour. Res.* 31, 2639–2644, 1995.
- Dagan G. and E. Bresler, Solute dispersion in unsaturated heterogeneous soil at field scale: I Theory. *Soil Sci. Soc. Am. J.* 43, 461–467, 1979.
- Dagan G., *Flow and Transport in Porous Formations*, Springer-Verlag, Berlin, 1989.
- Durner, W., Hydraulic conductivity estimation for soils with heterogeneous pore structure. *Water Resour. Res.* 32, 211–223, 1994.
- Fayer, M.J., UNSAT-H Version 3.0: Unsaturated Soil Water and Heat Flow Model. Theory, User Manual, and Examples. *Pacific Northwest National Laboratory* 13249, 2000.
- Fitter, A.H. and Sutton, S.D., The use of the Freundlich isotherm for soil phosphate sorption data. *J. Soil Sci.* 26, 241–246, 1975.
- Flerchinger, G.N., Hanson, C.L. and Wight, J.R., Modeling evapotranspiration and surface energy budgets across a watershed. *Water Resour. Res.* 32, 2539–2548, 1996.
- Flury, M., Fluhler, H., Jury, W. and Leuenberger, J., Susceptibility of soils to preferential flow of water: a field study. *Water Resour. Res.* 30, 1945–1954, 1994.
- Freundlich, H., *Kapillarchemie; eine Darstellung der Chemie der Kolloide und verwandter Gebiete*, Akademische Verlagsgesellschaft, Leipzig, Germany, 1909.
- Gelhar, L.W., Mantaglo, A., Welty, C. and Rehfeldt, K.R., A review of Field Scale Physical Solute Transport Processes in Unsaturated and Saturated Porous Media, EPRI Topic Report EA-4190, Electric Power Research Institute, Palo Alto, CA, 1985.
- Gerke, H.H. and van Genuchten, M.Th., A dual-porosity model for simulating the preferential movement of water and solutes in structured porous media. *Water Resour. Res.* 29, 305–319, 1993a.
- Gerke, H.H. and van Genuchten, M.Th., Evaluation of a first-order water transfer term for variably saturated dual-porosity flow models. *Water Resour. Res.* 29, 1225–1238, 1993b.
- Gerke, H.H. and van Genuchten, M.Th., Macroscopic representation of structural geometry for simulating water and solute movement in dual-porosity media. *Adv. Water Resour.* 19, 343–357, 1996.
- Germann, P.F., Kinematic wave approach to infiltration and drainage into and from soil macropores. *Trans. ASAE* 28, 745–749, 1985.
- Germann, P.F. and Beven, K., Kinematic wave approximation to infiltration into soils with sorbing macropores. *Water Resour. Res.* 21, 990–996, 1985.
- Glotfelty, D.E. and Schomburg, C.J., Volatilization of pesticides. In B.L. Sawhney and K. Brown (Eds.), *Reactions and Movement of Organic Chemicals in Soils*, Special Publ. No. 22, Soil Science Society of America, Madison, WI, pp. 181–208, 1989.
- Grolimund, D., Borkovec, M., Barmettler, K. and Sticher, H., Colloid-facilitated transport of strongly sorbing contaminants in natural porous media; a laboratory column study. *Environ. Sci. Technol.* 30, 3118–3123, 1996.
- Gunary, D., A new adsorption isotherm for phosphate in soil. *J. Soil Sci.* 21, 72–77, 1970.
- Gwo, J.P., Jardine, P.M., Wilson, G.V. and Yeh, G.T., A multiple-pore-region concept to modeling mass transfer in subsurface media. *J. Hydrol.* 164, 217–237, 1995.
- Hanselman, T.A., Graetz, D.A. and Wilkie, A.C., Manure-borne estrogens as potential environmental contaminants: A review. *Environ. Sci. Technol.* 37, 5471–5478, 2003.
- Hansen, S., Jensen, H.E., Nielsen, N.E. and Svendsen, H., DAISY: Soil Plant Atmosphere System Model. *NPO Report No. A 10*. The National Agency for Environmental Protection, Copenhagen, 272 pp, 1990.
- Haws, N.W., Rao, P.S.C., Šimůnek, J., and Poyer, I.C., Single-porosity and dual-porosity modeling of water flow and solute transport in subsurface-drained fields using effective field-scale parameters. *J. Hydrol.* 313, 257–273, 2005.

- Healy, R.W., Simulation of solute transport in variably saturated porous media with supplemental information on modifications to the U.S. Geological Survey's computer program VS2D. U.S. Geological Survey, *Water-Resources Investigation Report 90-4025*: 125, 1990.
- Hendrickx, J.M.H., Dekker, L.W. and Boersma, O.H., Unstable wetting fronts in water repellent field soils. *J. Env. Qual.* 22, 109–118, 1993.
- Hendrickx, J.M.H. and Flury, M., Uniform and preferential flow, Mechanisms in the vadose zone. In *Conceptual Models of Flow and Transport in the Fractured Vadose zone*. National Research Council, National Academy Press, Washington, D.C., pp. 149–187, 2001.
- Hopmans, J.W., Šimůnek, J., Romano, N. and Durner, W., Inverse modeling of transient water flow. In J.H. Dane and G.C. Topp (Eds.), *Methods of Soil Analysis, Part 1, Physical Methods*, Chapter 3.6.2, Third edition, SSSA, Madison, WI, pp. 963–1008, 2002.
- Hutson, J.L. and Wagenet, R.J., A multiregion model describing water flow and solute transport in heterogeneous soils. *Soil Sci. Soc. Am. J.* 59, 743–751, 1995.
- Huyakorn, P.S. and Pinder, G.F., *Computational Methods in Subsurface Flow*, Academic Press, London, 1983.
- HydroGeoLogic, Inc., MODFLOW-SURFACT, A Comprehensive MODFLOW-Based Flow and Transport Simulator, Version 2.1, HydroGeoLogic, Inc., 1996.
- Jacques, D., Šimůnek, J., Mallants, D. and van Genuchten, M. Th., Multicomponent transport model for variably-saturated porous media: Application to the transport of heavy metals in soils. In S.M., Hassanzadeh, R.J. Schotting, W.G. Gray, and G.F. Pinder (Eds.) *Computational Methods in Water Resources*, XIVth International Conference on Computational Methods in Water Resources, June 23–28, 2002, Delft, The Netherlands. Elsevier, *Developments in Water Science*, 47, 555–562, 2002.
- Jacques, D., Šimůnek, J., Mallants, D. and van Genuchten, M. Th., The HYDRUS-PHREEQC multicomponent transport model for variably-saturated porous media: Code verification and application, MODFLOW and More 2003: Understanding through Modeling, Conference Proceedings, E. Poeter, Ch. Zheng, M. Hill, and J. Doherty, (Eds.), Int. Ground Water Modeling Center, Colorado School of Mines, pp. 23–27, 2003.
- Jacques, D. and Šimůnek, J., Multicomponent — Variable Saturated Transport Model, Description, Manual, Verification and Examples, Waste and Disposal, SCKCEN, BLG-998, Mol, Belgium, 79 pp, 2005.
- Jansson, P.-E. and Karlberg, L., Coupled heat and mass transfer model for soil-plant-atmosphere systems. Royal Institute of Technology, Dept of Civil and Environmental Engineering, Stockholm 325 pp, 2001.
- Jarvis, N.J., The MACRO model (Version 3.1), Technical description and sample simulations. Reports and Dissertations 19. Dept. Soil Sci., Swedish Univ. Agric. Sci., Uppsala, Sweden, 51 pp, 1994.
- Jarvis, N. J., Modeling the impact of preferential flow on nonpoint source pollution. In H.M. Selim and L. Ma (Eds.), *Physical Nonequilibrium in Soils: Modeling and Application*. Ann Arbor Press, Chelsea, MI, pp. 195–221, 1998.
- Jiang, S. and Corapcioglu, M.Y., A hybrid equilibrium model of solute transport in porous media in the presence of colloids. *Coll. Surf. A: Physicochem. Eng. Aspects* 73, 275–286, 1993.
- Jury, W.A., Simulation of solute transport using a transfer function model. *Water Resour. Res.* 18, 363–368, 1982.
- Jury, W.A., Spencer, W.F. and Farmer, W.J., Behavior assessment model for trace organics in soil: I. Description of model. *J. Env. Qual.* 12, 558–564, 1983.
- Jury, W.A., Spencer W.F. and Farmer, W.J., Behavior assessment model for trace organics in soil: I. Chemical classification and parameter sensitivity. *J. Env. Qual.* 13, 567–572, 1984.
- Jury W.A. and Sposito, G., Field calibration and validation of solute transport models for the unsaturated zone. *Soil Sci. Soc. Am. J.* 49, 1331–1341, 1985.
- Jury, W.A. and Horton, R., *Soil Physics*, Sixth ed, John Wiley & Sons, Inc., New York, 2004.

- Kan, A.T. and Tomson, M.B., Ground water transport of hydrophobic organic compounds in the presence of dissolved organic matter. *Env. Toxicol. Chem.* 9, 253–263, 1990.
- Kodešová, R., Kozák, J., Šimůnek, J. and Vacek, O., Field and numerical study of chlorotoluron transport in the soil profile: Comparison of single and dual-permeability model. *Plant, Soil Environ.* 51, 310–315, 2005.
- Köhne, J.M., Köhne, S., Mohanty, B.P. and Šimůnek, J., Inverse mobile–immobile modeling of transport during transient flow: Effect of between-domain transfer and initial water content. *Vadose Zone J.* 3, 1309–1321, 2004a.
- Köhne, J.M., B. Mohanty, J. Šimůnek and H.H. Gerke, Numerical evaluation of a second-order water transfer term for variably saturated dual-permeability models, *Water Resour. Res.* 40, doi:10.1029/2004WR00385, 2004b.
- Köhne, J.M. and Mohanty, B.P., Water flow processes in a soil column with a cylindrical macropore: Experiment and hierarchical modeling. *Water Resour. Res.* 41, W03010, doi:10.1029/2004WR003303, 2005.
- Köhne, S., Lennartz, B., Köhne, J.M. and Šimůnek, J., Bromide transport at a tile-drained field site: Experiment, one- and two-dimensional equilibrium and non-equilibrium numerical modeling. *J. Hydrol.* 321, 390–408, 2006.
- Köhne, J.M., Köhne, S. and Šimůnek, J., Multi-process herbicide transport in structured soil columns: Experiment and model analysis. *J. Contamin. Hydrol.* 85, 1–32, 2006.
- Kosugi, K., Lognormal distribution model for unsaturated soil hydraulic properties, *Water Resour. Res.* 32, 2697–2703, 1996.
- Kung, K.-J.S., Preferential flow in a sandy vadose zone. 2. Mechanism and implications. *Geoderma* 46, 59–71, 1990.
- Lai, Sung-Ho and Jurinak, J.J., Numerical approximation of cation exchange in miscible displacement through soil columns. *Soil Sci. Soc. Am. Proc.* 35, 894–899, 1971.
- Langmuir, I., The adsorption of gases on plane surfaces of glass, mica, and platinum. *J. Am. Chem. Soc.* 40, 1361–1403, 1918.
- Lapidus, L. and Amundson, N.R., Mathematics of adsorption in beds: VI. The effect of longitudinal diffusion in ion exchange and chromatographic columns. *J. Phys. Chem.* 56, 984–988, 1952.
- Leij, F.J. and Bradford, S.A., 3DADE: A computer program for evaluating three-dimensional equilibrium solute transport in porous media. *Research Report No.* 134, U.S. Salinity Laboratory, USDA, ARS, Riverside, CA, 1994.
- Leij, F.J. and Toride, N., N3DADE: A computer program for evaluating nonequilibrium three-dimensional equilibrium solute transport in porous media. *Research Report No.* 143, U.S. Salinity Laboratory, USDA, ARS, Riverside, CA, 1997.
- Leij, F.J., Priesack, E. and Schaap, M.G., Solute transport modeled with Green's functions with application to persistent solute sources. *J. Contam. Hydrol.* 41, 155–173, 2000.
- Lenhard, R.J., Oostrom, M. and Dane, J.H., Multi-fluid flow. In J.H. Dane and G.C. Topp (Eds.), *Methods of Soil Analysis, Part 4, Physical Methods*, SSSA Book Series 5, Soil Sci. Soc. Am., Madison, WI, pp. 1537–1619, 2002.
- Lichtner, P.C., Continuum formulation of multicomponent-multiphase reactive transport. In P.C. Lichtner, C.I. Steefel and E.H. Oelkers (Eds.), *Reactive Transport in Porous Media, Reviews in Mineralogy*. Chapter 1, Vol. 34. Mineralogical Society of America, pp. 1–81, 1996.
- Lichtner, P.C. and Seth, M.S., User's manual for MULTIFLO: Part II MULTIFLOW 1.0 and GEM 1.0. Multicomponent-multiphase reactive transport model. CNWRA 96-010, Center for Nuclear Waste Regulatory Analyses, San Antonio, TX, 1996.
- Lichtner, P.C., FLOWTRAN User's manual. Los Alamos National Laboratory Document. NM, 2000.
- Lindstrom, F.T., Haque, R., Freed, V.H. and Boersma, L., Theory on the movement of some herbicides in soils. *Environ. Sci. Technol.* 1, 561–565, 1967.

- Lindstrom, F.T., Boersma, L. and Stockard, D., A theory on the mass transport of previously distributed chemicals in a water saturated sorbing porous medium. Isothermal cases, *Soil Sci.* 112, 291–300, 1971.
- Lindqvist, R. and Enfield, C.G., Biosorption of dichlorodiphenyltri-chloroethane and hexachlorobenzene in groundwater and its implications for facilitated transport. *Appl. Environ. Microbiol.* 58, 2211–2218, 1992.
- Liu, C.W. and Narasimhan, T.N., Redox-controlled multiple-species reactive chemical transport. 1. Model development. *Water Resour. Res.* 25:869–882, 1989.
- Liu, C., Szecsody, J.E., Zachara, J.M. and Ball, W.P., Use of the generalized integral transform method for solving equations of solute transport in porous media. *Adv. Water. Resour.* 23, 483–492, 2000.
- Logan, B.E., Jewett, D.G., Arnold, R.G., Bouwer, E.J. and O'Melia, C.R., Clarification of clean-bed filtration models. *J. Environ. Eng.* 121, 869–873, 1995.
- Magee, B.R., Lion, L.W. and Lemley, A.T., Transport of dissolved organic macromolecules and their effect on the transport of phenanthrene in porous media. *Environ. Sci. Technol.* 25, 323–331, 1991.
- Mangold, D.C. and Tsang, A summary of subsurface hydrological and hydrochemical models. *Rev. Geophys.* 29, 51–79, 1991.
- Mansfeldt, T., Leyer, H., Barmettler, K. and Kretzschmar, R., Cyanide leaching from soil developed from coking plant purifier waste as influenced by citrate. *Vadose Zone J.* 3, 471–479, 2004.
- Marquardt, D.W., An algorithm for least-squares estimation of nonlinear parameters. *SIAM J. Appl. Math.* 11, 431–441, 1963.
- Mayer, U., Frind, E.O. and Blowes, D.W., Multicomponent reactive transport modeling in variably saturated porous media using a generalized formulation for kinetically controlled reactions. *Water Resour. Res.* 38, 1174, 2002.
- McDonald, M.G. and Harbaugh, A.W., A modular three-dimensional finite-difference ground-water flow model, U.S. Geological Survey Techniques of Water-Resources Investigations Book 6, Chapter A1, 586 p, 1988.
- MDH Engineered Solutions Corp., Evaluation of Computer Models for Predicting the Fate and Transport of Salt in Soil and Groundwater, Phase I Report. Pub. No: T/403, ISBN No. 0-7785-2493-0 (Printed Edition), ISBN No. 0-7785-2494-9 (On-line Edition), 2003.
- Millington, R.J. and J.M. Quirk, Permeability of porous solids. *Trans. Faraday Soc.* 57, 1200–1207, 1961.
- Mills W.M., Liu S. and Fong, F.K., Literature review and model (COMET) for colloid/metals transport in porous media. *Groundwater* 29, 199–208, 1991.
- Moldrup, O., Poulsen, T.G., Schjonning, P., Olesen, T. and Yamaguchi, T., Gas permeability in undisturbed soils: measurements and predictive models. *Soil Sci.* 163, 180–189, 1998.
- Narasimhan, T.N., Buckingham, 1907; An appreciation. *Vadose Zone J.* 4, 434–441, 2005.
- Nielsen, D.R., van Genuchten, M. Th. and Biggar, J.W., Water flow and solute transport processes in the unsaturated zone. *Water Resour. Res.* 22, 89S–108S, 1986.
- Noell, A.L., Thompson, J.L., Corapcioglu, M.Y. and Triay, I.R., The role of silica colloids on facilitated cesium transport through glass bead columns and modeling. *J. Contam. Hydrol.* 31, 23–56, 1998.
- Parker, J.C. and van Genuchten, M.Th., Determining transport parameters from laboratory and field tracer experiments. *Bull. 84-3, Va Agric. Exp. St.*, Blacksburg, VA, 1984.
- Parkhurst, D.L. and Appelo, C.A.J., User's guide to PHREEQC (version 2) — A computer program for speciation, batch-reaction, one-dimensional transport, and inverse geochemical calculations. Water Resources Investigation, Report 99-4259, Denver, Co, USA, 312 pp, 1999.
- Perrochet, P. and Berod, D., Stability of the standard Crank–Nicolson–Galerkin scheme applied to the diffusion-convection equation: some new insights. *Water Resour. Res.* 29, 3291–3297, 1993.
- Pinder, G.F. and Gray, W.G., *Finite Element Simulation in Surface and Subsurface Hydrology*, Academic Press, New York, 1977.
- Pitzer, K.S., Activity coefficients in electrolyte solutions. In R.M. Pitkowicz, *Activity Coefficients in Electrolyte Solutions*. CRC Press, Inc., Boca Raton, FL, pp. 157–208, 1979.

- Pot, V., Šimůnek, J., Benoit, P., Coquet, Y., Yra, A. and Martínez-Cordón, M.-J., Impact of rainfall intensity on the transport of two herbicides in undisturbed grassed filter strip soil cores. *J. Contam. Hydrol.* 81, 63–88, 2005.
- Pruess, K., Tough2 — A general-purpose numerical simulator for multiphase fluid and heat flow. *Rep. LBL-29400*, Lawrence Berkeley Laboratory, Berkeley, CA, 1991.
- Pruess, K., The TOUGH codes — A family of simulation tools for multiphase flow and transport processes in permeable media. *Vadose Zone J.* 3, 738–746, 2004.
- Pruess, K. and Wang, J.S.Y., Numerical modeling of isothermal and non-isothermal flow in unsaturated fractured rock — a review. In D.D. Evans and T.J. Nicholson (Eds.), *Flow and Transport Through Unsaturated Fractured Rock, Geophysics Monograph 42*, American Geophysical Union, Washington, D.C., pp. 11–22, 1987.
- Rajagopalan, R. and Tien, C., Trajectory analysis of deep-bed filtration with the sphere-in-cell porous media model. *AIChE J.* 22, 523–533, 1976.
- Rathfelder, K.M., Lang, J.R. and Abriola, L.M., A numerical model (MISER) for the simulation of coupled physical, chemical and biological processes in soil vapor extraction and bioventing systems. *J. Contam. Hydrol.* 43, 239–270, 2000.
- Richards, L.A., Capillary conduction of fluid through porous mediums. *Physics* 1, 318–333, 1931.
- Saiers, J.E., Hornberger, G.M. and Harvey, C., Colloidal silicakaolinite in the transport of cesium through laboratory sand columns. *Water Resour. Res.* 32, 33–41, 1996.
- Saiers, J.E. and Hornberger, G.M., The role of colloidal kaolinite in the transport of cesium through laboratory sand columns. *Water Resour. Res.* 32, 33–41, 1996.
- Scanlon, B.R., Christman, M., Reedy, R., Porro, I., Šimůnek, J. and Flerchinger, G., Intercode comparisons for simulating water balance of surficial sediments in semiarid regions. *Water Resour. Res.* 38, 1323, doi:10.1029/2001WR001233, 59.1–59.16, 2002.
- Schaerlaekens, J., Mallants, D., Šimůnek, J. van Genuchten, M.Th. and Feyen, J. Numerical simulation of transport and sequential biodegradation of chlorinated aliphatic hydrocarbons using CHAIN_2D. *J. Hydrologi. Proc.* 13, 2847–2859, 1999.
- Schaap, M.G., Leij, F.J. and van Genuchten, M.Th., Rosetta: A computer program for estimating soil hydraulic parameters with hierarchical pedotransfer functions. *J. Hydrol.* 251, 163–176, 2001.
- Schijven, J.F. and Hassanzadeh, S.M., Removal of viruses by soil passage: Overview of modeling, processes and parameters. *Crit. Rev. Environ. Sci. Technol.* 30, 49–127, 2000.
- Schijven, J. and Šimůnek, J., Kinetic modeling of virus transport at field scale. *J. Contam. Hydrol.* 55, 113–135, 2002.
- Selim, H.M., Davidson, J.M. and Mansell, R.S., Evaluation of a two-site adsorption-desorption model for describing solute transport in soils, paper presented at Proceedings, Summer Computer Simulation Conference, *Nat. Sci. Found.*, Washington, DC, 1976.
- Seuntjens, P., Tirez, K., Šimůnek, J., van Genuchten, M.Th., Cornelis, C. and Geuzens, P., Aging effects on cadmium transport in undisturbed contaminated sandy soil columns. *J. Environ. Qual.* 30, 1040–1050, 2001.
- Shapiro, R.E. and Fried M., Relative release and retentiveness of soil phosphates. *Soil Sci. Soc. Am. Proc.* 23, 195–198, 1959.
- Sibbesen, E., Some new equations to describe phosphate sorption by soils. *J. Soil Sci.* 32, 67–74, 1981.
- Šimůnek, J. and Suarez, D.L., Modeling of carbon dioxide transport and production in soil: 1. Model development. *Water Resour. Res.* 29, 487–497, 1993.
- Šimůnek, J. and Suarez, D.L., Two-dimensional transport model for variably saturated porous media with major ion chemistry. *Water Resour. Res.* 30, 1115–1133, 1994.
- Šimůnek, J. and van Genuchten, M.Th., The CHAIN_2D code for simulating two-dimensional movement of water flow, heat, and multiple solutes in variably-saturated porous media, Version 1.1, *Research Report No 136*, U.S. Salinity Laboratory, USDA, ARS, Riverside, California, 205 pp, 1994.
- Šimůnek, J., Suarez, D.L. and Šejna, M., The UNSATCHEM software package for simulating one-dimensional variably saturated water flow, heat transport, carbon dioxide production and transport,

- and multicomponent solute transport with major ion equilibrium and kinetic chemistry. Version 2.0. Research Report No. 141, U.S. Salinity Laboratory, USDA, ARS, Riverside, California, 186 pp, 1996.
- Šimůnek, J. and Suarez, D.L., Sodic soil reclamation using multicomponent transport modeling. *ASCE J. Irrig. Drain. Eng.* 123, 367–376, 1997.
- Šimůnek, J., Šejna, M. and van Genuchten, M.Th. The HYDRUS-1D software package for simulating the one-dimensional movement of water, heat, and multiple solutes in variably-saturated media. Version 2.0, *IGWMC-TPS-70*, International Ground Water Modeling Center, Colorado School of Mines, Golden, Colorado, 202pp, 1998a.
- Šimůnek, J., Wendroth, O. and van Genuchten, M.Th., A parameter estimation analysis of the evaporation method for determining soil hydraulic properties. *Soil Sci. Soc. Am. J.* 62, 894–905, 1998b.
- Šimůnek, J., van Genuchten, M.Th., Gribb, M.M. and Hopmans, J.W., Parameter estimation of unsaturated soil hydraulic properties from transient flow processes. *Soil Tillage Res.* 47, Special issue “State of the art in soil physics and in soil technology of aridic soils”, 27–36, 1998c.
- Šimůnek, J., van Genuchten, M.Th., Šejna, M., Toride, N. and Leij, F.J., The STANMOD computer software for evaluating solute transport in porous media using analytical solutions of convection-dispersion equation. Versions 1.0 and 2.0, *IGWMC-TPS-71*, International Ground Water Modeling Center, Colorado School of Mines, Golden, Colorado, 32 pp, 1999a.
- Šimůnek, J., Šejna, M. and van Genuchten, M.Th., The HYDRUS-2D software package for simulating two-dimensional movement of water, heat, and multiple solutes in variably saturated media. Version 2.0, *IGWMC-TPS-53*, International Ground Water Modeling Center, Colorado School of Mines, Golden, Colorado, 251pp, 1999b.
- Šimůnek, J., Wendroth, O., Wypler, N. and van Genuchten, M.Th., Nonequilibrium water flow characterized from an upward infiltration experiment. *Eur. J. Soil Sci.* 52, 13–24, 2001.
- Šimůnek, J. and Hopmans, J.W., Parameter optimization and nonlinear fitting. In: J.H. Dane and G.C. Topp, (Eds.) *Methods of Soil Analysis, Part 1, Physical Methods*, Chapter 1.7, 3rd ed., SSSA, Madison, WI, pp. 139–157, 2002.
- Šimůnek, J., Jacques, D., Hopmans, J.W., Inoue, M., Flury, M. and van Genuchten, M.Th., Solute transport during variably-saturated flow — inverse methods. In J.H. Dane and G.C. Topp, (Eds.), *Methods of Soil Analysis, Part 1, Physical Methods*, Chapter 6.6, 3rd ed., SSSA, Madison, WI, pp. 1435–1449, 2002.
- Šimůnek, J. and Valocchi, A.J., Geochemical transport, In J. H. Dane and G.C. Topp, (Eds.), *Methods of Soil Analysis, Part 1, Physical Methods*, Chapter 6.9, 3rd ed., SSSA, Madison, WI, pp. 1511–1536, 2002.
- Šimůnek, J., Jarvis, N.J., van Genuchten, M.Th. and Gärdenäs, A., Nonequilibrium and preferential flow and transport in the vadose zone: review and case study. *J. Hydrol.* 272, 14–35, 2003.
- Šimůnek, J., van Genuchten, M.Th. and Šejna, M., The HYDRUS-1D software package for simulating the one-dimensional movement of water, heat, and multiple solutes in variably-saturated media. Version 3.0, *HYDRUS Software Series 1*, Department of Environmental Sciences, University of California Riverside, Riverside, CA, 270pp, 2005.
- Šimůnek, J., Models of water flow and solute transport in the unsaturated zone. In: M.G. Anderson and J.J. McDonnell (Eds.), *Encyclopedia of Hydrological Sciences*, John Wiley & Sons, Ltd., Chichester, England, 1171–1180, 2005.
- Šimůnek, J., Jacques, D., van Genuchten, M.Th. and Mallants, D., Multicomponent geochemical transport modeling using the HYDRUS computer software packages. *J. Am. Water Resour. Assoc.* 2006.
- Sips, R., On the structure of the catalyst surface, II. *J. Chem. Phys.* 18, 1024–1026, 1950.
- Skaggs, T.H. and Leij, F.J., Solute transport: Theoretical background. In J.H. Dane and G.C. Topp, (Eds.), *Methods of Soil Analysis, Part 1, Physical Methods*, Chapter 6.3, 3rd ed., SSSA, Madison, WI, pp. 1353–1380, 2002.
- Sposito, G. and Barry, D.A., On the Dagan model of solute transport I groundwater: Foundational aspects. *Water Resour. Res.* 23, 1867–1875, 1987.

- Steefel, C.I. and Yabusaki, S.B., OS3D/GIMRT, software for multicomponent–multidimensional reactive transport. User Manual and Programmer's Guide, PNL-11166, Pacific Northwest National Laboratory, Richland, WA, 1996.
- Steefel, C.I. and MacQuarrie, K.T.B., Approaches to modeling of reactive transport in porous media. In P.C. Lichtner, C.I. Steefel and E.H. Oelkers (Eds.), *Reactive Transport in Porous Media. Reviews in Mineralogy*. Chapter 2, Vol. 34, Mineralogical Society of America, pp. 83–129, 1996.
- Thiele-Bruhn, S., Pharmaceutical antibiotic compounds in soils — a review. *J. Plant Nutr. Soil Sci.* 166, 145–167, 2003.
- Tolls, J., Sorption of veterinary pharmaceuticals in soils: A review. *Environ. Sci. Technol.* 35, 3397–3406, 2001.
- Toride, N., Leij, F.J. and van Genuchten, M.Th., A comprehensive set of analytical solutions for nonequilibrium solute transport with first-order decay and zero-order production. *Water Resour. Res.* 29, 2167–2182, 1993.
- Toride, N., Leij, F.J. and van Genuchten, M.Th., The CXTFIT code for estimating transport parameters from laboratory or field tracer experiments. Version 2.0, *Research Report No. 137*, U. S. Salinity Laboratory, USDA, ARS, Riverside, CA, 1995.
- Truesdell, A.H. and Jones, B.F., Wateq, a computer program for calculating chemical equilibria of natural waters. *J. Res. U.S. Geol. Surv.* 2, 233–248, 1974.
- Ungs, M., Wilkinson, K. and Summers, K., Users guide for MYGRT, Version 3.0: Software for simulating migration of organic and inorganic chemicals in groundwater, *EPRI Report TR-111748*, Electric Power Research Institute, Palo Alto, CA, 1998.
- van Dam, J.C., Huygen, J., Wesseling, J.G., Feddes, R.A., Kabat, P., van Valsum, P.E.V. and Groenendijk, P. and van Diepen, C.A., Theory of SWAP, version 2.0. Simulation of water flow, solute transport and plant growth in the Soil–Water–Atmosphere–Plant environment. Dept. Water Resources, WAU, *Report 71*, DLO Winand Staring Centre, Wageningen, Technical Document 45, 1997.
- Vanderborght, J., Kasteel, R., Herbst, M., Javaux, M., Thiery, D., Vanclooster, M., Mouvet, C. and Vereecken, H., A set of analytical benchmarks to test numerical models of flow and transport in soils. *Vadose Zone J.* 4, 206–221, 2005.
- van de Weerd, H., Leijnse, A. and van Riemsdijk, W.H., Transport of reactive colloids and contaminants in groundwater: effect of nonlinear kinetic reactions. *J. Contam. Hydrol.* 32, 313–331, 1998.
- van Genuchten, M.Th., Davidson, J.M. and Wierenga, P.J., An evaluation of kinetic and equilibrium equations for the prediction of pesticide movement through porous media. *Soil Sci. Soc. Am. Proc.* 38, 29–35, 1974.
- van Genuchten, M.Th. and Wierenga, P.J., Mass transfer studies in sorbing porous media, I. Analytical solutions. *Soil Sci. Soc. Am. J.* 40, 473–481, 1976.
- van Genuchten, M.Th., A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. *Soil Sci. Soc. Am. J.* 44, 892–898, 1980a.
- van Genuchten, M.Th., Determining transport parameters from solute displacement experiments. *Research Report No. 118*, U.S. Salinity Laboratory, USDA, ARS, Riverside, CA, 1980b.
- van Genuchten, M.Th., Non-equilibrium transport parameters from miscible displacement experiments. *Research Report No. 119*, U.S. Salinity Laboratory, USDA, ARS, Riverside, CA, 1981.
- van Genuchten, M.Th. and Parker, J.C., Boundary conditions for displacement experiments through short laboratory soil columns. *Soil Sci. Soc. Am. J.* 48, 703–708, 1984.
- van Genuchten, M.Th., Convective-dispersive transport of solutes involved in sequential first-order decay reactions. *Comp. Geosci.* 11, 129–147, 1985.
- van Genuchten, M.Th. and Wagenet, R.J., Two-site/two-region models for pesticide transport and degradation: Theoretical development and analytical solutions. *Soil Sci. Soc. Am. J.* 53, 1303–1310, 1989.
- van Genuchten, M.Th. and J. Šimůnek, Evaluation of pollutant transport in the unsaturated zone, In: P.E. Rijtema and V. Eliáš (Eds.), *Regional Approaches to Water Pollution in the Environment*, NATO ASI Series: 2. Environment, Kluwer, Dordrecht, The Netherlands, pp. 139–172, 1996.

- van Genuchten, M.Th. and Šimůnek, J., Integrated modeling of vadose zone flow and transport processes. Proc. *Unsaturated Zone Modelling: Progress, Challenges and Applications*, Eds. R.A. Feddes, G.H. de Rooij, and J.C. van Dam, Wageningen, The Netherlands, October 3–5, 2004, pp. 37–69, 2004.
- Vinten, A.J.A., Yaron, B. and Nye, P.H., Vertical transport of pesticides into soil when adsorbed on suspended particles. *J. Agric. Food Chem.* 31, 662–664, 1983.
- Von Gunten, H.R., Weber, V.E. and Krahenbuhl, U., The reactor accident at Chernobyl: A possibility to test colloid-controlled transport of radionuclides in a shallow aquifer. *J. Contam. Hydrol.* 2, 237–247, 1998.
- Wagenet, R.J. and Hutson, J.L., LEACHM: Leaching Estimation And CHemistry Model, A process-based model of water and solute movement, transformations, plant uptake and chemical reactions in the unsaturated zone. Continuum 2. Water Resour. Inst., Cornell University, Ithaca, New York, 1987.
- Westall, J.C., Zachary, J.L. and Morel, F.M.M. MINEQL: A computer program for the calculation of chemical equilibrium composition of aqueous systems. Tech. Note 18, 91pp, Dep. of Civ. Eng., Mass. Inst. of Technol., Cambridge, MA, 1976.
- White, N. and Zelazny, L.W., Charge properties in soil colloids. In D.L. Sparks (Ed.) *Soil Physical Chemistry*. CRC Press, Boca Raton, Florida, pp. 39–81, 1986.
- Wheatcraft, S.W. and Cushman, J.H., Hierarchical approaches to transport in heterogeneous porous media. *Rev. Geophys.* 29 (Suppl.), 261–267, 1991.
- Wilson, G.V., Albright, W.H., Gee, G.W., Fayer, M.J. and Ogan, B.D., Alternative cover assessment project, Phase 1 Report. Final Technical Report prepared by Desert Research Institute, NV and Pacific Northwest National Laboratory, WA for U.S. Environmental Protection Agency under Contract No. 44-0000-2038, 202p, 1999.
- Yeh, G.T., Analytical transient one-, two-, and three-dimensional simulation of waste transport in the aquifer system, ORNL-5602, Oak Ridge National Laboratory, Oak Ridge, TN, 1981.
- Yeh, G.T. and Tripathi, V.S., HYDROGEOCHEM: A coupled model of hydrological and geochemical equilibrium of multicomponent systems. Rep. ORNL-6371, Oak Ridge Natl. Lab., Oak Ridge, TN, 1990.
- Yeh, G.T., Salvage, K.M., Gwo, J.P., Zachara, J.M. and Szecsody, J.E., HYDROBIOGEOCHEM: A coupled model of hydrological transport and mixed biochemical kinetic/equilibrium reactions in saturated–unsaturated media. Rep. ORNL/TM-13668, Oak Ridge Natl. Lab., Oak Ridge, TN, 1998.
- Zhang, P., Šimůnek, J. and Bowman, R.S., Nonideal transport of solute and colloidal tracers through reactive zeolite/iron pellets. *Water Resour. Res.* 40, doi:10.1029/2003WR002445, 2004.
- Zienkiewicz, O.C., *The Finite Element Method*, 3rd ed., McGraw-Hill, London, 1977.
- Zyvoloski, G.A., Robinson, B.A., Dash, Z.V. and Trease, L.L., Summary of the models and methods for the FEHM application — a finite element heat- and mass-transfer code. Los Alamos National Laboratory Rept. LA-13307-MS, NM, 1997.

23

Groundwater Modeling

23.1	Introduction.....	23-1
23.2	Models	23-2
23.3	Flow and Transport Processes.....	23-3
23.4	Governing Equations	23-3
	Groundwater-Flow Equation • Seepage Velocity •	
	Solute-Transport Equation	
23.5	Numerical Methods to Solve Equations.....	23-7
	Basics of Finite-Difference Methods • Basics of	
	Finite-Element Methods • Basics of	
	Method-of-Characteristics Methods • Matrix Solution	
	Techniques • Boundary and Initial Conditions • Generic	
	Model Verification	
23.6	Model Design, Development, and Application	23-19
	Grid Design • Model Calibration • Model Error • Mass	
	Balance • Sensitivity Tests • Calibration Criteria •	
	Predictions and Postaudits • Model Validation	
23.7	Overview of Representative Generic Models.....	23-28
	MODFLOW-2000 • SUTRA	
23.8	Case Histories	23-31
	Flow and Transport in a Shallow Unconfined Aquifer •	
	Delineation of Contributing Areas to Water-Supply Wells,	
	Rhode Island • Seawater Intrusion in a Coastal Aquifer	
23.9	Available Groundwater Software.....	23-44
	Glossary.....	23-45
	Acknowledgments.....	23-46
	References	23-46
	Further Information	23-50

Leonard F. Konikow,
Thomas E. Reilly,
Paul M. Barlow, and
Clifford I. Voss
U.S. Geological Survey

23.1 Introduction

Effective management of groundwater requires the ability to predict subsurface flow and transport of solutes, and the response of fluid and solute flux to changes in natural or human-induced stresses. One popular type of tool that has been evolving since the mid-1960s is the deterministic, distributed-parameter, computer simulation model for analyzing flow and solute-transport in groundwater systems.

The development of the computer simulation model has paralleled the development and increasing availability of faster, more capable, yet less expensive computer systems.

The purpose of this chapter is to review the state of the art in deterministic modeling of groundwater flow and transport processes. This chapter is aimed at practitioners and is intended to describe the types of models that are available and how to apply them to complex field problems. Our discussions of transport modeling are focused on nonreactive dissolved constituents, and we do not address multicomponent reactive transport or multiphase flow phenomena, such as the transport of nonaqueous phase liquids (NAPL). The chapter discusses the philosophy and theoretical basis of deterministic modeling, the advantages and limitations of models, the use and misuse of models, how to select a model, and how to calibrate and evaluate a model. However, as this chapter is only a review, it cannot offer comprehensive and in-depth coverage of this complex topic; instead, it guides the reader to references that provide more details.

23.2 Models

The word **model** has so many definitions and is so overused that it is sometimes difficult to discern its meaning (Konikow and Bredehoeft 1992). A model is perhaps most simply defined as an approximate representation of a real system or process. A **conceptual model** is a hypothesis for how a system or process operates. This hypothesis can be expressed quantitatively as a mathematical model. **Mathematical models** are abstractions that represent processes as equations, physical properties as constants or coefficients in the equations, and measures of state or potential in the system as variables.

Most groundwater models in use today are deterministic mathematical models. **Deterministic models** are based on conservation of mass, momentum, and energy and describe cause and effect relations. The underlying assumption is that given a high degree of understanding of the processes by which stresses on a system produce subsequent responses, the response of the system to any set of stresses can be predetermined, even if the magnitude of the new stresses falls outside the historically observed range.

Deterministic groundwater models generally require the solution of partial differential equations. Exact solutions can often be obtained analytically, but **analytical models** require that the parameters and boundaries be highly idealized. Some deterministic models treat the properties of porous media as lumped parameters (essentially, as a black box), but this precludes explicit representation of heterogeneous hydraulic properties in the model. Heterogeneity, or variability in aquifer properties, is characteristic of all geologic systems and is recognized as important in affecting groundwater flow and solute transport. Thus, it is often preferable to apply distributed-parameter models, which allow the representation of more realistic distributions of system properties. Numerical methods yield approximate solutions to the governing equation (or equations) through the **discretization** of space and time. Within the discretized problem domain, the varying internal properties, boundaries, and stresses of the hydrogeologic system are approximated. Deterministic, distributed-parameter, **numerical models** can relax the rigid idealized conditions of analytical models or lumped-parameter models, and they can, therefore, be more realistic and flexible for simulating field conditions (if applied properly).

The number and types of equations to be solved are determined by the dominant governing processes. The coefficients of the equations are parameters that are measures of the properties, boundaries, and stresses of the system; the dependent variables of the equations are measures of the state of the system and are mathematically determined by the solution of the equations. When a numerical algorithm is implemented in a computer code to solve one or more partial differential equations, the resulting computer code can be considered a **generic model**. When the grid dimensions, boundary conditions, and other parameters (such as hydraulic conductivity and storativity) are specified in an application of a generic model to represent a particular geographical area, the resulting computer program is a **site-specific model**. The capacity of generic models to solve the governing equations accurately is typically demonstrated by example applications to simplified problems. This capacity does not guarantee a similar level of accuracy when the model is applied to a complex field problem.

If the user of a model is unaware of or ignores the details of the numerical method, including the derivative approximations, the scale of discretization, and the matrix solution techniques, appreciable errors can be introduced and remain undetected. For example, if the groundwater-flow equation is solved iteratively, but the convergence criterion is too coarse, then the numerical solution may converge to an inaccurate solution. The inaccuracy of the solution may or may not be reflected in the mass-balance error. The mass-balance error itself may not be readily observed by inexperienced model users. Moreover, the potential for unrecognized errors in numerical models is increasing because “user-friendly” commercially available graphical user interfaces (GUIs) make it easier for models to be used (and to be misused). These interfaces effectively place more “distance” between the modeler and the numerical method that lies at the core of the model. When developing a complex three-dimensional (3D) groundwater model, however, a GUI may be a necessity.

23.3 Flow and Transport Processes

The process of groundwater flow is generally assumed to be governed by the relations expressed in Darcy’s law (see Chapter 3) and the conservation of mass. However, Darcy’s law does have limits on its range of applicability, and these limits must be evaluated in any model application.

The purpose of a model that simulates solute transport in groundwater is to compute the concentration of a dissolved chemical species in a groundwater system at any specified time and place. The theoretical basis for the equation describing solute transport has been well documented in the literature (e.g., Bear 1979; Domenico and Schwartz 1997). Reilly et al. (1987) provide a conceptual framework for analyzing and modeling physical solute-transport processes in groundwater. Changes in chemical concentration occur within a dynamic groundwater system primarily because of four distinct processes: (1) advective transport, in which dissolved chemicals are moving with the flowing groundwater; (2) hydrodynamic dispersion, in which molecular and ionic diffusion and small-scale variations in the flow velocity through the porous media cause the paths of dissolved molecules and ions to diverge or spread from the average advective flow path; (3) fluid sources, where water of one composition is introduced into and mixed with water of a different composition; and (4) reactions, in which some amount of a particular dissolved chemical species may be added to or removed from the groundwater as a result of chemical, biological, and physical reactions in the water or between the water and the solid aquifer materials or other separate liquid phases.

The subsurface environment constitutes a complex, 3D, heterogeneous hydrogeologic setting. This variability strongly affects groundwater flow and transport, and such settings can be described accurately only through careful hydrogeologic practice in the field. Regardless of how much data are collected, uncertainty always remains about the properties and boundaries of the groundwater system of interest. **Stochastic** approaches have resulted in many significant advances in characterizing subsurface heterogeneity and dealing with uncertainty (see Gelhar 1993). Parameters often are treated in stochastic modeling as random variables and deterministically solve the governing equations for many realizations.

23.4 Governing Equations

The mathematical equations that describe groundwater flow (see Chapter 4) and transport processes (see Chapter 18 and Chapter 19) may be developed from the fundamental principle of conservation of mass of fluid or of solute. Given a representative elementary volume (REV) of porous medium, a general equation for conservation of mass for the volume may be expressed as

$$\begin{aligned} \text{rate of mass inflow} - \text{rate of mass outflow} + \text{rate of mass production/consumption} \\ = \text{rate of mass accumulation} \end{aligned} \quad (23.1)$$

This statement of conservation of mass (or continuity equation) may be combined with a mathematical expression of the relevant process to obtain a differential equation that describes flow or transport.

23.4.1 Groundwater-Flow Equation

The rate of flow of water through a porous media is related to the properties of the water, the properties of the porous media, and the gradient of the hydraulic head, which can be written using Darcy's law as

$$q_i = -K_{ij} \frac{\partial h}{\partial x_j} \quad (23.2)$$

where q_i is the specific discharge, LT^{-1} ; K_{ij} is the hydraulic conductivity of the porous medium (a second-order tensor), LT^{-1} ; h is the hydraulic head, L; and x_j are the Cartesian coordinates, L.

A general form of the equation describing the transient flow of a slightly compressible fluid in a nonhomogeneous anisotropic aquifer may be derived by combining Darcy's law with the continuity equation. A general groundwater-flow equation may be written in Cartesian tensor notation as

$$\frac{\partial}{\partial x_i} \left(K_{ij} \frac{\partial h}{\partial x_j} \right) = S_s \frac{\partial h}{\partial t} + W^* \quad (23.3)$$

where S_s is the specific storage, L^{-1} ; t is time, T; and W^* the volumetric flux per unit volume (positive for outflow and negative for inflow), T^{-1} . The summation convention of Cartesian tensor analysis is implied in Equation 23.2 and Equation 23.3. Equation 23.3 can generally be applied if isothermal conditions prevail, the porous medium only deforms vertically, the volume of individual grains remains constant during deformation, Darcy's law applies (and gradients of hydraulic head are the only driving force), and fluid properties (density and viscosity) are homogeneous and constant. Aquifer properties can vary spatially and fluid stresses (W^*) can vary in space and time.

If the aquifer is relatively thin compared to its lateral extent, it may be appropriate to assume that groundwater flow is areally two dimensional. This assumption allows the 3D flow equation to be reduced to the case of two-dimensional (2D) areal flow, for which various additional simplifications are possible. The advantages of reducing the dimensionality of the equation include less stringent data requirements, smaller computer memory requirements, and shorter computer execution times to achieve numerical solutions.

An expression similar to Equation 23.3 may be derived for the 2D areal flow of a homogeneous fluid in a confined aquifer and written as

$$\frac{\partial}{\partial x_i} \left(T_{ij} \frac{\partial h}{\partial x_j} \right) = S \frac{\partial h}{\partial t} + W \quad (23.4)$$

where T_{ij} is the transmissivity, L^2T^{-1} ; and $T_{ij} = K_{ij}b$; b is the saturated thickness of the aquifer, L; S is the storage coefficient (dimensionless); and $W = W^*b$ is the volume flux per unit area, LT^{-1} .

When Equation 23.4 is applied to an unconfined (water-table) aquifer, it must be assumed that flow is horizontal and equipotential lines are vertical, that the horizontal hydraulic gradient equals the slope of the water table, and that the storage coefficient is equal to the specific yield (S_y) (Anderson and Woessner 1992). Note that in an unconfined system, the saturated thickness changes as the water-table elevation (or head) changes. Thus, the transmissivity also can change over space and time (i.e., $T_{ij} = K_{ij}b$, where $b(x, y, t) = h - h_b$, and h_b is the elevation of the bottom of the aquifer).

The cross-product terms of the hydraulic conductivity tensor drop out when the coordinate axes are aligned with the principal axes of the tensor; that is, $K_{ij} = 0$ when $i \neq j$. Therefore, the only hydraulic conductivity terms with possible nonzero values are K_{xx} and K_{yy} . Under this assumption, Equation 23.4

may be simplified to

$$\frac{\partial}{\partial x} \left(T_{xx} \frac{\partial h}{\partial x} \right) + \frac{\partial}{\partial y} \left(T_{yy} \frac{\partial h}{\partial y} \right) = S \frac{\partial h}{\partial t} + W \quad (23.5)$$

for 2D flow.

In some field situations, fluid properties, such as density and viscosity, may vary significantly in space or time. This may occur where water temperature or dissolved-solids concentration changes significantly. When the fluid properties are heterogeneous and (or) transient, the relations among water levels, hydraulic heads, fluid pressures, and flow velocities are neither simple nor straightforward. In such cases, it is impossible to solve the groundwater-flow equation in terms of the fluid potential (hydraulic head), and the flow equation is written and solved in terms of fluid pressures, fluid densities, and the intrinsic permeability of the porous media as

$$(\rho S_{0p}) \frac{\partial p}{\partial t} + \left(\varepsilon \frac{\partial \rho}{\partial \omega} \right) \frac{\partial \omega}{\partial t} - \frac{\partial}{\partial x_i} \left[\left(\frac{k_{ij} \rho}{\mu} \right) \left(\frac{\partial p}{\partial x_j} + \rho g \frac{\partial z}{\partial x_j} \right) \right] = Q_p \quad (23.6)$$

where S_{0p} is the specific pressure storativity [$\text{ML}^{-1}\text{T}^{-2}]^{-1}$; ρ is the fluid density [ML^{-3}]; p is the fluid pressure [$\text{ML}^{-1}\text{T}^{-2}$]; Q_p is a fluid mass source [$\text{ML}^{-3}\text{T}^{-1}$]; k_{ij} is the intrinsic permeability of the porous medium (a second-order tensor) [L^2]; μ is the dynamic fluid viscosity [$\text{ML}^{-1}\text{T}^{-1}$]; g is the gravitational acceleration [LT^{-2}]; and ω is the solute mass fraction (mass of solute/mass of fluid) [dimensionless]. The specific pressure storativity is related to the specific storage coefficient as $S_s = \rho |g| S_{0p}$.

23.4.2 Seepage Velocity

The migration and mixing of chemicals dissolved in groundwater is affected by the velocity of the flowing groundwater. The specific discharge calculated from Equation 23.2 is sometimes called the Darcy velocity. However, this nomenclature can be misleading because q_i does not actually represent the speed of water movement. Rather, q_i represents a volumetric flux per unit cross-sectional area. To calculate the mean seepage velocity, one must account for the cross-sectional area through which flow is occurring by dividing the specific discharge by the effective porosity of the porous medium, ε , resulting in

$$V_i = \frac{q_i}{\varepsilon} = -\frac{K_{ij}}{\varepsilon} \frac{\partial h}{\partial x_j} \quad (23.7)$$

where V_i is the seepage velocity (also commonly called average linear velocity or average interstitial velocity), LT^{-1} . For variable-density flow, the fluid velocity is given by

$$V_i = -\left(\frac{k_{ij} \varepsilon \rho}{\mu} \right) \left(\frac{\partial p}{\partial x_j} + \rho g \frac{\partial z}{\partial x_j} \right) \quad (23.8)$$

23.4.3 Solute-Transport Equation

An equation describing the transport and dispersion of a dissolved chemical in flowing groundwater may be derived from the principle of conservation of mass by considering all fluxes into and out of a REV, as described by Bear (1979, p. 29). A generalized form of the solute-transport equation is presented by Grove (1976), in which terms are incorporated to represent chemical reactions and solute concentration both in the pore fluid and on the solid surface, as

$$\frac{\partial (\varepsilon C)}{\partial t} = \frac{\partial}{\partial x_i} \left(\varepsilon D_{ij} \frac{\partial C}{\partial x_j} \right) - \frac{\partial}{\partial x_i} (\varepsilon C V_i) - C^* W^* + \text{CHEM} \quad (23.9)$$

where CHEM equals one or more of the following:

$$-\rho_b \frac{\partial \bar{C}}{\partial t} \quad \text{for linear equilibrium controlled sorption or ion-exchange reactions,}$$

$$\sum_{k=1}^s R_k \quad \text{for } s \text{ chemical rate-controlled reactions, and (or)}$$

$$-\lambda(\varepsilon C + \rho_b \bar{C}) \quad \text{for decay,}$$

and where C is the volumetric concentration (mass of solute per unit volume of fluid, ML^{-3}), D_{ij} is the coefficient of hydrodynamic dispersion (a second-order tensor), L^2T^{-1} ; C^* is the concentration of the solute in the source or sink fluid; \bar{C} is the concentration of the species adsorbed on the solid (mass of solute/mass of solid); ρ_b is the bulk density of the sediment, ML^{-3} ; R_k is the rate of production of the solute in reaction k , $\text{ML}^{-3}\text{T}^{-1}$; and λ is the decay constant (equal to $\ln 2/\text{half life}$), T^{-1} (Grove 1976). For cases of variable-density fluids without reactions, the solute mass balance may be expressed as

$$(\varepsilon\rho) \frac{\partial \omega}{\partial t} = \frac{\partial}{\partial x_i} \left(\varepsilon\rho D_{ij} \frac{\partial \omega}{\partial x_j} \right) - \varepsilon\rho V_i \frac{\partial \omega}{\partial x_i} + Q_p(\omega^* - \omega) \quad (23.10)$$

The first terms on the right sides of Equation 23.9 and Equation 23.10 represent the change in concentration because of hydrodynamic dispersion. This expression is analogous to Fick's law describing diffusive flux. This Fickian model assumes that spreading is driven by the concentration gradient and that the dispersive flux occurs in a direction from higher concentrations toward lower concentrations. However, this assumption is not always consistent with field observations and is the subject of much ongoing research and field study (see, e.g., Gelhar et al. 1992; Carrera et al. 2005). The coefficient of hydrodynamic dispersion is defined as the sum of mechanical dispersion and molecular diffusion (Bear 1979), commonly expressed as

$$D_{ij} = \alpha_{ijmn} \frac{V_m V_n}{|V|} + D^* \quad i, j, m, n = 1, 2, 3 \quad (23.11)$$

where α_{ijmn} is the dispersivity of the porous medium (a fourth-order tensor), L ; V_m and V_n are the components of the flow velocity of the fluid in the m and n directions, respectively, LT^{-1} ; D^* is the effective coefficient of molecular diffusion, L^2T^{-1} ; and $|V|$ is the magnitude of the velocity vector, LT^{-1} , defined as $|V| = \sqrt{V_x^2 + V_y^2 + V_z^2}$ (Bear 1979; Domenico and Schwartz 1997). Mechanical dispersion is a function both of the intrinsic properties of the porous medium (such as heterogeneities in hydraulic conductivity and porosity) and of the fluid flow. Molecular diffusion in a porous medium will differ from that in free water because of the effects of tortuous paths of fluid connectivity in porous media.

The dispersivity of an isotropic porous medium can be defined by two constants. These constants are the longitudinal dispersivity of the medium, α_L , and the transverse dispersivity of the medium, α_T . These constants are related to the longitudinal and transverse dispersion coefficients by $D_L = \alpha_L |V|$ and $D_T = \alpha_T |V|$. Most documented applications of transport models to groundwater contamination problems have been based on this conventional formulation, even for cases in which the hydraulic conductivity is assumed to be anisotropic, despite the conceptual inconsistency. However, some models (e.g., Voss and Provost 2002) incorporate an additional level of sophistication by allowing α_L and (or) α_T to vary with direction. Flow-direction-dependent dispersivity is indeed a useful generalization of the classical dispersion paradigm. For example, it may be expected that flow parallel to layering in a layered medium and flow perpendicular to the layering do not necessarily have the same longitudinal dispersivity. For seawater intrusion, as considered in cross section, solutes may travel several kilometers inland from the

sea within a given aquifer layer with an effective longitudinal dispersivity of hundreds of meters. After transport inland, the solutes may be induced to travel upward to an overlying aquifer layer when shallow wells are pumped; this transport occurs over the distance of aquifer thickness of a few tens of meters. A dispersivity value of hundreds of meters is not appropriate or meaningful for quantifying transport with flow over a vertical distance of tens of meters. Rather, a dispersivity of a few meters is more reasonable. Thus, for simulating this hydrogeologic system using flow-direction-dependent dispersivity, longitudinal dispersivity for flow in the horizontal direction would be specified to be about 100 m and for flow in the vertical direction would be specified to be about 1 m.

Although conventional theory holds that α_L is generally an intrinsic property of the aquifer, it is found in practice to be dependent on and proportional to the scale of the measurement. Most reported values of α_L fall in a range from 0.01 to 1.0 times the scale of the measurement, although the ratio of α_L to scale of measurement tends to decrease at larger scales (see Anderson 1984; Gelhar et al. 1992). Field-scale dispersion (commonly called macrodispersion) results from large-scale spatial variations in hydraulic properties. Consequently, the use of relatively large values of dispersivity together with uniform hydraulic properties (K_{ij} and ϵ) is inappropriate for describing transport in geological systems (Smith and Schwartz 1980). Part of the scale dependence of dispersivity may be explained as an artifact of the models used, in that a scaling up of dispersivity will occur whenever an $(n - 1)$ -dimensional model is calibrated or used to describe an n -dimensional system (Domenico and Robbins 1984). Furthermore, if mean values are applied in a model of a system with variable hydraulic conductivity and thereby does not explicitly represent the variability, the model calibration will likely yield values for the dispersivity coefficients that are larger than would be measured locally in the field area. Similarly, representing a transient flow field by a mean steady-state flow field, as is commonly done, inherently ignores some of the variability in velocity and must be compensated for by using increased values of dispersivity (primarily transverse dispersivity) (Goode and Konikow 1990). Overall, the more accurately a model can represent or simulate the true velocity distribution in space and time, the less of a problem will be the uncertainty concerning representation of dispersion processes.

The mathematical solute-transport model requires at least two partial differential equations. One is the equation of flow, from which groundwater-flow velocities are obtained, and the second is the solute-transport equation, whose solution gives the chemical concentration in groundwater. If the properties of the water are affected appreciably by changes in solute concentration, as in a seawater intrusion problem in a coastal aquifer, then the flow and transport equations should be solved simultaneously (or at least iteratively). For such a variable-density problem, the flow and transport equations must be linked by a functional relation between concentration and fluid density. If the transport problem is such that the properties of the water remain constant, then the flow and transport equations can be decoupled and solved sequentially, which is simpler numerically.

23.5 Numerical Methods to Solve Equations

The partial differential equations describing groundwater flow and transport can be solved mathematically using either analytical or numerical solutions. The advantages of an analytical solution, when it is possible to apply one, are that it usually provides an exact solution to the governing equation and is often relatively simple and efficient to use. Many analytical solutions have been developed for the flow equation; however, most applications have been limited to well hydraulics problems involving radial symmetry and to 1D and 2D problems involving stream-aquifer interactions. The familiar Theis type curve represents the solution of one such analytical model. Analytical solutions are also available to solve the solute-transport equation (e.g., Bear 1979; Javandel et al. 1984; Wexler 1992). In general, obtaining the exact analytical solution to the partial differential equation requires that the properties and boundaries of the flow system be highly and perhaps unrealistically idealized. For modeling most field problems, the mathematical benefits of obtaining an exact analytical solution can be outweighed by the errors introduced by the simplifying assumptions required to apply the analytical approach to the field setting.

For problems where simplified analytical models cannot describe the physics of the situation, the partial differential equations of groundwater flow and transport can be approximated numerically. In so doing, the continuous variables are replaced with discrete variables that are defined throughout a spatial grid. Thus, the continuous differential equation, which defines hydraulic head or solute concentration everywhere in the groundwater system, is replaced by a finite number of algebraic equations that defines the hydraulic head or concentration at specific points. This system of algebraic equations generally is solved using matrix techniques. This approach constitutes a numerical model.

Two major classes of numerical methods have come to be the most widely used and accepted for solving the groundwater-flow equation. These classes are finite-difference and finite-element methods. Each of these two major classes of numerical methods includes a variety of subclasses and implementation alternatives. Comprehensive treatments of the application of these numerical methods to groundwater problems are presented by Remson et al. (1971) and Wang and Anderson (1982). Both these numerical approaches require that the area of interest be subdivided into a grid (or mesh) of smaller sub-areas (cells or elements) that are associated with nodal points (either at the centers or peripheries of the sub-areas).

In addition to finite-difference and finite-element methods, boundary integral equation methods and analytical element methods can also be applied to solve the flow equation (e.g., Haitjema 1995). The main advantage of these methods is that, for homogeneous regions, they can provide precise solutions without discretization. Thus, if the heterogeneity of a groundwater system can be adequately represented by using only a few very large elements, the methods can be efficient in terms of computer time. If heterogeneities are such that a large number of elements are required to describe them adequately, then finite-difference or finite-element methods may be preferable.

Finite-difference methods approximate the first derivatives in the partial differential equations as difference quotients (the differences between values of the independent variable at adjacent nodes with respect to the distance between the nodes, and at two successive time levels with respect to the duration of the time-step increment). Assumed functions of the dependent variable and parameters are used in the finite-element methods to evaluate equivalent integral formulations of the partial differential equations. Huyakorn and Pinder (1983) present a comprehensive analysis of the application of finite-element methods to groundwater problems. In both numerical approaches, the discretization of the space and time dimensions allows the continuous boundary-value problem for the solution of the partial differential equation to be reduced to the simultaneous solution of a set of algebraic equations. These equations can then be solved using either iterative or direct matrix methods.

Each method described here has advantages and disadvantages, but there are very few groundwater problems for which either is clearly superior. In general, the finite-difference methods are conceptually and mathematically simpler, and are easier to program. The finite-difference equations typically are derived for a relatively simple, rectangular grid, which also eases data entry. Finite-element methods generally require the use of more sophisticated mathematics but, for some problems, may be more accurate numerically than standard finite-difference methods. A major advantage of the finite-element methods is the flexibility of the finite-element grid, which allows a close spatial approximation of irregular boundaries of the aquifer and (or) of parameter zones within the aquifer when they are considered. However, the construction and specification of an input data set are much more difficult for an irregular finite-element grid than for a regular rectangular finite-difference grid. Thus, the use of a model preprocessor, which includes a mesh generator and a scheme to number the nodes and elements of the mesh and to specify the spatial coordinates of each node, is recommended. A hypothetical aquifer system that has been discretized using finite-difference and finite-element grids is illustrated in Figure 23.1. Conceptually how the finite-difference and finite-element grids can be adjusted to use a finer mesh spacing in selected areas of interest are illustrated in Figure 23.1b and Figure 23.1c. The rectangular finite-difference grid approximates the aquifer boundaries in a step-wise manner, resulting in some nodes and cells that lie outside of the aquifer. In contrast, the finite-element mesh can closely follow the outer boundary of the aquifer using a minimal number of nodes.

The solute-transport equation is more difficult to solve numerically than the groundwater-flow equation, largely because the mathematical properties of the transport equation vary depending upon

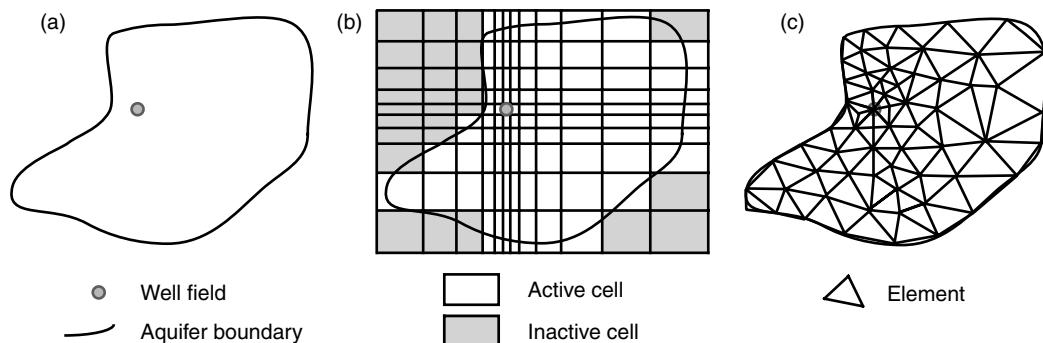


FIGURE 23.1 Application of a numerical model to a simple hypothetical problem, showing (a) an irregularly bounded aquifer discretized using (b) a finite-difference grid and (c) a finite-element mesh. (From Konikow, L. F. (1996) In *Manual on Mathematical Models in Isotope Hydrogeology*, International Atomic Energy Agency, Vienna, Austria.)

which terms in the equation are dominant in a particular situation. When solute transport is dominated by advective transport, as is common in many field problems, then Equation 23.9 or Equation 23.10 approximates a hyperbolic type of equation (similar to equations describing the propagation of a wave or of a shock front). However, if transport is dominated by dispersive fluxes, such as might occur where fluid velocities are relatively low and aquifer dispersivities are relatively high, then Equation 23.9 or Equation 23.10 becomes more parabolic in nature (similar to the transient groundwater-flow equation).

The numerical methods that work best for parabolic partial differential equations are not best for solving hyperbolic equations, and vice versa. Thus, no single numerical method or simulation model will be ideal for the entire spectrum of groundwater-transport problems likely to be encountered in the field. Further compounding this difficulty is that in the field, the seepage velocity of groundwater can be highly variable because of heterogeneous aquifer properties and (or) boundary conditions. Thus, in low permeability zones or near stagnation points, the velocity may be close to zero and the transport processes will be dominated by dispersion processes; in high permeability zones or near stress points (such as pumping wells), the velocity may be several meters per day and the transport processes will be advection-dominated. In other words, for the same system, the governing equation may be more hyperbolic in one area (or at one time) and more parabolic in another. Therefore, no numerical method will be ideal or optimal over the entire domain of the problem, and appreciable numerical errors may be introduced somewhere in the solution. The transport modeling effort must recognize this inherent difficulty and strive to minimize and control the numerical errors.

Additional complications arise when the solutes of interest are reactive. The reaction terms included in Equation 23.9 are mathematically simple ones that may not necessarily represent the true complexities of many reactions. Also, particularly difficult numerical problems arise when reaction terms are highly nonlinear, or if the concentration of the solute of interest is strongly dependent on the concentration of other chemical constituents. In reality, the transfer of solute between the liquid and solid phases may not be linear or equilibrium controlled. For field problems in which reactions appreciably affect solute concentrations, simulation accuracy is less limited by mathematical constraints than by data constraints. That is, the types and rates of reactions for the specific solutes and minerals in the particular groundwater system of interest are rarely known and often highly variable, and require an extensive amount of data to assess accurately.

Finite-difference and finite-element methods also can be applied to solve the transport equation, particularly when dispersive transport is large compared to advective transport. However, numerical errors, such as numerical dispersion and oscillations, may be important for some problems. The numerical errors can generally be reduced by using a finer discretization (either shorter time steps or finer spatial grid). An example of a documented 3D, transient, finite-difference model that simultaneously solves the

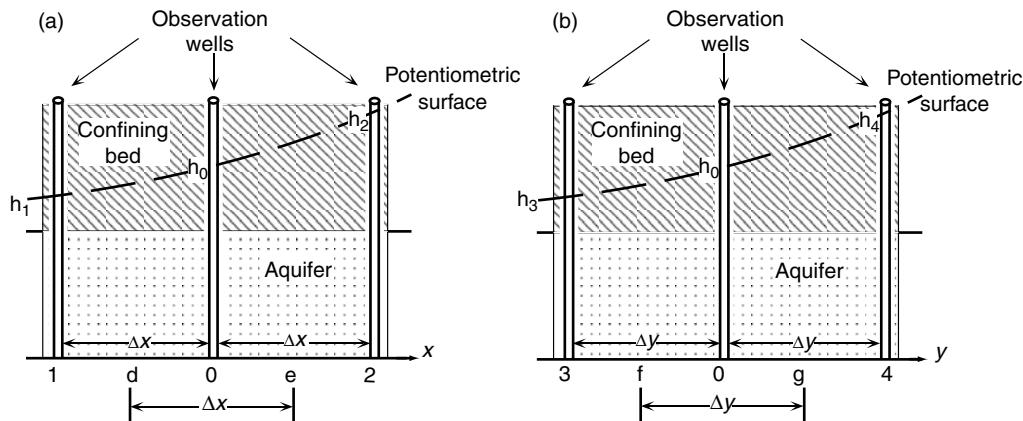


FIGURE 23.2 Schematic cross section through confined aquifer to illustrate numerical approximation to derivatives of head, (a) $\partial h/\partial x$ and (b) $\partial h/\partial y$. (Adapted from Bennett, G. D. (1976) Introduction to ground-water hydraulics: A programmed text for self-instruction, *Techniques of Water-Res. Invests.* of the U.S. Geol. Survey, Book 3, Ch. B2.)

fluid pressure, energy-transport, and solute-transport equations for nonhomogeneous miscible fluids is HST3D (Kipp 1987). An example of a 3D finite-element transport model is SUTRA, documented by Voss and Provost (2002).

Although finite-difference and finite-element models are commonly applied to transport problems, other types of numerical methods also have been used, including the method of characteristics, random walk, Eulerian–Lagrangian methods, and adaptive grid methods. All of these methods have the capacity to track sharp fronts accurately with a minimum of numerical dispersion. Documented models based on variants of these approaches include Konikow et al. (1996), Prickett et al. (1981), Zheng and Wang (1999), and Heberton et al. (2000). None of the standard numerical methods is ideal for a wide range of transport problems and conditions. Thus, there is still much research being done to develop better mixed or adaptive methods that aim to minimize numerical errors and combine the best features of alternative standard numerical approaches.

23.5.1 Basics of Finite-Difference Methods

The partial differential equations describing the flow and transport processes in groundwater include terms representing derivatives of continuous variables in space and time. Finite-difference methods are based on the approximation of these derivatives (or slopes of curves) by discrete linear changes over discrete intervals of space or time. If the intervals are sufficiently small, then all of the linear increments will represent a good approximation of the true curvilinear surface or hydrograph.

If we consider three observation wells spaced an equal distance apart in a confined aquifer, as illustrated in Figure 23.2a, Bennett (1976) shows that a reasonable approximation for the derivative of head, $\partial h/\partial x$, at a point (d) midway between wells 1 and 0 is

$$\left(\frac{\partial h}{\partial x}\right)_d \approx \frac{h_0 - h_1}{\Delta x} \quad (23.12)$$

Similarly, a reasonable approximation for the second derivative, $\partial^2 h/\partial x^2$, at point 0 (the location of the center well) can be given as

$$\left(\frac{\partial^2 h}{\partial x^2}\right)_0 \approx \frac{(\partial h/\partial x)_e - (\partial h/\partial x)_d}{\Delta x} \approx \frac{((h_2 - h_0)/\Delta x) - ((h_0 - h_1)/\Delta x)}{\Delta x} = \frac{h_1 + h_2 - 2h_0}{(\Delta x)^2} \quad (23.13)$$

If we also consider wells 3 and 4 shown in Figure 23.2b, located on a line parallel to the y -axis, we can similarly approximate $\partial^2 h / \partial y^2$ at point 0 (the same point 0 as in Figure 23.2a) as (Bennett 1976):

$$\left(\frac{\partial^2 h}{\partial y^2} \right)_0 \approx \frac{h_3 + h_4 - 2h_0}{(\Delta y)^2} \quad (23.14)$$

If we consider a case of 2D horizontal flow only for the uniform spacing of wells in Figure 23.2b (i.e., $\Delta x = \Delta y = a$), then we can develop the following approximation:

$$\frac{\partial^2 h}{\partial x^2} + \frac{\partial^2 h}{\partial y^2} \approx \frac{h_1 + h_2 + h_3 + h_4 - 4h_0}{a^2} \quad (23.15)$$

These approximations can also be obtained through the use of Taylor series expansions. A certain error is involved in approximating the derivatives by finite-differences, but this error will generally decrease as a (or Δx and Δy) is given smaller and smaller values. This error is called a “truncation error” because the replacement of a derivative by a difference quotient is equivalent to using a truncated Taylor series, so that the exact solution of a difference equation differs from the solution of the corresponding differential equation (Peaceman 1977). Also, it may not be possible to achieve an “exact” solution of the difference equation because of precision limits in storing numbers in a digital computer. In solving a large set of difference equations, many arithmetic operations are performed, and round-off errors may sometimes accumulate.

Next, consider the construction of a rectangular finite-difference grid. Two possible modes of grid construction are illustrated in two dimensions in Figure 23.3a and Figure 23.3b. In Figure 23.3a, the calculation points (or nodes) are located at the centers of the blocks (or cells) formed by the grid lines. This type of grid is commonly called a block-centered grid. In the second grid type (Figure 23.3b), the nodes are considered to be located at the intersections of the grid lines. This type has been variously called a point-, node-, mesh-, or lattice-centered grid. Although there is no overall inherent advantage of one type over the other, there will be some operational differences between the two approaches in the treatment of boundaries and in areas of influence around nodes. Most, but not all, finite-difference groundwater models are based on the use of block-centered grids. Double indexing is normally used to identify functions and variables within the 2D region. For example, $h_{i,j}$ is the head at node i, j , where i and j are the row and column locations in the finite-difference grid. This procedure is easily extended to three dimensions, as illustrated in Figure 23.3c. Here the vertical dimension (or z -direction) is indexed by the subscript k and $h_{i,j,k}$ would represent the head at node i, j, k .

We must also consider the discretization of time, which may be viewed as another dimension, and, hence, represented by another index. If we consider a representative segment of a hydrograph (see Figure 23.4), in which head is plotted against time for a transient flow system, n is the index or subscript used to denote the time at which a given head value is observed. The slope of the hydrograph at any point is the derivative of head with respect to time, and it can be approximated as $\partial h / \partial t \approx \Delta h / \Delta t$. In terms of the

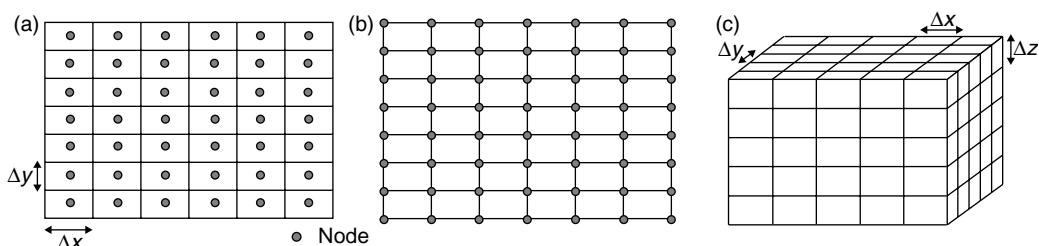


FIGURE 23.3 Examples of finite-difference grids: (a) 2D block-centered grid, (b) 2D node-centered grid, and (c) 3D block-centered grid. (Modified from Konikow, L. F. (1996) In *Manual on Mathematical Models in Isotope Hydrogeology*, International Atomic Energy Agency, Vienna, Austria.)

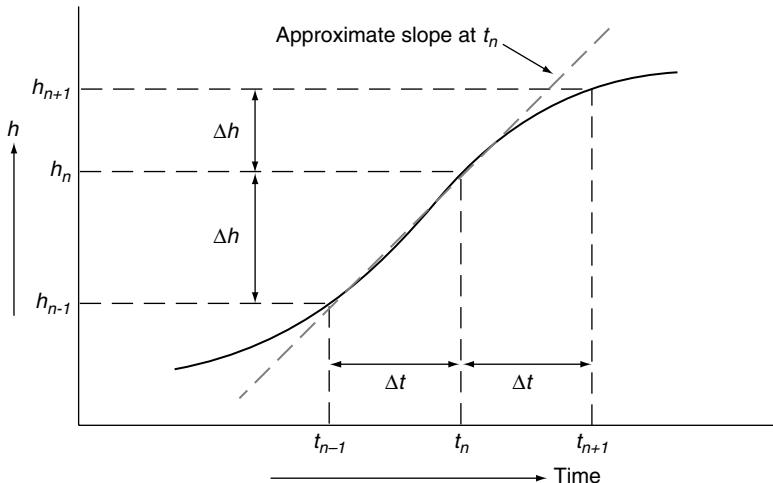


FIGURE 23.4 Part of a hydrograph showing that the derivative (or slope, $\partial h / \partial t$) at time t_n may be approximated by $\Delta h / \Delta t$. (From Konikow, L. F. (1996) In *Manual on Mathematical Models in Isotope Hydrogeology*, International Atomic Energy Agency, Vienna, Austria.)

heads calculated at specific time increments (or time steps), the slope of the hydrograph at time n can be approximated by

$$\left(\frac{\partial h}{\partial t} \right)_{n\Delta t} \approx \frac{h_{n+1} - h_n}{\Delta t} \quad (23.16)$$

or

$$\left(\frac{\partial h}{\partial t} \right)_{n\Delta t} \approx \frac{h_n - h_{n-1}}{\Delta t} \quad (23.17)$$

Calculation of the derivative at $t = n\Delta t$ using Equation 23.16 takes a “forward difference” from time n to time $n+1$, and calculation using Equation 23.17 takes a “backward difference.” In terms of solving the groundwater-flow equation for a node (i, j) of a 2D finite-difference grid, we have to consider heads at five nodes and at two time levels, as illustrated in Figure 23.5. In Figure 23.5a, we have expressed the spatial derivatives of head at time level n , where all values are known, and the time derivative as a forward difference to the unknown head at time step $n+1$. Then for every node of the grid, we will have a separate difference equation, each of which contains only one unknown variable. Thus, these equations can be solved explicitly. Explicit finite-difference equations are simple and straightforward to solve, but may have stability criteria associated with them. If time increments are too large, small numerical errors or perturbations may propagate into larger errors at later stages of the computations.

In Figure 23.5b, we have expressed the time derivative as a backward difference from the heads at time level n , which are the unknown heads, to the heads at the previous time level, $n - 1$, which are known (either from initial conditions specified for the first time step or from the solution for previous time steps). The spatial derivatives of head are written at time level n , where all values are unknown; so, for every node of the grid, we will have one difference equation that contains five unknowns, which cannot be solved directly. However, for the entire grid, which contains N nodes, we would have a system of N equations containing a total of N unknowns. Such a system of simultaneous equations, together with specified boundary conditions, can be solved implicitly. Although implicit solutions are more complicated, they also have the advantage of generally being unconditionally stable. This stability implies that a solution will be obtained; although not necessarily that the estimate of the derivative that is calculated will be accurate,

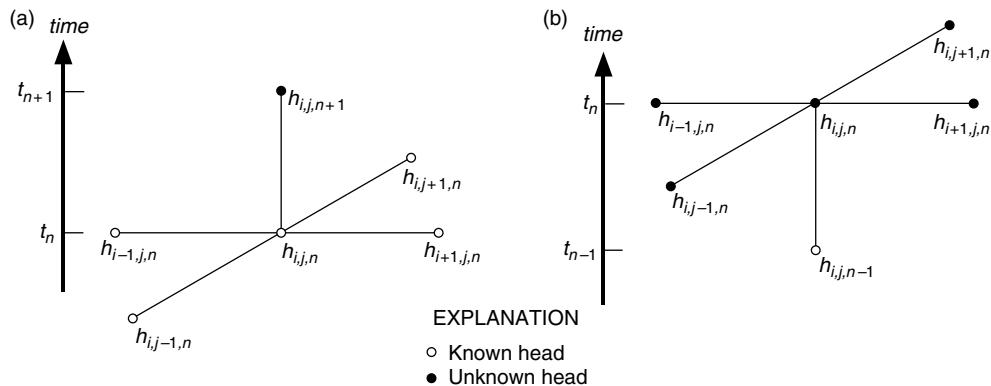


FIGURE 23.5 Grid stencil showing discretization of time at node (i, j) in a 2D finite-difference grid: (a) explicit (forward-difference) formulation, and (b) implicit (backward-difference) formulation. (From Konikow, L. F. (1996) In *Manual on Mathematical Models in Isotope Hydrogeology*, International Atomic Energy Agency, Vienna, Austria.)

even if the time steps are large relative to the rate of change of head. Most available groundwater-flow models solve an implicit finite-difference approximation to the flow equation.

We may next consider Equation 23.5, a 2D groundwater-flow equation for a heterogeneous, anisotropic aquifer, in which the coordinate system is aligned with the major axes of the transmissivity tensor. This may be approximated by the following implicit finite-difference equation for representative node (i, j) as

$$\begin{aligned} & T_{xx[i-1/2,j]} \left(\frac{h_{i-1,j,n} - h_{i,j,n}}{(\Delta x)^2} \right) + T_{xx[i+1/2,j]} \left(\frac{h_{i+1,j,n} - h_{i,j,n}}{(\Delta x)^2} \right) \\ & + T_{yy[i,j-1/2]} \left(\frac{h_{i,j-1,n} - h_{i,j,n}}{(\Delta y)^2} \right) + T_{yy[i,j+1/2]} \left(\frac{h_{i,j+1,n} - h_{i,j,n}}{(\Delta y)^2} \right) \\ & = S \left(\frac{h_{i,j,n} - h_{i,j,n-1}}{\Delta t} \right) - \frac{q_{i,j}}{\Delta x \Delta y} - \frac{K_z}{m} (H_{s[i,j]} - h_{i,j,n}) \end{aligned} \quad (23.18)$$

where $q_{i,j}$ is the volumetric rate of withdrawal (negative in sign) or recharge (positive) at the i, j node, L^3T^{-1} . This formulation assumes that any stresses, such as represented by $q_{i,j}$, are applied over the entire surface area of cell i, j rather than at a point (or at node i, j). This condition implies that if a pumping well is represented at node i, j , then the head will be calculated as if it were being withdrawn from a well that had a cross-sectional area for the borehole equal to $\Delta x \Delta y$, rather than the actual cross-sectional area of the well. In Equation 23.18, the transmissivity terms represent the harmonic means of the transmissivity of the two adjacent cells. The harmonic mean can be shown to be appropriate and consistent with the assumption that transmissivity is constant and uniform within each cell but may differ between cells. Other types of means for interblock transmissivity may be more appropriate for other assumptions about the transmissivity distribution, such as smoothly varying transmissivity (Goode and Appel 1992).

23.5.2 Basics of Finite-Element Methods

The finite-element method (FEM) is a numerical analysis technique for obtaining approximate solutions to a wide variety of problems in physics and engineering. The method was originally applied to structural mechanics but is now used in all fields of continuum mechanics. Huebner (1975) describes four different approaches to formulate the finite-element method for a problem, which are, the direct approach, the variational approach, the weighted residual approach, and the energy balance approach. In groundwater problems, the approach frequently used is either the weighted residual or variational approach.

The concept of “piecewise approximation” is used in the FEM. The domain of the problem, that is, the extent of the groundwater system to be simulated, is divided into a set of elements or pieces. In theory,

the elements can be of different shapes and sizes. Most FEM computer programs use one shape element, most commonly either triangular or quadrilateral elements. In the groundwater model MODFE (Torak, 1993 and Cooley, 1992) triangular elements are used, whereas in the groundwater model SUTRA (Voss and Provost, 2002), quadrilateral (2D) or hexahedral (3D) elements are used. Point values of the dependent variable (e.g., head, pressure, or concentration) are calculated at nodes, which are the corners or vertices of the elements; a simple equation is used to describe the value of the dependent variable within the element. This simple equation is called a basis function and each node that is part of an element has an associated basis function. The simplest basis functions that are usually used are linear functions. The solution to the differential equation for flow (Equation 23.3 or Equation 23.6) or transport (Equation 23.9 or Equation 23.10) is approximated by a set of elements in which the dependent variable only varies linearly within the element, but the entire set of elements approximates the complex distribution of head or concentration. The approximate modeled hydraulic head distribution (Figure 23.6c) is comprised of a set of triangular elements (Figure 23.6a) having a linear approximation of head variation within each element (Figure 23.6b).

In the method of weighted residuals, the “piecewise” continuous surface is obtained by minimizing the difference between the approximate surface and the continuous surface. The method of weighted residuals is summarized by Huyakorn and Pinder (1983, p. 39) as follows. Any differential equation $L(h)$, such as the steady-state form of Equation 23.3 (the groundwater-flow equation) can be written as

$$L(h) = 0 \quad (23.19)$$

over the domain of the problem R . The first step in obtaining the approximate solution is to define the approximate solution as the sum of all the simple basis functions as

$$\hat{h} = \sum_{i=1}^n N_i Z_i \quad (23.20)$$

where \hat{h} is the approximate solution, n is the number of linearly independent basis functions, N_i are the linearly independent basis functions defined over the entire domain, and Z_i are the unknown coefficients to be determined (there is one coefficient for each node in the finite-element mesh). The trial function \hat{h} is an approximation, so that when it is substituted into Equation 23.19 there will be some error, ξ , defined as

$$\xi = L(\hat{h}) \quad (23.21)$$

The method of weighted residuals determines the unknown coefficients by minimizing the error. This minimization is accomplished by weighting the error, integrating the error, and setting the error equal to zero over the entire domain. A weighting function, W_i , can be specified for each basis function and the resulting integration is

$$\int_R W_i \xi \, dR = \int_R W_i L(\hat{h}) \, dR = 0, \quad i = 1, 2, \dots, n \quad (23.22)$$

Equation 23.20 is substituted into Equation 23.22 and weighting functions are specified. There are then n equations and n unknowns. The selection of the weighting functions and the simplification of the integral in Equation 23.22 into a linear algebraic equation is mathematically straightforward, but not intuitive. In the Galerkin method, the weighting functions are chosen to be identical to the basis functions and Equation 23.22 is simplified by using integration by parts. Because the basis functions and weighting functions are defined to be of a specific algebraic form (e.g., linear basis functions), the modified integral is straightforward to solve and becomes a set of n simultaneous algebraic equations.

After Equation 23.22 is mathematically evaluated into a set of n simultaneous equations, they are solved using matrix solution techniques for the n unknown coefficients Z_i , and the approximate solution \hat{h} is

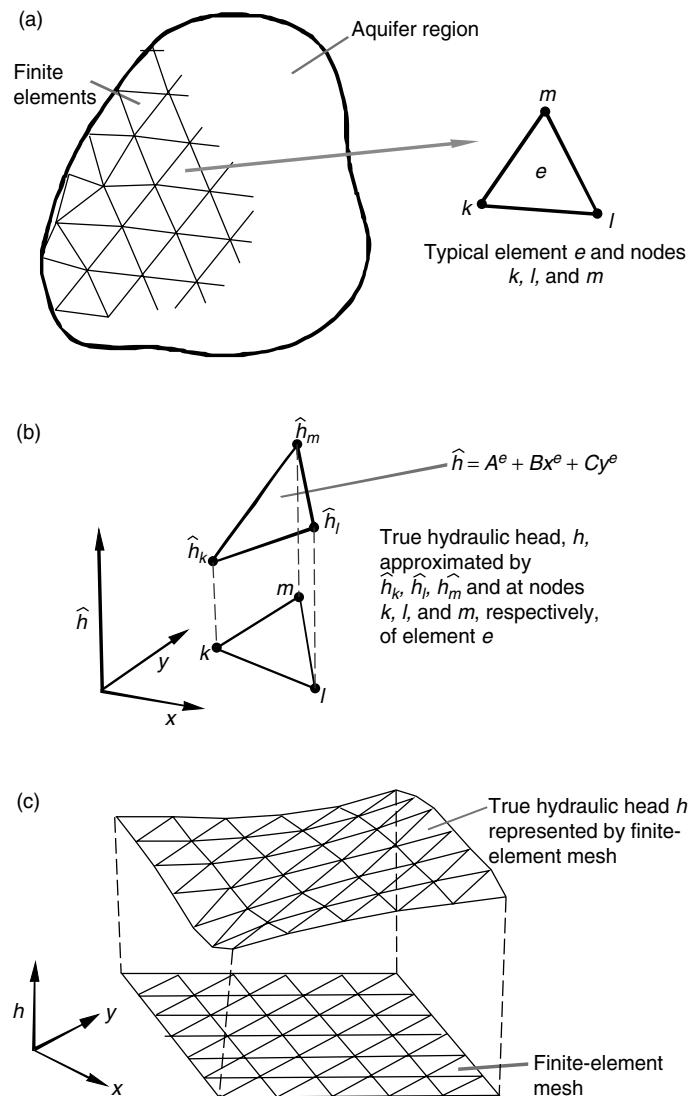


FIGURE 23.6 Diagram showing (a) aquifer region partially subdivided by finite elements and typical element e , (b) finite-element representation of hydraulic head \hat{h} , and (c) finite-element mesh configuration for approximating true hydraulic head. (From Torak, L. J. (1993) A modular finite-element model (MODFE) for areal and axisymmetric ground-water-flow problems, Part 1: Model description and user's manual, *Techniques of Water-Res. Invests. of the U.S. Geol. Survey*, Book 6, Ch. A3.)

determined at each node. The time derivative is frequently approximated by finite differences as discussed in the previous section. Huyakorn and Pinder (1983), Huebner (1975), Zienkiewicz (1971), Wang and Anderson (1982), and Cooley (1992) provide more comprehensive explanations of the method.

23.5.3 Basics of Method-of-Characteristics Methods

The method-of-characteristics (MOC) was developed to solve hyperbolic differential equations (advectively dominated transport equations). A major advantage of this over other methods is that MOC minimizes numerical dispersion (Garder et al. 1964; Reddell and Sunada 1970; Zheng and Bennett 2002). The approach taken with the MOC is not to solve Equation 23.9 directly, but rather to solve an equivalent

system of ordinary differential equations. A form of Equation 23.9, accounting for equilibrium-controlled sorption or exchange and first-order irreversible rate reactions, can be further modified for improved compatibility with this method by expanding the advection term, substituting relations from Darcy's law and the flow equation, and rearranging terms to obtain

$$\frac{\partial C}{\partial t} = \frac{1}{R_f} \frac{\partial}{\partial x_i} \left(D_{ij} \frac{\partial C}{\partial x_j} \right) - \frac{V_i}{R_f} \frac{\partial C}{\partial x_i} + \frac{W^*(C - C^*)}{\varepsilon R_f} - \lambda C \quad (23.23)$$

where R_f is defined as a dimensionless retardation factor, $R_f = 1 + (\rho_b K_d / \varepsilon)$, and K_d is the distribution coefficient, $L^3 M^{-1}$. If we consider the material derivative of concentration with respect to time, dC/dt , as describing the change in concentration of a parcel of water moving at the seepage velocity of water, it may be defined for a 3D system as

$$\frac{dC}{dt} = \frac{\partial C}{\partial t} + \frac{\partial C}{\partial x} \frac{dx}{dt} + \frac{\partial C}{\partial y} \frac{dy}{dt} + \frac{\partial C}{\partial z} \frac{dz}{dt}. \quad (23.24)$$

The second, third, and fourth terms on the right side include the material derivatives of position, which are defined by the velocity in the x , y , and z directions. We then have

$$\frac{dx}{dt} = \frac{V_x}{R_f} \quad (23.25)$$

$$\frac{dy}{dt} = \frac{V_y}{R_f} \quad (23.26)$$

$$\frac{dz}{dt} = \frac{V_z}{R_f} \quad (23.27)$$

and

$$\frac{dC}{dt} = \frac{1}{R_f} \frac{\partial}{\partial x_i} \left(D_{ij} \frac{\partial C}{\partial x_j} \right) + \frac{W^*(C - C^*)}{\varepsilon R_f} - \lambda C \quad (23.28)$$

The solutions of the system of equations comprising Equation 23.25 through Equation 23.28 may be given as $x = x(t)$, $y = y(t)$, $z = z(t)$, and $C = C(t)$, and are called the characteristic curves of Equation 23.23. Given solutions to Equation 23.25 through Equation 23.28, a solution to the partial differential equation may be obtained by following the characteristic curves, which are defined by the particle pathlines. This may be accomplished by introducing a set of moving points (or reference particles) that can be traced within the stationary coordinates of a finite-difference grid. Each particle corresponds to one characteristic curve, and values of x , y , z , and C are obtained as functions of t for each characteristic (Garder et al. 1964). Each point has a concentration and position associated with it and is moved through the flow field in proportion to the flow velocity at its location (see Figure 23.7). The concentrations at the nodes of the fixed finite-difference grid may then be estimated as an arithmetic or weighted mean of the concentrations of all particles contained within the cell area for that node. In the traditional MOC, the nodal concentrations are used to estimate spatial concentration gradients for an explicit or implicit finite-difference solution to Equation 23.28. However, the transfer between the moving grid of particles and the fixed grid of finite-difference nodes involves a loss of accuracy (generally seen as a small amount of numerical dispersion). An alternate approach is to use a particle strength exchange method, which uses particle positions at the end of each time increment as quadrature points in what is essentially a finite-element solution to Equation 23.28 (Zimmermann et al. 2001). This eliminates the problems involved in interpolating from particles to nodes and back again to compute the changes in concentration caused by hydrodynamic dispersion, but does so at a cost of increased computational effort.

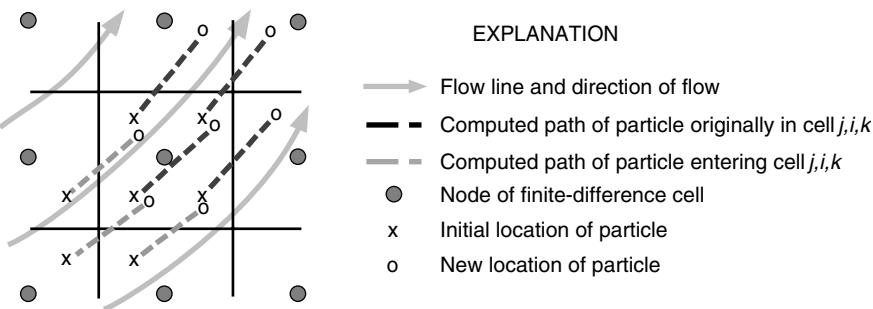


FIGURE 23.7 Part of a hypothetical finite-difference grid showing relation of flow field to movement of particles in method of characteristics model for simulating solute transport. (Adapted from Konikow, L. F. and Bredehoeft, J. D. (1978) Computer model of two-dimensional solute transport and dispersion in ground water, *Techniques of Water-Res. Invests. of the U.S. Geol. Survey*, Book 7, Ch. C2.)

23.5.4 Matrix Solution Techniques

As indicated previously, the finite-difference and finite-element approximations lead to an algebraic equation for each node point. The set of algebraic equations may be solved numerically by one of two basic methods: direct or iterative. In direct methods, a sequence of operations is performed only once to solve the matrix equation, providing a solution that is exact, except for machine round-off error. Iterative methods arrive at a solution by a process of successive approximation. The process involves making an initial guess at the solution, then improving this guess by some iterative process until an error criterion is satisfied. Therefore, in these techniques, convergence and the rate of convergence are of concern.

Direct methods have two main disadvantages when compared to indirect methods. The first problem is one of computer resource requirements, including large storage (memory) requirements and long computation times for large problems. The matrix is sparse (contains many zero values) and to minimize computational effort, various techniques have been proposed. However, for finite-difference and finite-element methods, storage requirements may still prove to be unavoidably large for 3D problems. The second disadvantage of direct methods is round-off error. As many arithmetic operations are performed, round-off errors can accumulate for certain types of matrices.

Iterative schemes avoid the need for storing large matrices, which make them attractive for solving large problems. Numerous schemes have been developed (Barrett et al. 1994). A few of the more commonly used ones include successive over-relaxation methods, iterative alternating-direction implicit procedure, the strongly implicit procedure, and conjugate-gradient methods with various types of preconditioning.

As iterative methods start with an initial estimate for the solution, the efficiency of the method depends somewhat on this initial guess. To speed up the iterative process, relaxation and acceleration factors are sometimes used. Unfortunately, the definition of best values for these factors commonly is problem dependent. In addition, iterative approaches require that an error tolerance be specified to stop the iterative process. An optimal value for the tolerance, which is used to evaluate when the iterative calculations have converged on a solution, may also be problem dependent. If the tolerance is set too large, then the iterations may stop before adequate numerical accuracy is achieved. If the tolerance is set too small, then the iterative process may consume excessive computational resources in striving for numerical precision that may be orders of magnitude smaller than the precision of the field data, or the iterative process may even fail to converge.

23.5.5 Boundary and Initial Conditions

To obtain a unique solution to a partial differential equation corresponding to a given physical process, additional information about the physical state of the process is required. This information is supplied

TABLE 23.1 Common Designations for the Three Common Mathematical Boundary Conditions Specified in Mathematical Analyses of Groundwater-Flow and Transport Systems (n is directional coordinate normal to the boundary (L))

Boundary Type and Name	Formal Name	Mathematical Designation (Equation for head and equation for concentration)
Type 1 Specified value	Dirichlet	$h(x, y, z, t) = \text{constant}$ $C(x, y, z, t) = \text{constant}$
Type 2 Specified flux	Neumann	$\frac{dh(x, y, z, t)}{dn} = \text{constant}$ $\frac{dC(x, y, z, t)}{dn} = \text{constant}$
Type 3 Value-dependent flux	Cauchy	$\frac{dh}{dn} + ah = \text{constant}$ $\frac{dC}{dn} + aC = \text{constant}$ (where a is also a constant)

Sources: Franke, O.L., Reilly, T.E. and Bennett, G.D. (1987) Definition of boundary and initial conditions in the analysis of saturated groundwater flow systems — An introduction, *Techniques of Water-Res. Invests. of the U.S. Geol. Survey*, Book 3, Ch. B5; and Reilly, T.E. (2001) System and boundary conceptualization in ground-water flow simulation: *Techniques of Water-Res. Invests. of the U.S. Geol. Survey*, Book 3, Ch. B8.

by boundary and initial conditions. For steady-state problems, only boundary conditions are required, whereas for transient problems, boundary and initial conditions must be specified.

Mathematically, the boundary conditions include the geometry of the boundary and the values of the dependent variable or its derivative normal to the boundary. Internal sources and sinks are also considered as boundary conditions in the solution to the governing equations. In physical terms, for groundwater-model applications, the boundary conditions are generally of three types (Table 23.1): (1) specified value (head or concentration), (2) specified flux (corresponding to a specified gradient of head or concentration), or (3) value-dependent flux (or mixed boundary condition, in which the flux across a boundary is related to both the normal derivative and the value) (Mercer and Faust 1981; Franke et al. 1987; Reilly 2001). The third type of boundary condition might be used, for example, to represent leakage or exchange between a stream and an adjacent aquifer, in which the leakage may change over time as the head in the aquifer changes, even though the head in the stream might remain fixed. A no-flow boundary is a special case of the second type of boundary condition. A fourth type of boundary condition that is important in many groundwater-flow problems is the free-surface boundary, in which head in the groundwater system is a function of only the potential energy of a water particle because of its position above some datum, that is, its elevation head (Franke et al. 1987). The most common type of free-surface boundary is the water table, which is the boundary surface between the saturated flow field and the atmosphere. The types of boundaries appropriate to a particular field problem require careful consideration.

The initial conditions are simply the values of the dependent variable specified everywhere inside the boundary at the start of the simulation. Preferably, the initial conditions are specified from a steady-state simulation of the groundwater system. If, however, initial conditions are directly specified on the basis of field measurements, they may not be consistent with the model conceptualization or the stresses specified in the model. In that event, transient flow and (or) transient solute transport may occur in the system at the start of the simulation in response to both the new hydrologic stresses and the initial conditions (Franke et al. 1987; Reilly and Harbaugh 2004).

23.5.6 Generic Model Verification

It must be demonstrated that the generic model accurately solves the governing equations for various boundary value problems, an evaluation that is often called model “verification.” This evaluation is checked by demonstrating that the code gives good results for problems having known solutions. This test is usually done by comparing the numerical model results to those of an analytical solution. Numerical accuracy is rarely a problem for the solution to the flow equation, but may sometimes be a major problem in transport modeling.

Numerical solutions are sensitive to spatial and temporal discretization. Therefore, even a perfect agreement for test cases only proves that the numerical code can accurately solve the governing equations, not that it will accurately solve problems under any and all circumstances.

Analytical solutions generally require simple geometry, uniform properties, and idealized boundary and initial conditions. The power of the numerical methods is that they are not constrained by the simplification imposed by analytical methods and allow the introduction of nonhomogeneous, anisotropic parameter sets, irregular geometry, mixed boundary conditions, and even nonlinearities into the boundary value problems. Usually, analytical solutions approximating these complexities are unavailable for comparison. Therefore, once these complexities are introduced, there is no definitive basis for verifying the numerical model.

One approach that improves confidence for complex heterogeneous problems is to compare the model results to experimental data, to results of other well-accepted models, or to some other accepted standard. Such evaluations might best be termed benchmarking. The HYDROCOIN Project used standardized problem definitions as a basis for intercode comparisons (Swedish Nuclear Power Inspectorate 1987). Although this type of benchmarking helps ensure consistency, it does not guarantee or measure accuracy (Konikow et al. 1997). A collection and detailed discussion of a number of classical groundwater problems that have been used historically as a basis of model evaluation are presented and documented by Ségol (1994).

23.6 Model Design, Development, and Application

The first step in model design and application is to define the nature of the problem and the purpose of the model (Figure 23.8). Although this may seem obvious, it is an important first step that is sometimes overlooked in a hasty effort to take action. This step is linked closely with the formulation of a conceptual model, which is required prior to the development of a mathematical model. In formulating a conceptual model, the analyst must evaluate which processes are important in the groundwater system being investigated for the particular problem at hand. Some processes may be important to consider at one scale of study, but negligible or irrelevant at another scale of investigation. The analyst must similarly decide on the appropriate dimensionality and resolution for the numerical model. Good judgment is required to evaluate and balance the tradeoffs between accuracy and cost, with respect to model development, model use, and data requirements. The key to efficiency and accuracy in modeling a system probably is more affected by the formulation of a proper and appropriate conceptual model than by the choice of a particular numerical method or code.

Once a decision to develop a model has been made, a code (or generic model) must be selected (or modified or constructed) that is appropriate for the given problem. Next, the generic code must be adapted to the specific site or region being simulated. Development of a numerical deterministic, distributed-parameter, simulation model involves selecting or designing spatial grids and time increments that will yield an accurate solution for the given system and problem. The analyst must then specify the properties of the system (and their distributions), stresses on the system (such as recharge and pumping rates), boundary conditions, initial conditions (for transient problems), and geochemical processes/reactions (if appropriate). All of the parameter specifications and boundary conditions must be consistent with the overall conceptual model of the system, and the initial numerical model reflects the analyst’s conceptual model of the system.

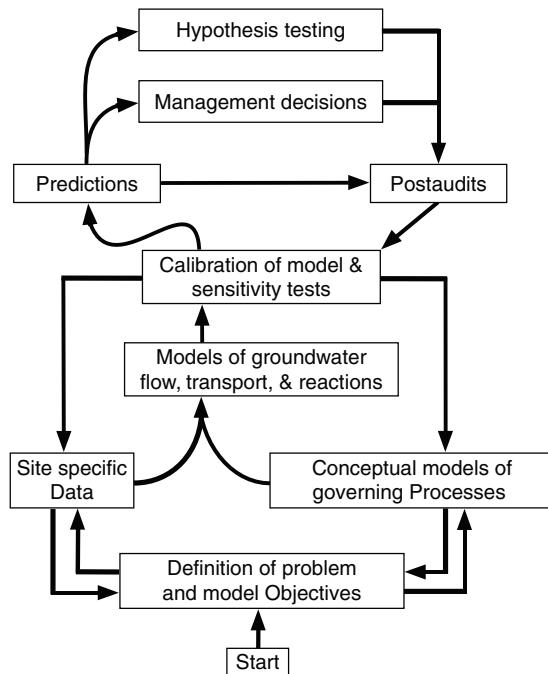


FIGURE 23.8 The use and application of models in the analysis of groundwater problems. (From Konikow, L. F. (1996) In *Manual on Mathematical Models in Isotope Hydrogeology*, International Atomic Energy Agency, Vienna, Austria.)

A model is a simplified approximation of a complex reality, but the model should capture the essential features and processes relative to the problem at hand. The selection of the appropriate model and appropriate level of model complexity (or, rather, simplicity) remains subjective and dependent on the judgment and experience of the analysts, the objectives of the study, the level of prior information available for the system of interest, and the complexity of the system being modeled. The trade-off between model accuracy and model cost will always be difficult to resolve, but will always have to be made. Water managers and other users of model results must be made aware that these trade-offs and judgments have been made and may affect model reliability. Reilly and Harbaugh (2004) discuss many of these design decisions and their trade-offs in the context of developing guidelines to aid water managers in their evaluation of the usefulness and appropriateness of a groundwater-flow model.

In general, it is more difficult to calibrate a solute-transport model of an aquifer than it is to calibrate a groundwater-flow model. Fewer parameters need to be defined to compute the head distribution with a flow model than are required to compute concentration changes with similar confidence using a solute-transport model. Also, in typical field problems, defining the source term for a solute-transport model is especially difficult for point-source contamination problems because the timing, location, and strength of releases of solute mass into an aquifer system are rarely known or reported accurately (and is commonly the very point of contention in litigation).

Groundwater seepage velocity is determined from the head distribution, and because both advective transport and hydrodynamic dispersion are functions of the seepage velocity, a model of groundwater flow typically is calibrated before a pathline or solute-transport model is developed. In a field environment, perhaps the single most important key to understanding a solute-transport problem is the development of an accurate description (or model) of flow. This is particularly relevant to transport in fractured rocks, where simulation is commonly based on porous-media concepts. In highly heterogeneous systems, the head distribution and flow directions often can be simulated fairly accurately, whereas the calculated velocity field still may be greatly in error, which results in considerable errors in transport simulations.

23.6.1 Grid Design

The dimensionality of the model (i.e., one, two, or three dimensions) should be selected during the formulation of the conceptual model. If a 1D or 2D model is selected, then it is important that the grid be aligned with the flow system so that there is no unaccounted flux into or out of the line or plane of the grid. For example, if a 2D areal model is applied, then there should be no significant vertical components of flow and any vertical leakage or flux must be accounted for by boundary conditions. If a 2D profile model is applied, then the line of the cross section should be aligned with an areal streamline, and there should not be any significant lateral flow into or out of the plane of the cross section.

To minimize a variety of sources of numerical errors, the model grid should be designed using the finest mesh spacing and time steps that are possible, given limitations on computer memory and computational time. The boundaries of the grid should be aligned, to the extent possible, with natural hydrologic and geologic boundaries of the groundwater system. Where it is impractical to extend the grid to a natural boundary, then an appropriate boundary condition should be imposed at the edge of the grid to represent the net effects of the continuation of the aquifer beyond the grid. This representation can typically be accomplished using head-dependent leakage (third type) boundary conditions. However, this would preclude calculating (and accounting for) any storage changes outside the active grid. These boundaries should also be placed as far as possible away from the area of interest and areas of stresses on the system, so as to minimize any effect of conceptual errors associated with these artificial boundary conditions. Note that it is possible for certain types of hydraulic boundaries, such as a groundwater divide, to change location over time if they are located near a major hydraulic stress. If such a boundary shift is anticipated, it might be preferable to extend the boundary of the grid some distance beyond the location of such a natural boundary.

In designing the grid, the length to width ratio (or aspect ratio) of cells or elements should be kept as close to one as possible. Long linear cells or elements can lead to numerical instabilities or errors, and should be avoided, particularly if the aspect ratio is greater than about five (Bear and Verruijt 1987). However, this is not a stringent guideline as aspect ratios exceeding 1000:1 are routinely used with the SUTRA code (Voss and Provost 2002) without numerical problems. For example, a ratio of 500:1 was used by Voss and Souza (1998) in a model of the Pearl Harbor aquifer, Oahu (Section 23.8.3).

In specifying boundary conditions for a particular problem and grid design, care must be taken not to over constrain the solution. That is, if dependent values are fixed at too many nodes of a grid, the model may have too little freedom to calculate a meaningful solution. At the extreme, by manipulating boundary conditions, one can force any desired solution at any given node. Although a forced solution may ensure a perfect match to observed data used for calibration, such a match is of course not an indicator of model accuracy or reliability, and can be meaningless (Franke and Reilly 1987).

To optimize computational resources in a model, it sometimes is advisable to use an irregular (or variably-spaced) mesh in which the grid is finest in areas of point stresses, where gradients are steepest, where data are most dense, where the problem is most critical, and (or) where greatest numerical accuracy is desired. Similarly, the length of time steps can often be increased geometrically during a transient simulation. At the initial times or after a change in the stress regime, small time steps should be imposed, as that is when changes in the dependent variable over time are the greatest. With increased elapsed time, the rate of change in head typically decreases, so time steps can often be safely increased.

Because transmissivity (or hydraulic conductivity) is a property of the porous media, the cross-product terms of the transmissivity tensor drop out of the governing flow equation that is solved in a model by aligning the model grid with the major axes of the transmissivity tensor (as represented in Equation 23.5). This makes the code simpler and more efficient, and is a required assumption for most finite-difference models. However, this same simplification typically is not possible for the dispersion tensor in the transport equation because it depends on the flow direction, which changes orientation over space and time. In general, it is not possible to design a fixed grid that will always be aligned with a changing flow field.

23.6.2 Model Calibration

Deterministic groundwater-simulation models impose large requirements for data to define all of the parameters at all of the nodes of a grid. To determine uniquely the parameter distribution for a field problem, so much expensive field testing would be required that it is seldom feasible either economically or technically. Therefore, the model typically represents an attempt, in effect, to solve a large set of simultaneous equations having more unknowns than equations. It is inherently impossible to obtain a unique solution to such a problem.

Uncertainty in parameters logically leads to a lack of confidence in the interpretations and predictions that are based on a model analysis, unless the model can be demonstrated to be a reasonably accurate representation of the real groundwater system. To demonstrate that a deterministic groundwater-simulation model is realistic, usually field observations of aquifer responses (such as changes in water levels and fluid flux for flow problems or changes in concentration for transport problems) are compared to corresponding values calculated by the model. The objective of this calibration procedure is to minimize differences between the observed data and calculated values. Usually, the model is considered calibrated when it reproduces historical data within some acceptable level of accuracy. The level of acceptability is, of course, determined subjectively. Although a poor match provides evidence of errors in the model, a good match in itself does not prove the validity or adequacy of the model (Konikow and Bredehoeft 1992).

Because of the large number of variables in the set of simultaneous equations represented in a model, calibration will not yield a unique set of parameters. Where the match is poor, it suggests (1) an error in the conceptual model, (2) an error in the numerical solution, and (or) (3) a poor set of parameter values. Even when the match to historical data is good, the model may still fail to predict future responses accurately, especially under a new or extended set of stresses than experienced during the calibration period.

The calibration of a deterministic groundwater model is often accomplished through a trial-and-error adjustment of the model input data (aquifer properties, sources and sinks, and boundary and initial conditions) to modify model output. As a large number of interrelated factors affect the output, trial-and-error adjustment may become a highly subjective and inefficient procedure. Advances in automated parameter-estimation procedures help to eliminate some of the subjectivity inherent in model calibration. The newer approaches generally treat model calibration as a statistical procedure using multiple regression approaches. This is sometimes called inverse modeling because instead of using defined parameter values to solve the governing flow equation for head, the procedure strives to achieve a best fit to observed heads by adjusting (in effect, solving for) unknown or uncertain parameter values. The procedure to achieve a best fit uses sensitivity coefficients that are based on the change in calculated value divided by the change in the parameter (e.g., the change in head with changing hydraulic conductivity). The sensitivity coefficients may be useful in the consideration of additional data collection. Hill (1998) provides an overview of methods and guidelines for effective model calibration using inverse modeling.

Parameter-estimation procedures improve the efficiency of model calibration and allow the simultaneous construction, application, and calibration of a model using uncertain data. Statistical measures are used to quantify the uncertainties both in model parameters and in predictions. One example of a parameter-estimation program is UCODE (Poeter and Hill 1998; Poeter et al. 2005), which performs inverse modeling using nonlinear regression. UCODE is a powerful public domain code because it will work with any model—not just a groundwater-flow model. An estimated parameter can be a quantity that appears in the input files of the model (such as hydraulic conductivity in a flow model); observations to be matched in the regression can be any quantity for which a simulated equivalent value can be produced (such as head or outflow to a stream) (Poeter and Hill 1998).

Even with regression modeling, the hydrologic experience and judgment of the modeler in formulating a proper conceptual model continues to be a major factor in calibrating a numerical simulation model both accurately and efficiently. The modeler should always be familiar with the specific field area being studied to ensure that both the database and the numerical model adequately represent prevailing field conditions. The modeler must also recognize that uncertainty in specification of sources, sinks, and boundary and initial conditions should be evaluated during the calibration procedure in the same manner

as uncertainty in aquifer properties. Failure to recognize the uncertainty inherent both in the input data and in the calibration data may lead to “fine-tuning” of the model through unjustifiably precise parameter adjustments strictly to improve the match between observed and calculated variables. This may serve only to provide a false confidence in the model without producing an equivalent (or any) increase in the predictive accuracy of the model or any improved conceptual understanding of the real groundwater system. Freyberg (1988) illustrated this in an exercise in which several groups were given the task of modeling a particular hypothetical groundwater problem. The group that achieved the best calibration, as measured by the minimum root mean square error, was not the group that developed the model that yielded the best prediction (measured by the same criterion). Freyberg (1988, p. 360) concluded that “simple measures of the goodness of a calibrated fit to head data are inadequate to evaluate the true worth of a calibrated parameter set.”

The use of deterministic models in the analysis of groundwater problems is illustrated, in a general sense, in Figure 23.8. Perhaps the greatest value of the modeling approach is its capacity to integrate site-specific data with equations describing the relevant processes as a quantitative basis for predicting changes or responses in a groundwater system. There must be allowances for feedback from the stage of interpreting model output both to the data collection and analysis phase and to the conceptualization and mathematical definition of the relevant governing processes. One objective of model calibration should be to improve the conceptual model of the system. As the model quantitatively integrates the effects of the many factors that affect groundwater flow or solute transport, the calculated results should be internally consistent with all input data, and it can be determined if any element of the conceptual model should be revised. Prior concepts or interpretations of aquifer parameters or variables, such as represented by potentiometric maps or the specification of boundary conditions, may be revised during the calibration procedure as a result of feedback from the model output. Another objective of modeling should be to define inadequacies in the database and help set priorities for the collection of additional data.

Automated parameter-estimation techniques improve the efficiency of model calibration and have two general components — one that calculates the best fit (sometimes called automatic history matching) and one that evaluates the statistical properties of the fit. The objective of automatic history matching is to obtain the estimates of system parameters that yield the closest match (minimize deviations) between observed data and model calculations. Least squares deviation is usually chosen as a criterion. The minimization procedure uses the sensitivity coefficients described previously.

Parameter uncertainty is commonly addressed using a sensitivity analysis. A major objective of sensitivity analysis of simulation models is to determine the change in model results as a result of changes in the model input or system parameters. Direct parameter sampling in which parameters are perturbed one by one and the complete set of system equations is resolved is used in conventional sensitivity analysis (Konikow and Mercer, 1988).

23.6.3 Model Error

Discrepancies between observed and calculated responses of a groundwater system are the manifestation of errors in the conceptual or mathematical model. In applying groundwater models to field problems, there are three sources of error, and it may not be possible to distinguish among them (Konikow and Bredehoeft, 1992). One source is conceptual errors — that is, misconceptions about the basic processes that are incorporated in the model. Conceptual errors include both neglecting relevant processes as well as inappropriate representation of processes. Examples of such errors include the use of a 2D model where appreciable flow or transport occurs in the third dimension, or the application of a model based upon Darcy's law to media or environments where Darcy's law is inappropriate. A second source of error involves numerical errors arising in the equation-solving algorithm. These errors include truncation errors, round-off errors, and numerical dispersion. A third source of error arises from uncertainties and inadequacies in the input data that reflect our inability to describe comprehensively and uniquely the aquifer properties, stresses, and boundaries. In most model applications, conceptualization problems and uncertainty concerning the input data are the most common sources of error.

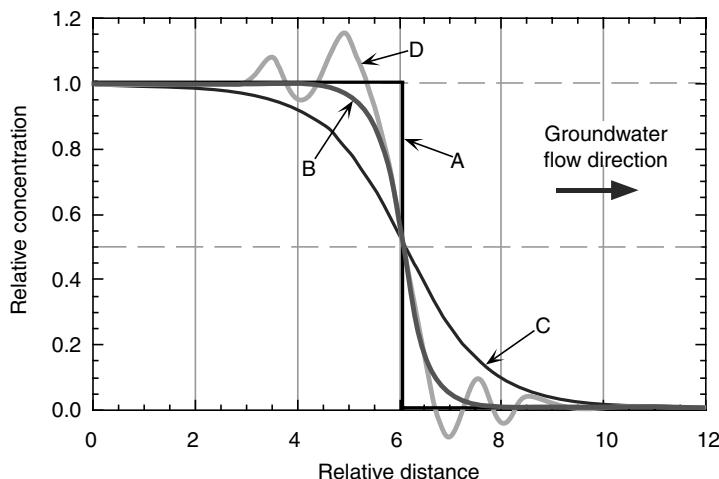


FIGURE 23.9 Representative breakthrough curves for a simple flow and transport problem to illustrate types of numerical errors that may occur in numerical solution to transport equation: (a) plug flow having no dispersion, (b) “exact” solution for transport with dispersion, (c) numerical solution for case B that exhibits effects of numerical dispersion, and (d) numerical solution for case B with oscillatory behavior. (From Konikow, L. F. (1996) In *Manual on Mathematical Models in Isotope Hydrogeology*, International Atomic Energy Agency, Vienna, Austria.)

Numerical methods, in general, yield approximate solutions to the governing equations. There are a number of possible sources of numerical error in the solution. If model users are aware of the source and nature of these errors, they can control them and interpret the results in light of the presence of error. For example, if the spatial grid spacing is too coarse, truncation errors affect the accuracy of the solution.

In solving advection-dominated transport problems in which a relatively sharp front (or steep concentration gradient) is moving through a groundwater system, it is numerically difficult to preserve the sharpness of the front. Obviously, if the width of the front is narrower than the node spacing, then it is inherently impossible to calculate the correct values of concentration in the vicinity of the sharp front. However, even in situations where a front is less sharp, the numerical solution technique can calculate a greater dispersive flux than would occur by physical dispersion alone or would be indicated by an exact solution of the governing equation. That part of the calculated dispersion (or spreading of solute about the center of mass) introduced solely by the numerical solution algorithm is called numerical dispersion.

Calculated breakthrough curves at a particular point in time for a hypothetical problem of 1D uniform flow and transport, after a tracer having a relative concentration of 1.0 was continuously injected at some point upstream, are illustrated in Figure 23.9. Curve A represents the breakthrough curve and position of a sharp front for a case having no dispersion (plug flow). Curve B represents an exact analytical solution for a nonzero dispersivity. Curve C illustrates the breakthrough curve calculated for the same conditions as B, but using a numerical method that introduces numerical dispersion. Major differences exist between the analytical solution (B) and the numerical solution (C) in parts of the domain. Therefore, care must be taken to assess and minimize such numerical errors that would artificially add “numerical” spreading or mixing to the calculated dispersion attributable to physical and chemical processes.

Numerical dispersion can be controlled by reducing the grid spacing (Δx , Δy , and Δz). However, reduction to a tolerable level may require an excessive number of nodes and render the computational costs unacceptably high. It may also be controlled in finite-element methods by using higher order basis functions or by adjusting the formulation of the equations (using different combinations of forward, backward, or centered in time and (or) space, or using different weighting functions). Unfortunately, many approaches that eliminate or minimize numerical dispersion introduce oscillatory behavior, causing overshoot behind a moving front and possibly undershoot ahead of the front (see curve D in Figure 23.9),

and vice versa. Undershoot can result in the calculation of negative concentrations, which are obviously unrealistic. Overshoot can introduce errors of equal magnitude that may go unnoticed because the value is positive in sign (although greater than the source concentration, so still unrealistic). Oscillations generally do not introduce any mass-balance errors, and often dampen out over simulation time. However, in some cases, oscillatory behavior can become unbounded, yielding an unstable solution or failure to converge numerically.

In solving the advective-dispersive transport equation, some numerical errors (mainly oscillations) can be related to two dimensionless parameter groups (or numbers). One is the Peclet number, P_e , which may be defined as $P_e = \Delta l/\alpha$, where Δl is a characteristic nodal spacing (although it should be noted that there are several alternative, though essentially equivalent, ways to define P_e). Anderson and Woessner (1992) recommend that the grid be designed so that $\Delta l < 4\alpha$ (or $P_e < 4$); Ségol (1994) recommends that the nodal spacing be adjusted so that $P_e \leq 2$. Similarly, time discretization can be related to the Courant number, C_o , which may be defined as $C_o = V\Delta t/\Delta l$ (Anderson and Woessner, 1992). Anderson and Woessner (1992) recommend that time steps be specified so that $\Delta t < \Delta l/V$ (or $C_o < 1.0$), which is equivalent to requiring that no solute be displaced by advection more than the distance across one grid cell or element during one time increment.

In finite-difference transport models, there may also be a grid-orientation effect in which the solute distribution, calculated for the same properties and boundary conditions, will vary depending on the angle of the flow relative to the grid. This effect is largely related to the cross-product terms in the governing equation, and generally is not a serious source of error, but the model user should be aware of this error.

23.6.4 Mass Balance

One measure of numerical accuracy is how well the model conserves mass. This can be measured by comparing the net fluxes calculated or specified in the model (e.g., inflow and sources minus outflow and sinks) with changes in storage (accumulation or depletion). Mass-balance calculations should always be performed and checked during the calibration procedure to help assess the numerical accuracy of the solution. For direct matrix equation solution methods, the mass-balance error should always be negligible. Mass-balance errors can be introduced by an inaccurate iterative solution of the matrix equations for flow or solute transport.

As part of these calculations, the hydraulic and chemical fluxes contributed by each distinct hydrologic component of the flow and transport model should be itemized separately to form hydrologic and chemical budgets for the system being modeled. The budgets are valuable assessment tools because they provide a measure of the relative importance of each component to the total budget.

Errors in the mass balance for flow models should generally be less than 0.1%. However, because the solute-transport equation is more difficult to solve numerically, the acceptable mass-balance error for a solute may be greater than for the fluid, but this will depend also on the nature of the numerical method implemented. Finite-difference and finite-element methods are inherently mass conservative, whereas some implementations of the method of characteristics and particle-tracking approaches may not be (or their mass-balance calculations themselves are only approximations). Also, whereas a large mass-balance error provides evidence of a poor convergence of the iterative solution, a perfect mass balance in itself does not and cannot prove that a true or accurate solution has been achieved or that the overall model is valid. That is, a perfect mass balance can be achieved if the model includes compensating errors. For example, the solutions C and D in Figure 23.9 that exhibit substantial numerical dispersion or oscillatory behavior arise from solutions that show a near-perfect mass balance, but they are still wrong.

23.6.5 Sensitivity Tests

Systematically varying values for given parameters helps to achieve another objective of the calibration procedure, namely to determine the sensitivity of the model to factors that affect groundwater flow and transport and to errors and uncertainty in the data. Evaluating the relative importance of each factor helps

determine which data must be defined most accurately and which data are already adequate or require only minimal further definition. If additional field data can be collected, such a sensitivity analysis helps in deciding which types of data are most critical and how to get the best information return on the costs of additional data collection. If additional data cannot be collected, then the sensitivity tests can help assess the reliability of the model by demonstrating the effect of a given range of uncertainty or error in the input data on the model output. The relative sensitivities of the parameters that affect flow and transport will vary from problem to problem. Furthermore, the sensitivities may change over time as the stress regime imposed on a system evolves. Therefore, a sensitivity analysis should be performed during the early stages of a model study.

The sensitivity of the solution to the spatial grid design (or spacing), time-step size, nature and placement of boundary conditions, and other numerical parameters also should be evaluated. This step is frequently overlooked, but failure to do so may cause critical design flaws to remain undetected. For example, parameter-estimation models cannot evaluate the sensitivity to grid spacing or certain boundary conditions that are fixed in the model by the user. It is generally recommended that after a preliminary calibration has been achieved, the model should be rerun for the same stresses and properties using a finer grid, smaller time steps, and perhaps alternative boundary conditions. If such a test yields significantly different results, then the model should be recalibrated using design criteria that yield a more accurate numerical solution. If such a test yields no significant differences, then the coarser design is probably adequate for that particular problem. For example, Bower et al. (2005) found that grid resolution was a critical issue for simulating groundwater flow and solute transport at Yucca Mountain, Nevada. The use of a GIS-based GUI with automatic regridding capability can greatly facilitate the assessment of grid sensitivity.

23.6.6 Calibration Criteria

Model calibration may be viewed as an evolutionary process in which successive adjustments and modifications to the model are based on the results of previous simulations. The modeler must decide when sufficient adjustments have been made to the representation of parameters and processes and at some time accept the model as being adequately calibrated (or perhaps reject the model as being inadequate and seek alternative approaches). This decision is often based on a mix of subjective and objective criteria. The achievement of a best fit between values of observed and computed variables is a regression procedure and can be evaluated as such. That is, the residual errors should have a mean that approaches zero and the deviations should be minimized. Cooley (1977) discusses several statistical measures that can be used to assess the reliability and “goodness of fit” of groundwater-flow models. The accuracy tests should be applied to as many dependent variables as possible. The types of observed data that are most valuable for model calibration include head and concentration changes over space and time, and the quantity and quality of groundwater discharges from the aquifer.

Although it is necessary to evaluate the accuracy of the model quantitatively, it is equally important to ensure that the dependent variables that serve as a basis for the accuracy tests are reliable indicators of model accuracy. For example, if a particular dependent variable was relatively insensitive to the governing parameters, then the presence of a high correlation between its observed and computed values would not necessarily be a reflection of a high level of accuracy in the overall model. Head values are the most common basis for model calibration, but water-level (head) measurements are subject to error, even if precisely measured. If the well is a long open borehole or has a long screen, then the measured water level represents some integrated mean head, and if the vertical discretization of the model does not correspond closely to the sampled interval, forcing a match may induce a bias in the model parameters. In addition, if the water levels were measured at various points in time, but the model is developed to simulate a long-term steady-state condition, then even if the water levels had been measured accurately, there may be a disconnect between the data and the supposedly corresponding model-computed values.

Similarly, caution must be exercised when the “observed data” contain an element of subjective interpretation. For example, matching an observed potentiometric surface or concentration distribution is

sometimes used as a basis for calibrating groundwater models. However, a contoured surface is itself interpretive and can be a weak basis for model calibration because it includes a variability or error introduced by the contouring process, in addition to measurement errors present in the observed data at the specific points.

23.6.7 Predictions and Postaudits

As model calibration and parameter estimation are keyed to a set of historical data, the confidence in and reliability of the calibration process is proportional to the quality and comprehensiveness of the historical record. The time over which predictions are made with a calibrated model should also be related to, and limited by, the length of the historical record. A reasonable guideline is to predict only for a period of time comparable in length to the period that was matched.

The accuracy of a model's predictions is the best measure of its reliability. Predictive accuracy, however, can be evaluated only after the fact. Anderson and Woessner (1992) summarize several published studies in which the predictive accuracy of a deterministic groundwater model was evaluated several years after the prediction had been made. The results suggest that extrapolations into the future were rarely accurate. Predictive errors often were related to having used a time period for history matching that was too short to capture an important element of the model or of the groundwater system, or to having an incomplete conceptual model. For example, processes and boundary conditions that are negligible or insignificant under the past and present stress regime may become important under a different set of imposed stresses. Thus, a conceptual model founded on observed behavior of a groundwater system may prove to be inadequate in the future, when present stresses are increased or new stresses are added. A major source of predictive error is sometimes attributable primarily to the uncertainty of future stresses, which is often affected by demographic, political, economic, and (or) social factors. If the range or probability of future stresses can be estimated, then the range or probability of future responses can be predicted. When using a model to predict future responses in a groundwater system, it is important to place confidence bounds on the predictions arising out of the uncertainty in parameter estimates. However, these confidence limits still would not bound errors arising from the selection of a wrong conceptual model or from problems in the numerical solution algorithms (Bredehoeft and Konikow 1993).

If a model is to be used for prediction relating to a problem or system that is of continuing interest or importance to society, then field monitoring should continue and the model should be periodically postaudited, or recalibrated, to incorporate new information, such as changes in imposed stresses or revisions in the assumed conceptual model. A postaudit offers a means to evaluate the nature and magnitude of predictive errors, which may itself lead to a large increase in the understanding of the system and in the value of a subsequently revised model. Revised predictions can then be made with greater reliability.

23.6.8 Model Validation

It is natural for people who apply groundwater models, as well as those who make decisions based on model results, to want assurance that the model is valid. Groundwater models are embodiments of various scientific theories and hypotheses. Popper (1959) argues that "*as scientists we can never validate a hypothesis, only invalidate it.*" The same philosophy has been applied specifically to groundwater models (Konikow and Bredehoeft 1992; Oreskes et al. 1994).

The criteria for labeling a model as validated are inherently subjective. In practice, validation is attempted through the same process that is typically and more correctly identified as calibration, that is, by comparing calculations with field or laboratory measurements. However, the non-uniqueness of model solutions means that a good comparison can be achieved with an inadequate or erroneous model. Also, because the definition of "good" is subjective, under the common operational definitions of validation, one competent and reasonable scientist may declare a model as validated, whereas another may use the same data to demonstrate that the model is invalid. To the general public, proclaiming that a groundwater model is validated carries with it an aura of correctness that many modelers would not claim (Bredehoeft

and Konikow 1993). Because labeling a model as having been “validated” has little objective or scientific meaning, such “certification” does little beyond instilling a false sense of confidence in such models. Konikow and Bredehoeft (1992) recommend that the term validation not be applied to groundwater models.

23.7 Overview of Representative Generic Models

A large number and variety of generic groundwater models are documented and available at the present time. Two widely used public domain models are explained in more detail as illustrative examples.

23.7.1 MODFLOW-2000

One of the most popular and comprehensive deterministic groundwater models available today is the United States Geological Survey’s (USGS) MODFLOW code (McDonald and Harbaugh 1988; Harbaugh and McDonald 1996; Harbaugh et al. 2000; Harbaugh 2005). This model is actually an integrated family of compatible codes that centers on an implicit finite-difference solution to the 3D flow equation. The model was coded in a modular style to allow and encourage the development of additional packages or modules that can be added on or linked to the original code. MODFLOW-2000 (Harbaugh et al. 2000) is a major revision of the previous MODFLOW computer programs (McDonald and Harbaugh 1988; Harbaugh and McDonald 1996). MODFLOW-2000 is organized by “Processes,” which currently (2005) include the Ground-Water Flow, Observation, Sensitivity, Parameter-Estimation, Ground-Water Management, and Ground-Water Transport Processes.

The Ground-Water Flow Process uses a block-centered finite-difference grid that allows variable spacing of the grid in three dimensions. Flow can be steady or transient. Layers can be simulated as confined or convertible between confined and unconfined. Aquifer properties can vary spatially and hydraulic conductivity (or transmissivity) can be anisotropic. Flow associated with external stresses, such as wells, areally distributed recharge, evapotranspiration, drains, lakes, and streams, can also be simulated through the use of specified head, specified flux, or head-dependent flux boundary conditions. The implicit finite-difference equations can be solved using one of several available solver packages. Although the input and output systems of the program were designed to permit maximum flexibility, usability and ease of interpretation of model results can be enhanced by using one of several commercially available preprocessing and postprocessing packages; some of these packages operate independently of MODFLOW, whereas others are directly integrated into reprogrammed and (or) recompiled versions of the MODFLOW code.

MODPATH (Pollock 1989, 1994) is a particle-tracking postprocessor for the MODFLOW Ground-Water Flow Process that calculates flowpaths and traveltimes of water particles through a simulated groundwater system for steady-state or transient conditions. Particles can be tracked in the direction of groundwater flow (forward tracking) or backwards towards points of recharge (backward tracking). A companion program, MODPATH-PLOT, provides a graphics interface package to visually display the results of a MODPATH simulation. A semi-analytical method is used in MODPATH to determine each particle’s flowpath within a grid cell. The method is based on heads calculated by MODFLOW, from which the volume flow rate and average linear velocity of groundwater are calculated across each of the six faces of a grid cell. An analytical expression is then obtained on the basis of the cell-face velocities to compute the 3D movement of a particle within each grid cell, given the initial particle position anywhere within the cell. A particle’s complete flowpath through the model domain is obtained by tracking particles from one cell to the next until the particle reaches a model boundary, an internal sink or source of water, or a termination site (or zone) specified by the user. The development of MODPATH in the late 1980s coincided with the initiation of many source-water and wellhead-protection programs in the United States designed to protect water supplies from contamination. Partly because of these programs, MODPATH has been used widely to determine the sources of water to wells, aquifers, and other features represented

in MODFLOW simulations, as well as to assist hydrologists and water managers understand and visualize water flow through complex groundwater systems.

The Observation, Sensitivity, and Parameter-Estimation Processes are all documented in Hill et al. (2000). The Observation Process generates model-calculated values for comparison with measured, or observed, quantities. The Sensitivity Process calculates the sensitivity of hydraulic heads throughout the model with respect to specified parameters. The Parameter-Estimation Process uses a modified Gauss-Newton method to adjust values of user-selected input parameters in an iterative procedure to minimize residuals. Parameters are defined in the Ground-Water Flow Process input files and can be used to calculate most model inputs such as horizontal hydraulic conductivity, horizontal anisotropy, vertical hydraulic conductivity or vertical anisotropy, specific storage, and specific yield. In addition, parameters can be defined to calculate the hydraulic conductance of the River, General-Head Boundary, and Drain Packages; areal recharge rates of the Recharge Package; maximum evapotranspiration of the Evapotranspiration Package; pumpage or the rate of flow at defined-flux boundaries of the Well Package; and the hydraulic head at constant-head boundaries.

Groundwater models often are developed to address specific questions (or problems) that arise in the management of groundwater systems. For example: What is the maximum amount of water that can be pumped from a particular aquifer without causing unacceptable groundwater-level declines or reductions in streamflow? What is the least-cost pumping strategy to capture or contain a contaminant plume? What is the best strategy to conjunctively use the groundwater and surface-water resources of a basin? During the past two decades, much effort has been directed toward developing computer models that can simultaneously account for the physical processes that affect flow and transport in a groundwater system and the large number of engineering, legal, and economic factors that can affect the development and management of groundwater resources. Such models couple the capabilities of groundwater simulation with mathematical optimization techniques (such as linear and nonlinear programming) that solve properly formulated management problems. An optimization problem (or formulation) is a model of the management decision-making process, and has three primary components — an objective function, a set of constraints, and a set of decision variables (Ahlfeld and Mulligan 2000). The decision variables of a management problem are the quantities that are determined by solution of the management problem — for example, the rates, locations, and timing of groundwater withdrawals from a set of available pumping wells. As implied by its name, the objective function defines the objective of the management problem — such as “maximize total pumping from the set of pumping wells.” The constraints of the problem limit the values that can be taken by the decision variables in the solution of the problem.

Several codes have been developed for groundwater simulation-optimization modeling (Lefkoff and Gorelick 1987; Greenwald 1998; Zheng and Wang 2002; Ahlfeld and Riefler 2003; Peralta 2004; Ahlfeld et al. 2005). These codes differ in (1) the numerical model that is used to represent the groundwater-flow system, (2) the types of groundwater-management problems that can be solved, and (3) the approaches that are used to mathematically solve the management problems. Recently, the Ground-Water Management (GWM) Process has been developed for MODFLOW-2000 (Ahlfeld et al. 2005). A response-matrix method is used in the GWM Process to solve various types of linear, nonlinear, and mixed-binary linear groundwater management-formulations. The method is based on the determination of functional relations (called response coefficients) between stresses imposed at the decision-variable locations and the resulting changes in state variables (heads or streamflows) at the constraint locations; these functional relations are calculated by the MODFLOW Ground-Water Flow Process. Decision variables that can be specified in GWM include withdrawal and injection wells, artificial-recharge basins, and imports and exports of water. The types of constraints that can be specified include upper and lower bounds on withdrawal and injection rates, water-supply demands, hydraulic-head constraints such as drawdowns and hydraulic gradients, and streamflow and streamflow-depletion constraints.

A variety of other MODFLOW accessory codes, packages, and features are available. Most of these were developed by the USGS and reports documenting them are available with the latest versions of MODFLOW-2000 and related models. Examples include coupled surface-water and groundwater flow, aquifer compaction, transient leakage from confining units, rewetting of dry cells, horizontal flow barriers,

alternative interblock transmissivity computations, cylindrical flow to a well, a statistical processor, a data input program, multi-node wells, and a program that calculates water budgets. The Ground-Water Transport Process (MODFLOW-GWT) includes the capability to solve the solute-transport equation using several optional solution methods, including the method of characteristics (Konikow et al. 1996).

Other packages have been developed by non-USGS sources to work with MODFLOW; one example is the advective-dispersive solute transport model MT3DMS (Zheng and Wang 1999). This model solves the solute-transport equation for one or more solutes on the basis of a flow field calculated using MODFLOW. MT3DMS offers several different numerical methods to solve the transport equation, including an advanced finite-difference method to reduce numerical dispersion. Langevin et al. (2003) have integrated MT3DMS directly into a modified version of MODFLOW-2000, which is called the Variable-Density Flow Process (VDF). In the VDF Process, the governing flow equation is modified for application to variable-density problems (Equation 23.6, but developed in terms of equivalent freshwater head rather than pressure), and the resulting code is called SEAWAT-2000. A comprehensive watershed simulation system based on MODFLOW has been documented by Panday and Huyakorn (2004). This model, called MODHMS, includes a 3D saturated-unsaturated flow equation for the subsurface, coupled with the diffusion wave equation for overland flow, both of which are coupled with the diffusion wave equation for flow through a network of streams, channels, and hydraulic structures (Panday and Huyakorn 2004).

23.7.2 SUTRA

The USGS's SUTRA code is one of the most widely used simulators in the world for seawater intrusion and other variable-density groundwater-flow problems based on solute transport or heat transport. Since its initial release in 1984, SUTRA (acronym for Saturated-Unsaturated TRAnsport) (Voss 1984; Voss and Provost 2002) has also been widely used for many other types of groundwater problems. The code is routinely used for hydrogeologic analyses and for many practical engineering studies.

The SUTRA code offers the means to numerically simulate a variety of subsurface physical processes and situations. It robustly and reliably provides simulation results for well-posed problems. Although the SUTRA code can be used to simulate saturated-unsaturated systems with subsurface energy transport or single-species reactive solute transport, this discussion focuses only on seawater intrusion. Application to seawater-intrusion problems is based on the capability of SUTRA to simulate saturated variable-density flow with non-reactive solute transport.

SUTRA is a finite-element code designed to solve a general set of single-phase subsurface fluid flow and single-species transport problems. The code was written to solve the basic equations presented by Bear (1979), which cover most types of groundwater flow and transport physics known. The two equations solved in 2D or 3D are a fluid mass balance for unsaturated and saturated groundwater flow, and a unified energy and solute mass balance. The unified balance provides a single equation that describes either solute or energy transport simply by changing the definition of parameters in the equation.

The coding style of SUTRA is modular with each routine serving a special function, and coding simplicity is stressed over code efficiency to allow users to make easy changes and additions of processes to the code. A robust direct solver was selected to avoid the convergence difficulties associated with iterative matrix solvers, despite that such a solver would be less efficient for problems with large meshes. The code is written in standard FORTRAN.

Recent developments for SUTRA include a complete GUI that provides a geographical information system (GIS) type of functionality together with automatic mesh generation (Voss et al. 1997; Winston and Voss 2003). The SUTRA GUI consists of public-domain USGS codes that connect SUTRA with the commercial Argus ONE™ software package that provides meshing and GIS capabilities. This GUI provides SUTRA users with new analytic power, making it easy to set up, run, and view results for even complex 3D SUTRA models.

Computer speed and size have been the only limiting factors for practical application since the code was released, and will likely continue to be so. Computer speed and size have increased substantially since SUTRA's original 1984 release, allowing more complex 2D and 3D simulations to be done regularly on a

personal computer. The numerical effort for a simulation (i.e., computer time and memory) is determined by the density of spatial discretization and the number of time steps simulated.

The numerical technique employed by the SUTRA code to solve the governing equations uses a modified 2D and 3D Galerkin finite-element method with bi- and tri-linear quadrilateral and hexahedral elements. Solution of the equations in the time domain is accomplished by the implicit finite-difference method. Modifications to the standard Galerkin method that are implemented in SUTRA are as follows: All terms of the equations not comprised of spatial gradients (e.g., time derivatives and sources) are assumed to be constant in the region surrounding each node (cell) in a manner similar to integrated finite differences. Parameters associated with the non-flux terms are thus specified nodewise, whereas parameters associated with flux terms are specified elementwise. This achieves some efficiency in numerical calculations while preserving the accuracy, flexibility, and robustness of the Galerkin finite-element technique. A very useful ad hoc generalization of the dispersion process is included in SUTRA, which allows the longitudinal and transverse dispersivities to vary in a time-dependent manner at any point *depending on the direction of flow* (as discussed above). Voss and Provost (2002) give a complete description of these numerical methods as used in the SUTRA code.

An important modification to the standard finite-element method that is required for variable-density flow simulation is implemented in SUTRA. This modification provides a velocity calculation within each finite element based on *consistent* spatial variability of pressure gradient, ∇p , and the buoyancy term, ρg , in Darcy's law. Without this "consistent velocity" calculation, the standard method generates spurious vertical velocities everywhere there is a vertical gradient of concentration within a finite-element mesh, even with a hydrostatic pressure distribution (Voss 1984; Voss and Souza 1987). The spurious velocities make it impossible to simulate, for example, a narrow transition between freshwater and seawater with the standard method, irrespective of how small a dispersivity is specified for the groundwater system.

The two governing equations, fluid mass balance and solute mass (or energy) balance, are solved sequentially for each iteration or time step. Iteration is carried out by the Picard method with linear half-time-step projection of nonlinear coefficients on the first iteration of each time step. Iteration to the solution for each time step is optional. Velocities required for solution of the transport equation are the result of the flow equation solution (i.e., pressures) from the previous iteration (or time step for non-iterative solution).

SUTRA is coded in a modular style, making it convenient for sophisticated users to modify the code (e.g., replace the existing direct banded-matrix solver) or to add new processes. Addition of new terms (i.e., new processes) to the governing equations is a straightforward process usually requiring changes to the code in very few lines. Also to provide maximum flexibility in applying the code, all boundary conditions and sources are allowed to be time-dependent in any user-specified manner.

23.8 Case Histories

A large number of documented examples of the application of groundwater models to a variety of hydrogeologic problems are available in the literature. Three case studies have been selected to help illustrate modeling philosophy and practice, including aspects of model conceptualization, model implementation, and interpretation of results.

23.8.1 Flow and Transport in a Shallow Unconfined Aquifer

Reilly et al. (1994) combined the application of environmental tracers and deterministic numerical modeling to analyze and estimate recharge rates, flow rates, flow paths, and mixing properties of a shallow groundwater system near Locust Grove in eastern Maryland. The study was undertaken as part of the USGS's National Water-Quality Assessment Program to provide flow paths and traveltimes estimates to be used in understanding and interpreting water-quality trends in monitoring wells and stream base flows. The study area encompassed about $2.6 \times 10^7 \text{ m}^2$ of mostly agricultural land on the Delmarva Peninsula.

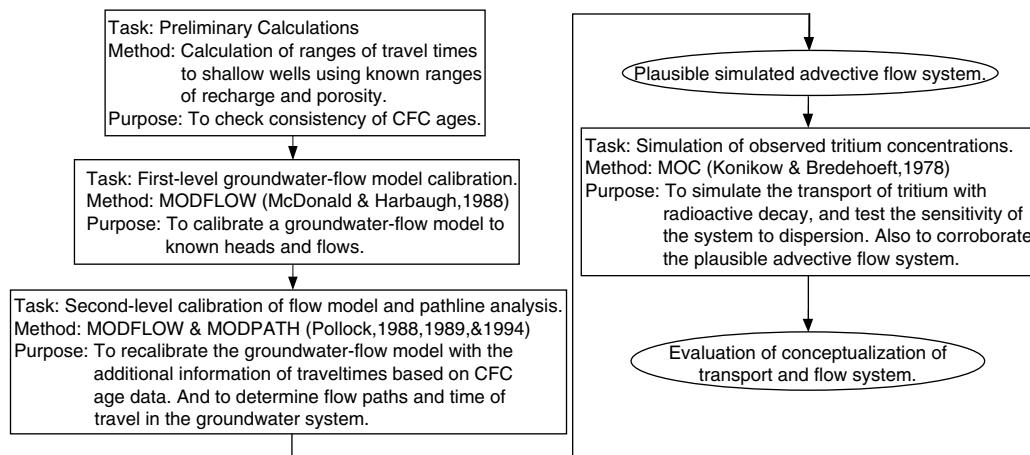


FIGURE 23.10 Flow diagram of the steps taken to quantify the flow paths in the Locust Grove, Maryland, groundwater-flow system. (Adapted from Reilly, T.E., Plummer, L.N., Phillips, P.J. and Busenberg, E. (1994) *Water Resour. Res.*, 30: 421–433.)

The surficial aquifer includes unconsolidated permeable sands and gravel that range in thickness from less than 6 m to more than 20 m. This surficial aquifer is underlain by relatively impermeable silt and clay deposits, which form a confining unit.

In this study, chlorofluorocarbons (CFCs) and tritium were analyzed from a number of water samples collected from observation wells to estimate the age of groundwater at each sampling location and depth. Because inaccuracies and uncertainty are associated with estimates of age based on environmental tracers, just as inaccuracies and uncertainty are associated with deterministic models of groundwater flow and transport, the authors applied a feedback or iterative process based on comparisons of independent estimates of traveltime. Their approach is summarized and outlined in Figure 23.10. Each task shown was designed to improve either the estimates of parameters or the conceptualization of the groundwater system.

The preliminary calculations (first task) were used to set bounds on the plausibility of the results of the more complex simulations and chemical analyses. The first-level calibration of a groundwater-flow model (second task) provided the initial system conceptualization. The third task was a second-level calibration and analysis involving simulation of advective transport, which provided quantitative estimates of flow paths and time of travel to compare with those obtained from the CFC analyses. The fourth task involved the application of a solute-transport model to simulate tritium concentrations in the groundwater-flow system as affected by the processes of advection, dispersion, radioactive decay, and time-varying input (source concentration) functions.

The sampling wells were located approximately along an areal flow line, and a 2D cross-sectional model was developed for the simulation of processes occurring along this flow line. The MODFLOW model (McDonald and Harbaugh 1988) was used to simulate groundwater flow and advective transport. The finite-difference grid consisted of 24 layers and 48 columns of nodes, with each cell having dimensions of 1.14 by 50.80 m, as shown in Figure 23.11, which also shows the wells that lie in the cross section. The simulation was designed to represent average steady-state flow conditions.

After the flow model was calibrated, 2D pathline and traveltime analyses were undertaken and comparisons to CFC age estimates were made. The pathlines calculated using MODPATH (Pollock 1989) after the second-level calibration with MODFLOW are shown in Figure 23.12. The comparison with CFC estimates was generally good. However, Reilly et al. (1994) note that close to the stream, many flow lines converge, and the convergence of pathlines representing the entire range of traveltimes present in the aquifer causes waters of different ages to be relatively near each other. Thus, at the scale and grid spacing of the model, in the area near the stream the convergent flow lines cannot be readily differentiated in the

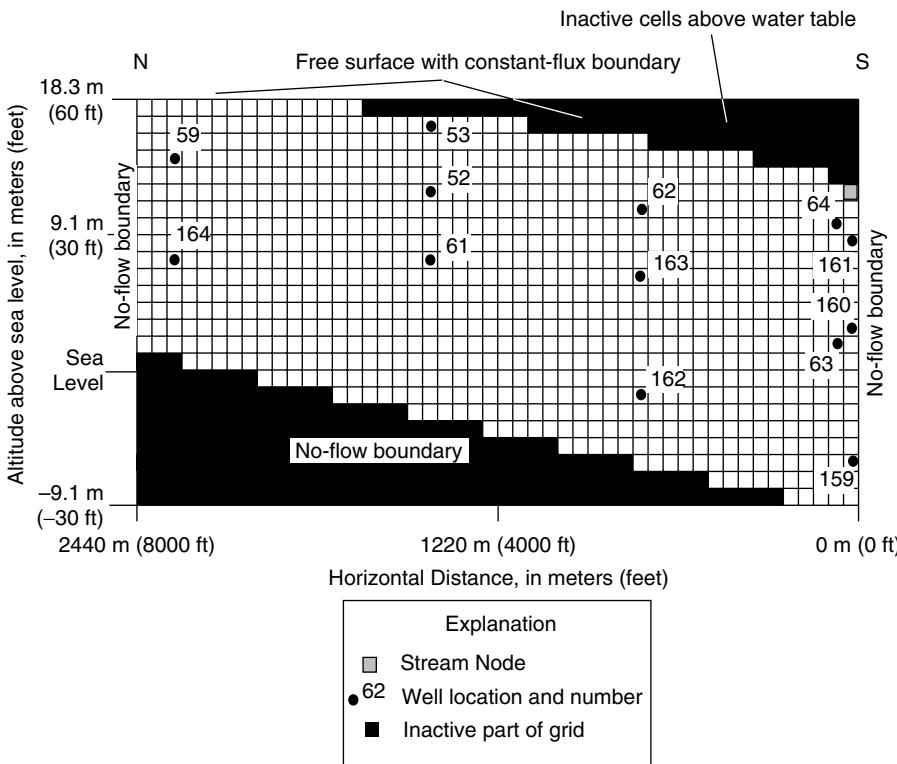


FIGURE 23.11 Model grid used to simulate Locust Grove cross section, showing well locations. (Adapted from Reilly, T.E., Plummer, L.N., Phillips, P.J. and Busenberg, E. (1994) *Water Resour. Res.*, 30: 421–433.)

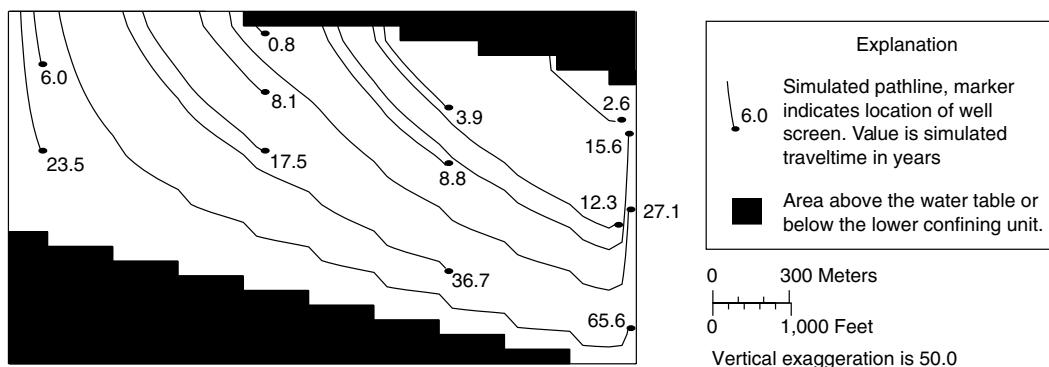


FIGURE 23.12 Pathlines (calculated using MODPATH after second-level calibration) in Locust Grove cross section to observation wells showing time of travel (in years) from the water table. (Adapted from Reilly, T.E., Plummer, L.N., Phillips, P.J. and Busenberg, E. (1994) *Water Resour. Res.*, 30: 421–433.)

model and the ages at individual well screens cannot be accurately represented directly under the stream. After the second-level calibration, the root mean squared error between the simulated ages and the CFC ages for the ten wells farthest from the stream (i.e., excluding wells 159, 160, and 161) was 3.4 years.

Tritium concentrations of recharge waters have varied considerably over the last 40 years. Thus, the time of travel would not always be readily apparent from the tritium concentration in a water sample. Also, mixing of waters recharged during periods of these relatively sharp changes of input concentrations

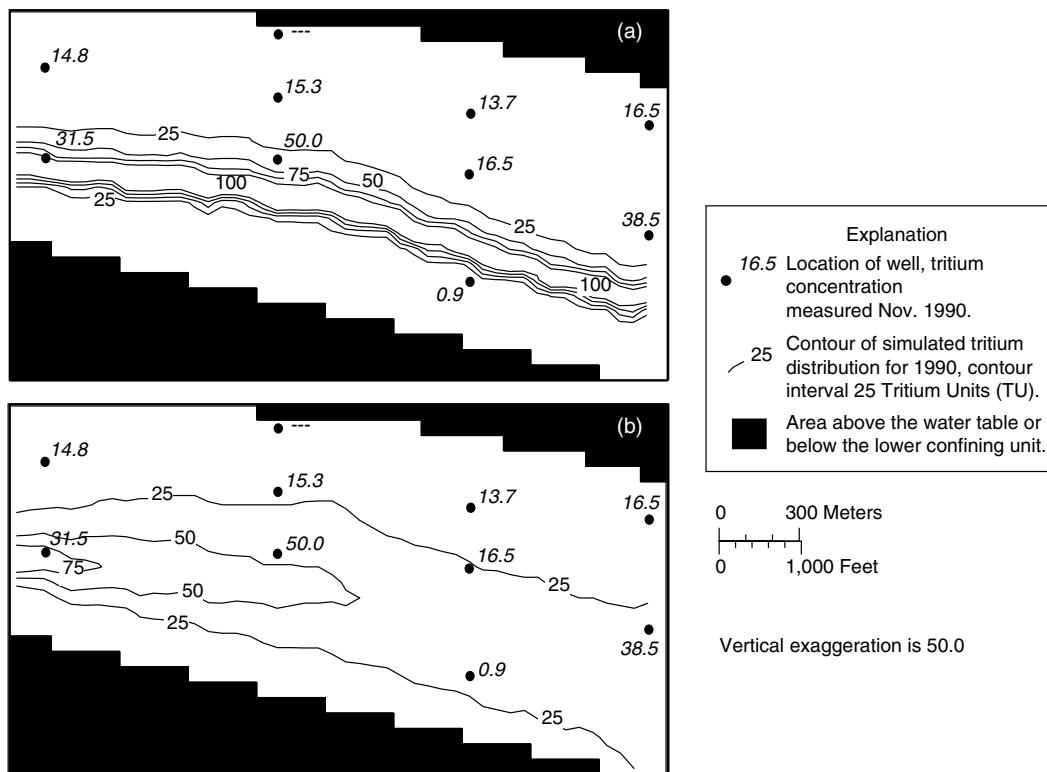


FIGURE 23.13 Simulated tritium distribution at the end of 1990: (a) with dispersivity $\alpha_L = 0.0 \text{ m}$ and $\alpha_T = 0.0 \text{ m}$, and (b) with dispersivity $\alpha_L = 0.15 \text{ m}$ and $\alpha_T = 0.015 \text{ m}$. Contour interval 25 tritium units (TU). (Adapted from Reilly, T.E., Plummer, L.N., Phillips, P.J. and Busenberg, E. (1994) *Water Resour. Res.*, 30: 421–433.)

can make the interpretation of time of travel from tritium concentrations even more uncertain. Thus, the investigators simulated solute transport of tritium within the system using a model that accounts for mixing (dispersion), radioactive decay, and transient input functions, which also allowed a further evaluation of consistency with the results of the previous flow and advective transport model. They applied the MOC solute-transport model of Konikow and Bredehoeft (1978) and Goode and Konikow (1989) for this purpose.

The results of the simulation of the tritium distribution assuming (1) no dispersion and (2) assuming α_L of 0.15 m and α_T of 0.015 m are shown in Figure 23.13. The limiting case simulation of no dispersion yielded acceptable results and was used as the best estimate of the tritium distribution in November 1990 (Reilly et al. 1994). This case reproduces the sharp concentration gradients required to reproduce the low tritium values that were observed. The MOC model was advantageous for this problem because it minimizes numerical dispersion and it can solve the governing equations for α_L of 0.0, which transport models based on finite-difference or finite-element methods often cannot solve. The results of the solute-transport simulation are consistent with the advective flow system determined by the second-level calibration and thus strengthen the case for the conceptual model. The coupling of the tritium analyses and the transport model indicates where discrepancies between the measured and simulated concentrations occur, where additional data collection would be most useful, and where refinement of the conceptual model may be warranted.

The case study described above illustrates that environmental tracers and numerical simulation methods in combination are effective tools that complement each other and provide a means to estimate the flow rate, time of travel, and path of water moving through a groundwater system. Reilly et al. (1994) found

that the environmental tracers and numerical simulation methods also provide a “feedback” that allows a more objective estimate of the uncertainties in the estimated rates and paths of movement. Together, the two methods enabled a coherent explanation of the flow paths and rates of movement while identifying weaknesses in the understanding of the system that require additional data collection and refinement of conceptual models of the groundwater system.

23.8.2 Delineation of Contributing Areas to Water-Supply Wells, Rhode Island

One type of water-resource management issue to which groundwater modeling and particle tracking are often applied is the delineation of contributing areas to water-supply wells. The contributing area of a well is defined as the surface area of the water table (and, possibly, the beds and banks of surface-water bodies) where water entering the groundwater system eventually flows to the well (Franke et al. 1998). The delineation of contributing areas has become an important component of many wellhead-protection programs because the land surface that overlies the contributing area to a well can be identified and protected from activities that may cause groundwater contamination.

Contributing areas to several supply wells in the Hunt River Basin of Rhode Island were determined to identify the sources of water to the wells, the land area that contributes recharge to the wells, and the travel times of recharge to the wells (Barlow and Dickerman 2001). The aquifer from which the wells pump consists of highly permeable, stratified sand-and-gravel sediments deposited by glacial meltwater. The aquifer covers a 19 square-mile area, which is about half the total drainage area of the basin; the remainder of the basin consists of upland areas of primarily till and bedrock that drain to the aquifer (Figure 23.14). The aquifer is unconfined and in direct hydraulic connection with several shallow rivers, streams, and ponds that generally penetrate only a thin portion of the top of the aquifer. However, this connection allows for extensive interaction between the groundwater and surface-water systems, which occurs as groundwater discharge (base flow) to the streams, streamflow leakage to the aquifer from natural stream-channel losses, and induced infiltration of streamflow caused by pumping at the supply wells. This hydraulic interaction was a key factor in the development of the conceptual and numerical models of the groundwater system.

Groundwater flow in the aquifer was simulated using the MODFLOW model (Harbaugh and McDonald 1996). The model grid was aligned approximately parallel to the northeast-trending Hunt River Valley (Figure 23.14). The model domain was discretized into a grid of 205 rows and 197 columns of square cells that were 200 ft on each side. Vertically, the model consists of a maximum of four layers that extend from the water table to the intersection of the aquifer with underlying bedrock. The uppermost layer of the model was relatively thin (a maximum of 10 ft) to simulate shallow groundwater flow near surface-water bodies as accurately as possible. The thickness of the remaining layers was specified partly on the basis of the vertical locations of the screened intervals of supply wells that pump from the aquifer. Areas of the aquifer where saturated thickness was less than 5 ft were made inactive to ensure numerical stability. This criterion resulted in many cells near the contact between the aquifer and adjoining upland areas being made inactive and a modeled area that was smaller than the measured extent of the aquifer (Figure 23.14).

Boundary conditions specified for the model consisted of several variants of the specified flow and head-dependent flow boundary condition types defined in Table 23.1. No-flow boundary conditions were specified along groundwater-flow lines and drainage divides that separate the modeled area from adjacent aquifer areas that were not simulated. Nonzero flow rates were specified to represent (1) groundwater inflow from the upland areas, which was simulated using injection wells located within the active area of the model at the contact between the aquifer and upland areas; (2) pumping from 18 supply wells; and (3) recharge at the water table. Two types of head-dependent flow boundary conditions were simulated — evapotranspiration from the water table and hydraulic interaction between the aquifer and overlying streams, rivers, and ponds. Interactions between groundwater and surface water were simulated by use of the streamflow-routing (STR) package of MODFLOW (Prudic 1989). This package calculates the flow rate (Q_s ; L^3/T) between each simulated stream reach and surrounding model cell by use of the following

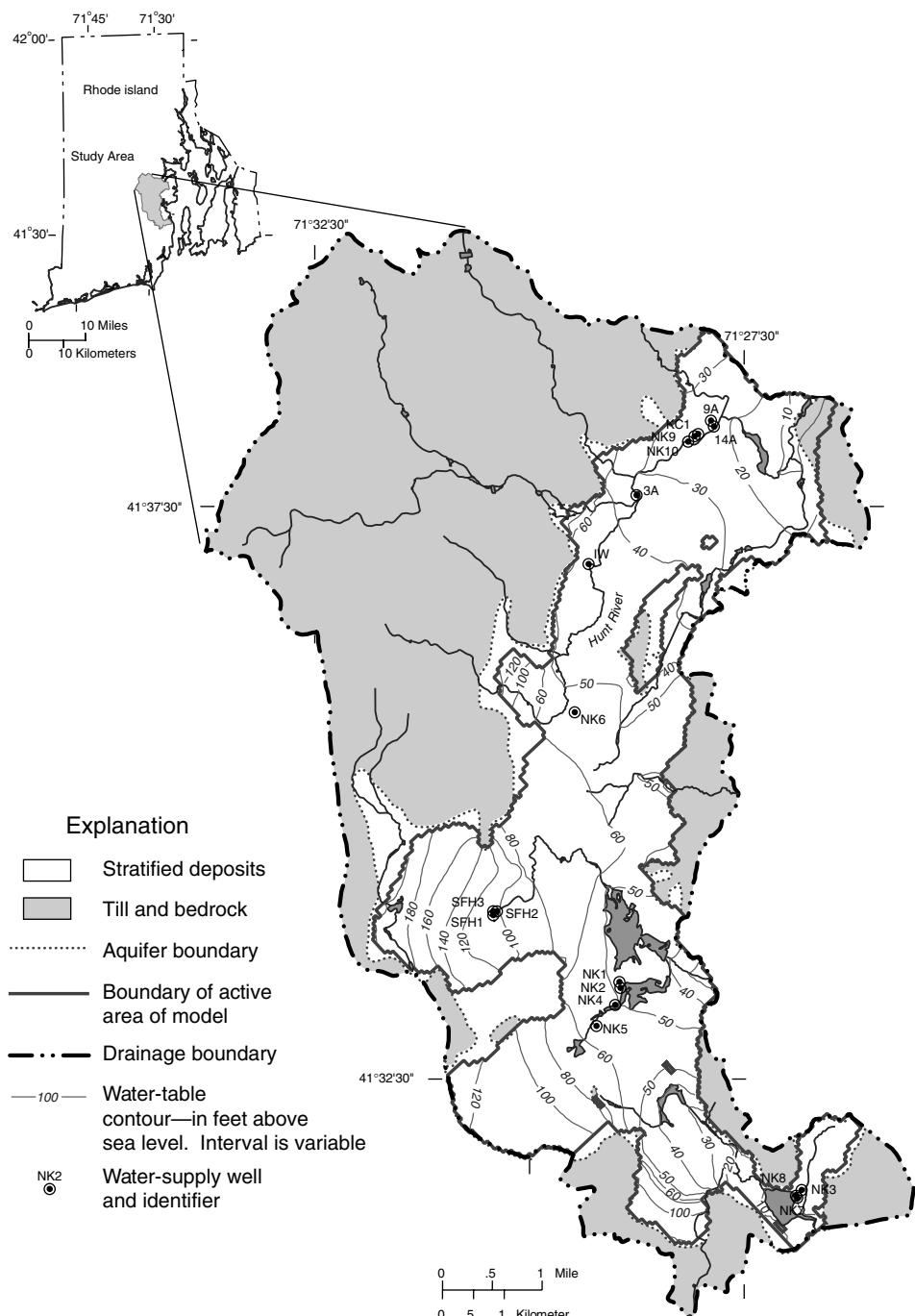


FIGURE 23.14 Model-calculated steady-state water table, Hunt River Basin, Rhode Island. (Modified from Barlow, P.M. and Dickerman, D.C. (2001) Numerical-simulation and conjunctive-management models of the Hunt-Annaquaticket-Pettaquamscutt stream-aquifer system, Rhode Island. U.S. Geol. Survey Prof. Paper 1636.)

equation:

$$Q_s = C_s(H_s - h) \quad (23.29)$$

where C_s is the streambed conductance of the stream reach (L^2/T), H_s is the average stream stage in the reach (L), and h is the groundwater head calculated for the model cell (L). Streambed conductance is a function of the thickness and hydraulic conductivity of the streambed and width and length of the stream reach within each model cell. C_s and H_s were estimated primarily on the basis of field data, including measurements of the hydraulic conductivity of streambed sediments, width of the streams, and stream-stage elevation, but the estimates of C_s were highly uncertain and subsequently modified during model calibration.

The steady-state model was calibrated to groundwater levels measured at 23 observation wells, average annual streamflow measured at a USGS gauging station on the Hunt River, and estimated average annual streamflows for four additional streams in the basin. During the calibration process, initial estimates of several model parameters were modified to improve the match between observed and simulated water levels and streamflow. These modifications included increasing the initial estimate of aquifer recharge by 10% (to 28.0 inches per year), increasing estimates of hydraulic conductivity in several areas, and decreasing the largest estimates of streambed conductance by about one-third.

Contributing areas and sources of water were delineated for 13 supply wells within the study area by use of the MODPATH particle-tracking program (Pollock 1994). Potential sources of water to the wells are recharge to the water table, streamflow leakage, and simulated groundwater inflow from uplands. The contributing area to each well was delineated by overlaying a 2×2 array of particles onto the simulated water table of each grid cell. Particles were then tracked in the forward direction from the water table to their points of discharge. The origin of those particles captured by each well defined the well's contributing area. The amount of streamflow leakage contributing to each well's withdrawal also was determined by tracking particles from losing stream reaches to their point of discharge from the flow system. Twenty-seven particles were distributed uniformly in a 3D array ($3 \times 3 \times 3$) within each losing stream cell. To estimate the amount of streamflow leakage reaching each of the supply wells, a volumetric flow rate was assigned to each particle by dividing the streamflow-leakage rate to the aquifer in the cell in which the particle originated by the number of particles placed in each cell (27). The contribution of streamflow leakage to each supply well was then calculated by summing the individual flow rates of all particles captured by each well.

The contributing areas for the four northernmost wells as determined using MODPATH are shown in Figure 23.15 to illustrate some of the complicated spatial patterns that are often determined for contributing areas, particularly for groundwater systems where several wells are located in close proximity to one another. Each well's contributing area is a function of various hydrogeologic and well-design factors, including the structure and hydraulic properties of the aquifer; the thickness, width, and hydraulic conductivity of streambed sediments near the wells; the withdrawal rate of each well; and the position of the screened interval of each well. This last factor is likely to be one of the primary reasons why the contributing areas for wells NK9, KC1, and 14A extend to the opposite side of the Hunt River from which the wells pump. Also note that the contributing area to well KC1 does not overlie the wellhead. The contributing area to well 9A is completely enveloped by that to well 14A; this results, in part, because the screened interval of well 9A is at a shallower depth than that of well 14A.

Streamflow leakage was determined to be a source of water to 7 of the 13 supply wells (Table 23.2). The largest contribution was calculated for well 3A, which captures an estimated 95% of its discharge from a nearby stream. Streamflow leakage also is a substantial source of water to wells 9A, 14A, and KC1.

The total travel time of each simulated water particle from its entry at the water table to its withdrawal at a supply well also was calculated by MODPATH, and zones of equal travel time to each well are shown in Figure 23.15. Model-calculated travel times for particles captured by the four wells (for a uniform porosity of 0.35) ranged from a minimum of 45 days (to well 14A) to a maximum of 24.7 years (to well NK9).

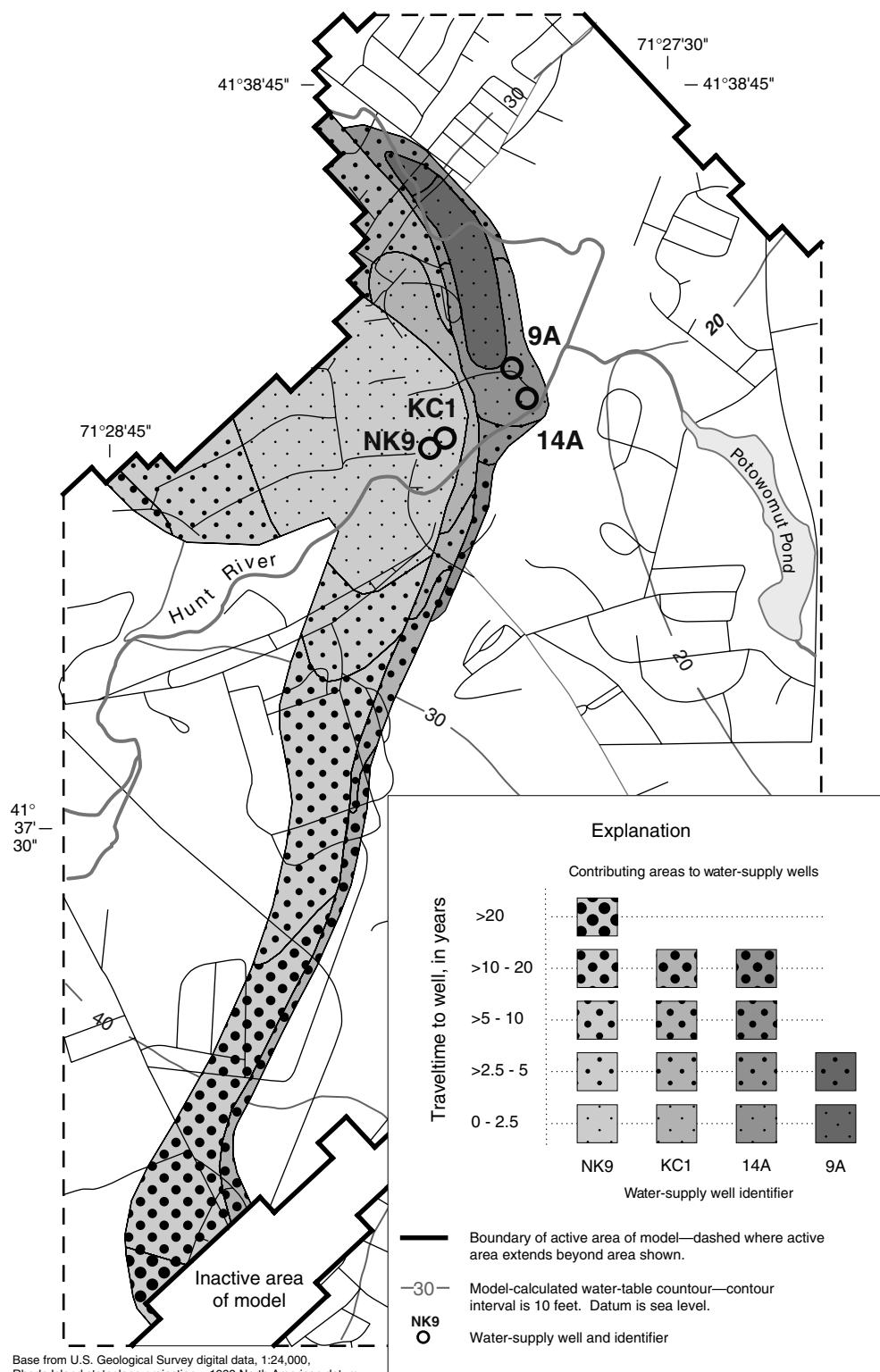


FIGURE 23.15 (See color insert following page 23-52) Model-calculated steady-state contributing areas and traveltimes to water-supply wells in the northern part of the Hunt River Basin, Rhode Island. (Modified from Barlow, P.M. and Dickerman, D.C. (2001) Numerical-simulation and conjunctive-management models of the Hunt-Annaquaticket-Pettaquamscutt stream-aquifer system, Rhode Island. U.S. Geol. Survey Prof. Paper 1636.)

TABLE 23.2 Model-Calculated Sources of Water to Supply Wells in the Hunt River Basin, Rhode Island. Values Shown are Percent of Well's Simulated Withdrawal Rate

Well Identifier	Percent Contribution from Recharge and Uplands	Percent Contribution from Streamflow Leakage
9A	24.1	75.9
14A	24.2	75.8
KC1	27.2	72.8
NK9	83.8	16.2
3A	5.3	94.7
NK6	100.0	0
NK1	100.0	0
NK2	100.0	0
NK4	100.0	0
NK5	100.0	0
SFH1	69.8	30.2
SFH2	100.0	0
NK3	92.6	7.4

Sources: Modified from Barlow, P.M. and Dickerman, D.C. (2001) Numerical-simulation and conjunctive-management models of the Hunt-Annaquatucket-Pettaquamscutt stream-aquifer system, Rhode Island. U.S. Geol. Survey Prof. Paper 1636.

23.8.3 Seawater Intrusion in a Coastal Aquifer

The main population center of the volcanic Hawaiian Island chain in the north-central Pacific Ocean, is on the island of Oahu. Oahu depends on groundwater for its potable water supply, and most of it is produced from coastal aquifers. Oahu has high rainfall, highly permeable aquifers, and a coastal semi-confining layer above many of the aquifers. The coastal confining unit keeps heads high at the coast, and creates a very thick freshwater lens. This combination of high recharge, high permeability, and impeded discharge provides an ample supply of fresh groundwater for both drinking and irrigation. Seawater is present at the base of Oahu's coastal aquifers, and water supplies are potentially threatened by seawater intrusion.

A major issue in Oahu that remains unresolved is how much groundwater may be safely withdrawn from each aquifer. This is an important issue as limited availability of freshwater may potentially limit further economic growth. On the other hand, over-development of groundwater may lead to deleterious impacts on local groundwater supplies from seawater intrusion. Such intrusion may be difficult to reverse quickly. A careful quantitative analysis of aquifer hydraulics and the physics of seawater intrusion will give water managers a reliable scientific basis for managing Oahu's aquifer systems. The primary aquifer of interest underlies the Pearl Harbor area of southern Oahu (Figure 23.16).

The SUTRA code was applied first to evaluate aquifer hydraulics, and then to evaluate seawater intrusion. The objective of the hydraulic analysis was to find the **simplest** possible variable-density flow and transport model that reproduces the measured large-scale hydraulic behavior of the system and to elucidate the major hydrologic controls on the system. SUTRA was modified to include a water table in cross-sectional simulation. The cross-sectional 2D-model domain represents the basal aquifer and caprock containing the basal freshwater lens, the transition zone (TZ) between freshwater and seawater, and deeper seawater to the right of the dike zone (Figure 23.17). A 2D representation is appropriate for evaluation of aquifer-wide response in this system because (1) flow is approximately perpendicular to the coastal boundary; (2) permeability anisotropy is primarily horizontal and vertical; (3) the caprock is parallel to the coast; and (4) sources and sinks are distributed evenly along bands parallel to the coast.

When analyzing water resources in a coastal aquifer system where water quality may be affected by seawater intrusion, relatively low concentrations of solute are critical for water supply. The limiting concentration of potable water is defined here as the equivalent of freshwater containing about 2%

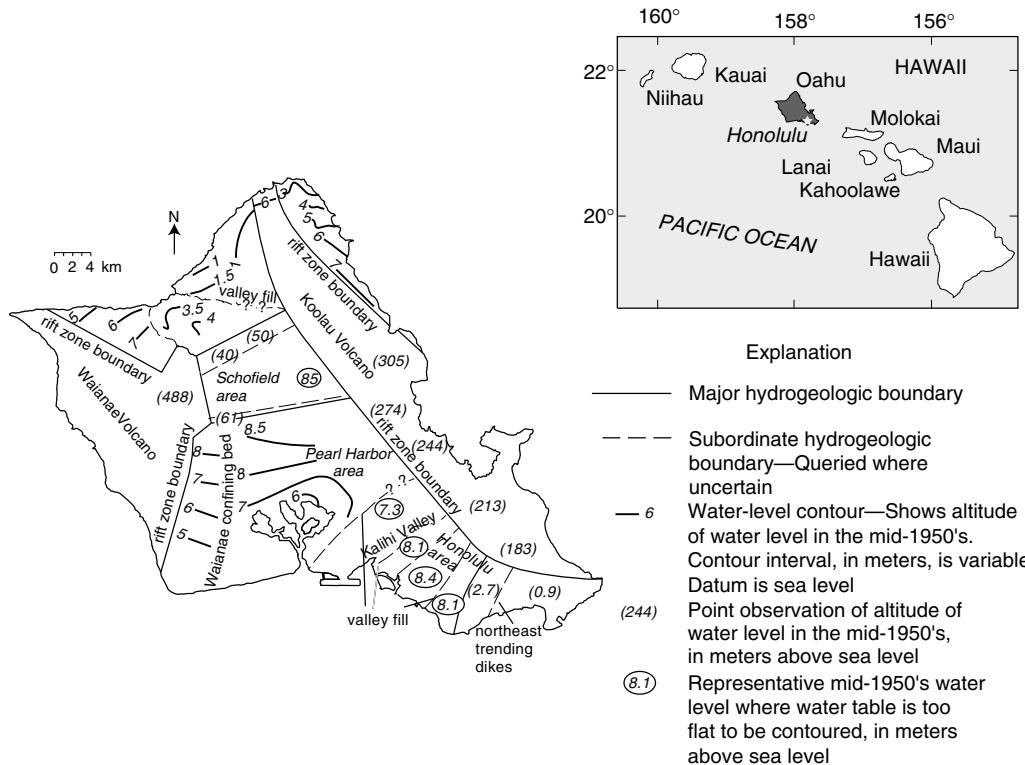


FIGURE 23.16 Location and relative positions of Hawaiian Islands and surface locations of hydrogeologic barriers, water levels in principal aquifers, and inferred freshwater flow directions on Oahu, Hawaii. (From Gingerich, S.B. and Voss, C.I. (2005) *Hydrogeol. J.*, 13: 436–450; after Hunt, C.D., Jr. (1996) Geohydrology of the island of Oahu, Hawaii, U.S. Geol. Surv. Prof. Paper 1412-B.)

seawater (~700 ppm total dissolved solids). Tracking the boundary of potable water as delineated by this relatively low concentration requires use of a numerically accurate solute-transport model with variable fluid density applied using a fine mesh to minimize numerical dispersion.

A 2D cross-sectional model was initially developed for analyzing the TZ and used to predict the general evolution of water quality that will be drawn from the coastal pumping zone. The finite-element mesh (Figure 23.18) encompasses the entire aquifer to a depth of 1800 m below sea level, the deepest extent of permeable basalts. The vertical discretization is only schematically represented in Figure 23.18; the actual discretization in the top 600 m of the mesh is actually 2 m. Recharge in the model-simulated aquifer enters through the water table and through the upstream boundary of the aquifer (representing inflow from dike compartments and the central plateau), both at prescribed rates. The seaward model boundary is held at hydrostatic seawater pressure and the top boundary above the caprock is held at a pressure of zero. The other boundaries are modeled as impermeable (no flow). More details are provided by Souza and Voss (1989) and Voss (1999).

The period 1980–2080 is considered, and the effect of various alternative pumping-recharge scenarios is evaluated (Souza and Voss 1989). Pumping is simulated by withdrawing water from the model area identified as the pumping zone in Figure 23.18, which represents a strip of scattered wells along the coast about 1.5 km (5000 ft) wide. In a practical sense, a regional-scale model cannot be used to predict the salinity of the water pumped from an individual well in the zone. Because of the large scale of processes modeled, only some average concentration in the zone may be considered when analyzing the general salinity changes expected. Thus, meaningful salinity values of the pumped water are the mean

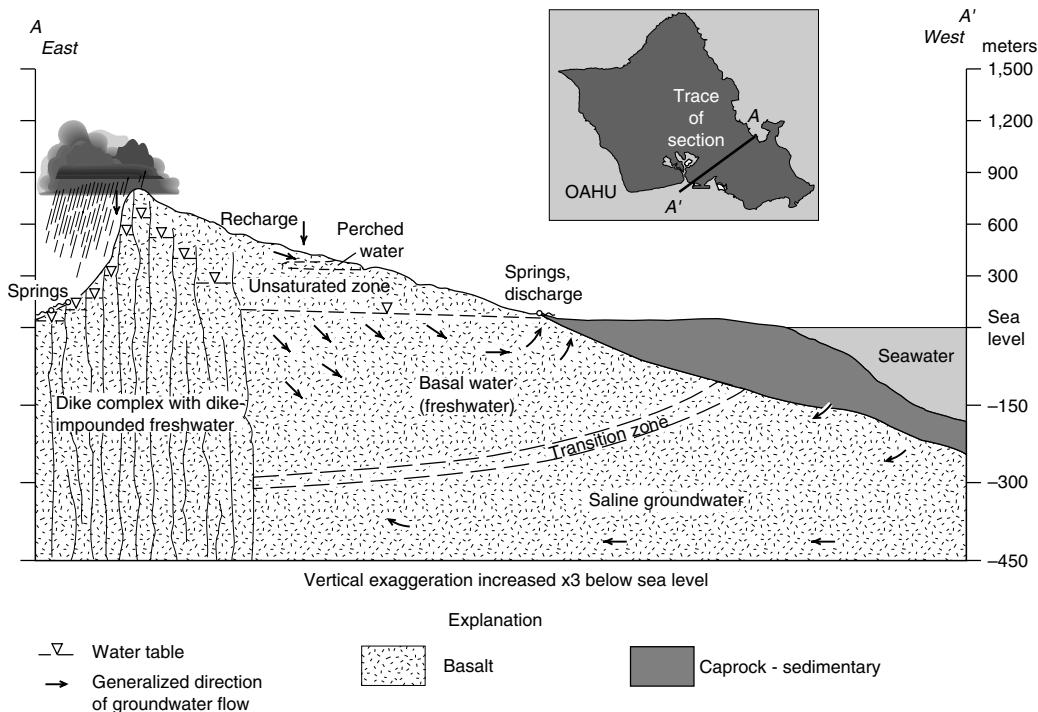


FIGURE 23.17 Schematic cross section of typical groundwater bodies in Hawaii. (From Gingerich, S.B. and Voss, C.I. (2005) *Hydrogeol. J.*, 13: 436–450; after Hunt, C.D., Jr. (1996) Geohydrology of the island of Oahu, Hawaii, U.S. Geol. Surv. Prof. Paper 1412-B.)

concentrations of the three nodes that represent the pumping center. This represents the average salinity of the water pumped from the areally extensive well field that is present in a strip across the aquifer system underlying the Pearl Harbor area.

The TZ evolution was modeled from predevelopment conditions (1880) through recent time (1990). The calculated distribution of seawater, freshwater, and irrigation water in 1990 (Figure 23.19) and the concurrent flow field (Figure 23.20) represented the starting conditions for the simulation of water-quality evolution in the pumping center for the period 1980–2080 for eight scenarios, which had different assumed withdrawal rates and different recharge rates for future conditions. These numerical experiments were used to help assess management alternatives. The results show that careful management of groundwater system withdrawals during droughts can preserve water quality.

An analysis of controls on water-quality evolution indicates that the key control on long-term saltwater intrusion is the net system discharge (equal to the difference between total recharge and pumpage), not the percentage of recharge pumped, and the maximum safe withdrawal will vary depending on the actual long-term recharge rate (Souza and Voss 1989). The effect is illustrated in Figure 23.21 wherein the average concentration of seawater in water produced at the pumping center in the year 2080 is a strong function of net system discharge. This curve is the result of several independent simulations, each with different recharge and withdrawals. The net system discharge is important because it controls the ultimate thickness of the freshwater lens downstream of the pumping center.

A “minimum safe discharge” from the aquifer underlying the Pearl Harbor area is defined on the basis of simple interpretations of what the average salinity of water produced in the coastal pumping zone means in terms of a “safe yield,” and indicates the least quantity of water that must be allowed to discharge naturally from the aquifer to ensure production of potable water in 2080 from the zone of pumping (Voss 1999). This modeling indicates that if the long-term net system discharge rate drops to within the range

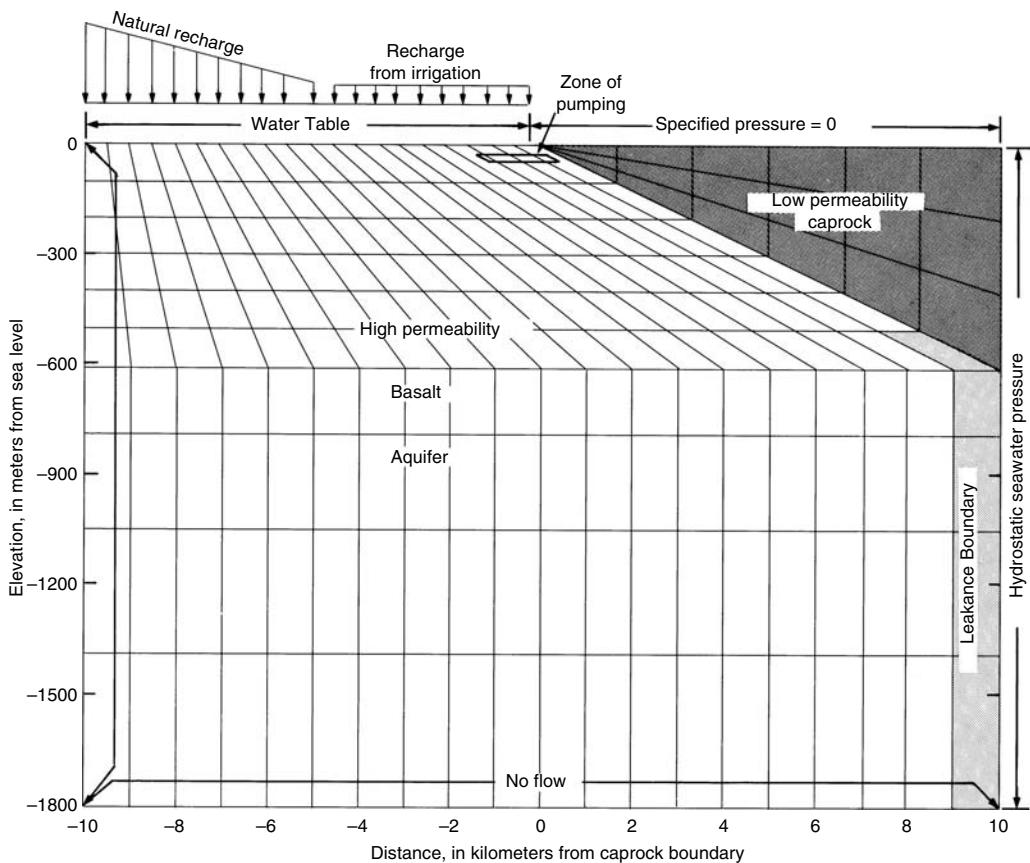


FIGURE 23.18 Schematic SUTRA finite-element mesh for transition zone modeling showing boundary conditions applied to model and regions of constant permeability. Actual vertical discretization is finer than illustrated. (From Voss, C.I. and Souza, W.R. (1998) Dynamics of a regional freshwater–saltwater transition zone in an anisotropic coastal aquifer system, U.S. Geol. Survey Open-File Rept. 98-398.)

of minimum safe discharge (1.7 to $2.9 \times 10^5 \text{ m d}^{-1}$), then significant seawater intrusion in parts of the Pearl Harbor area pumping zone will occur by 2080.

A 3D model of the aquifer underlying the Pearl Harbor area was developed subsequently as a lateral expansion of the 2D cross-sectional model by incorporating realistic areal distributions of hydrogeologic structure, recharge, and pumping (Gingerich and Voss 2005). One obvious advantage of 3D modeling is the ability to include realistic areal distributions allowing the model to be used as a management tool for testing or optimizing various pumping distribution schemes. An objective in constructing the 3D model was to maintain the general spatial geometry and boundary conditions of the 2D cross-sectional model as well as the 2D physical property and parameter values, and to add no additional complexity except for the realistic areal geometry of model features and the extra parameters required in 3D (one additional permeability and several dispersivity parameters). It was expected that, without significant adjustments, this 3D model should be a good representation of the aquifer because of the descriptive strength and functional simplicity of the 2D model upon which it is based.

One of the most useful applications of a 3D variable-density model is to simulate the response of the TZ to pumping stresses on the aquifer. The simulated rise and spread of the TZ at well 2300-18 (an open borehole) during 100 years of pumping is illustrated in Figure 23.22 where salinity concentrations equivalent to 2, 10, 50, and 90% seawater concentration are shown. The simulated trends of the four

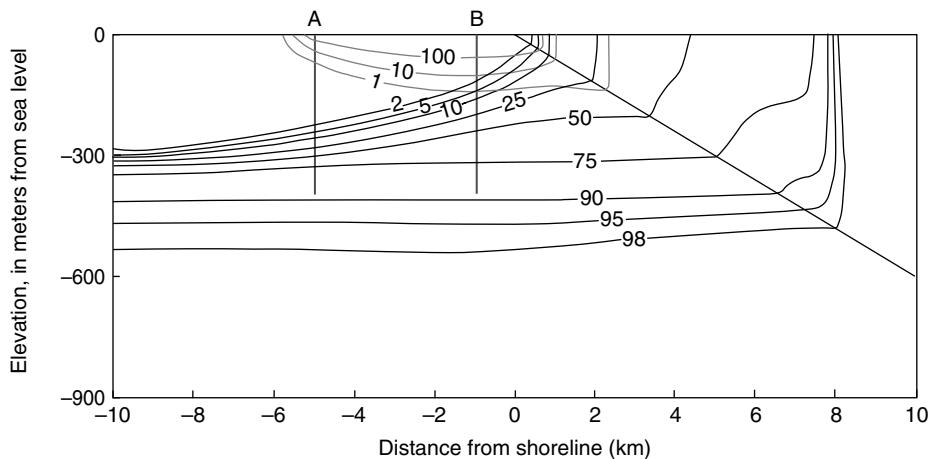


FIGURE 23.19 Modeled cross section of aquifer underlying Pearl Harbor area showing seawater distribution from 1880 to 1990 transient simulation (lower set of contours in units of percent seawater), irrigation recharge water distribution (upper set of contours in units of percent irrigation water), and location of monitor wells. Upper half of model section is shown. (Adapted from Voss, C.I. and Wood, W.W. (1994) In *Mathematical Models and their Applications to Isotope Studies in Groundwater Hydrology*, International Atomic Energy Agency (IAEA) Vienna, Austria, IAEA-TECDOC-777, pp. 147–178.)

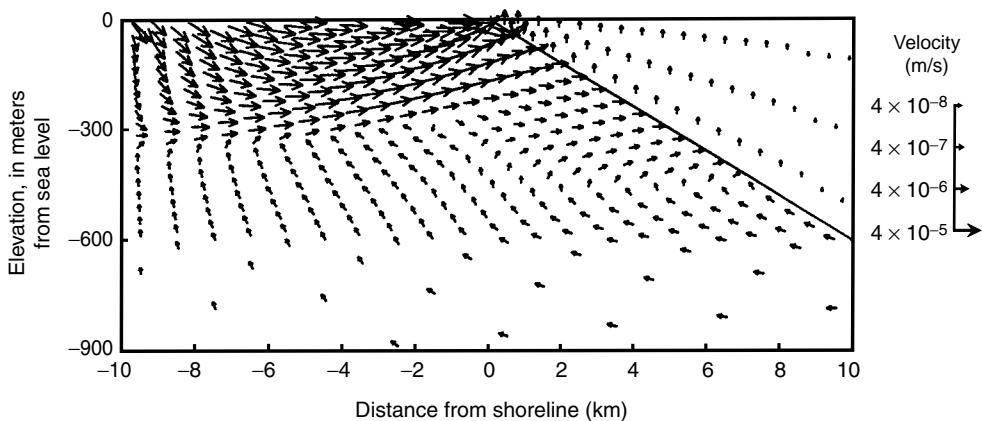


FIGURE 23.20 Modeled cross section of aquifer underlying Pearl Harbor area showing velocity distribution in 1990 from transient simulation. Upper half of model section is shown. (Adapted from Voss, C.I. and Wood, W.W. (1994) In *Mathematical Models and Their Applications to Isotope Studies in Groundwater Hydrology*, International Atomic Energy Agency (IAEA) Vienna, Austria, IAEA-TECDOC-777, pp. 147–178.)

salinity-concentration lines are generally parallel to each other, but show more spreading below the 50% line. The periods of most rapid rise in the TZ near the well were between 1910–15 (5 m/year) and 1970–75 (2.6 m/year) (Gingerich and Voss 2005). In many field situations, where TZ information from wells is lacking, the Ghyben–Herzberg (G–H) relation is commonly used to estimate the position of the TZ and the amount of freshwater available in the aquifer. The relation indicates that the depth of the interface between freshwater and saltwater below sea level is 40 times the altitude of the water table above sea level. The validity of this assumption can be checked by comparing the depth of this theoretical freshwater–saltwater interface with the simulated salinity-concentration lines (Figure 23.22). In this simulation, the

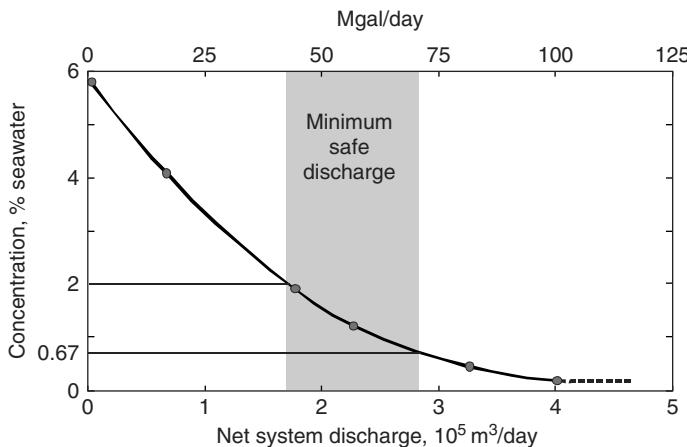


FIGURE 23.21 Dependence of average salinity of water produced in the Pearl Harbor area aquifer's zone of pumping after 100 years of pumping (1980–2080) on the quantity of water that still discharges naturally from the aquifer after water is removed by pumping (net system discharge). (From Voss, C.I. (1999) In J. Bear et al. (eds.) *Seawater Intrusion in Coastal Aquifers*, Kluwer Academic Publ., Dordrecht, pp. 249–313.)

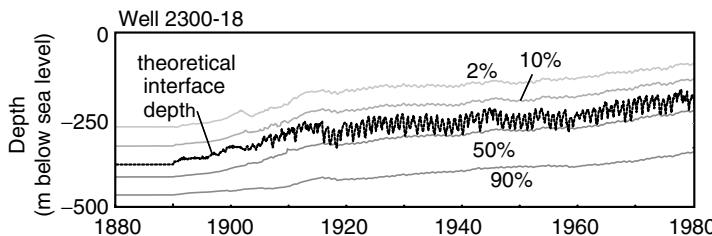


FIGURE 23.22 Position of simulated transition zone and theoretical (Ghyben–Herzberg) interface depth at well 2300-18, 1880–1980. (From Gingerich, S.B. and Voss, C.I. (2005) *Hydrogeol. J.*, 13: 436–450.)

depth of the theoretical interface is nearly always shallower than the 50% seawater concentration depth, indicating that the G–H relation is not a good predictor of the transition-zone position.

The general approach to coastal seawater intrusion modeling, as applied in Oahu, provides reasonable predictions of changes in the regional position of the potable-water boundary above the transition zone that result from changes in stress in a coastal aquifer. A strategy that controls withdrawals on the basis of (1) accurate estimates of recharge and (2) water-quality predictions from coastal aquifer modeling should be a major component of a total management scheme for a coastal aquifer water resource.

23.9 Available Groundwater Software

A large number of generic deterministic groundwater modeling software packages, based on a variety of numerical methods and a variety of conceptual models, are available. The selection of a numerical method or generic model for a particular field problem depends on several factors, including accuracy, efficiency/cost, and usability. The first two factors are related primarily to the nature of the field problem, availability of data, and scope, or intensity of the investigation. The usability of a method may depend partly on the mathematical background of the modeler, as it is preferable for the model user to understand the nature of the numerical methods implemented in a code. In selecting a model that is appropriate for a particular application, it is most important to choose one that incorporates the proper conceptual model; one must avoid force fitting an inappropriate model to a field situation solely because of the

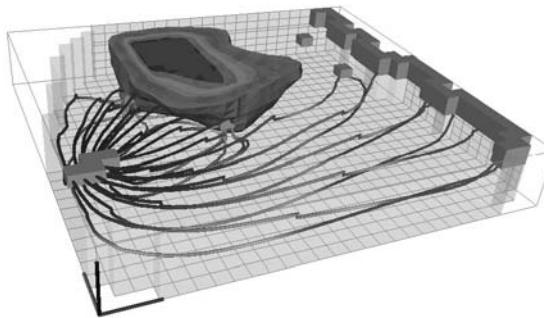


FIGURE 23.23 (See color insert following page 23-52) Example plot generated with Model Viewer software showing the grid shell, selected model features and boundary conditions, particle pathlines (shaded by time of travel), and a solid rendering of a contaminant plume for a 3D MODFLOW simulation. (From Hsieh, P.A. and Winston, R.B. (2002) User's guide to Model Viewer, a program for three-dimensional visualization of ground-water model results, U.S. Geol. Survey Open-File Rept. 02-106.)

model's convenience, availability, or familiarity to the user. Usability is also enhanced by the availability of preprocessing and postprocessing programs or features, and by the availability of comprehensive yet understandable documentation. Many groundwater models are in the public domain and available for downloading from the internet at no cost; some are proprietary codes available for purchase. However, most GUIs, GIS Packages, and visualization software are commercial products that must be purchased.

One available public domain postprocessor, which works with output from several different specific models, is Model Viewer (Hsieh and Winston 2002). Scalar data (such as hydraulic head or solute concentration) may be displayed as a solid or a set of isosurfaces, using a red-to-blue color spectrum to represent a range of scalar values. Vector data (such as velocity or specific discharge) are represented by lines oriented to the vector direction and scaled to the vector magnitude. Model Viewer can also display pathlines, cells, or nodes that represent model features (such as constant-head boundary cells, streams, and wells), and auxiliary graphic objects (such as grid lines and coordinate axes). Figure 23.23 represents an illustrative example showing the grid outline, selected model features and boundary conditions, particle pathlines, and a solid rendering of a contaminant plume. Users may rotate and crop the model grid in different orientations to examine the interior structure of the data. For transient simulations, Model Viewer can animate the time evolution of the simulated quantities.

Glossary

Analytical Model A closed-form exact mathematical solution, which is continuous in space and time.

Conceptual Model A hypothesis for how a system or process operates.

Deterministic Model A mathematical model based on conservation of mass, momentum, and energy.

Discretization The process of representing a continuous system by a set of discrete blocks, cells, or elements.

Generic Model The computer code used to solve one or more governing differential equations.

Mathematical Model A set of equations, which include mathematical variables, constants, and coefficients, that represents relevant processes.

Model A representation of a real system or process.

Numerical Model An approximate solution of a differential equation obtained by replacing the continuous variables with a set of discrete variables defined at grid blocks, cells, or nodes.

Site-Specific Model A numerical model with the parameters (such as hydraulic conductivity, dispersivity, etc.), boundary conditions, and grid dimensions of the generic model specified to represent a particular geographical area.

Stochastic Involving random variables or random processes, and amenable to analysis using probabilistic concepts.

Acknowledgments

The authors appreciate the helpful review comments provided by Robert Nicholson and Michael Planert of the USGS, and by Mary P. Anderson of the University of Wisconsin-Madison, on the first edition of this chapter, and by Brendan A. Zinn, Philip T. Harte, and Angel Martin, Jr. of the USGS on this revised version for the second edition.

References

- Ahlfeld, D.P. and Mulligan, A.E. (2000) *Optimal Management of Flow in Groundwater Systems*, Academic Press, San Diego.
- Ahlfeld, D.P. and Riefler, G. (2003) Documentation for MODFOC — A program for solving optimal flow control problems based on MODFLOW simulation, version 2.3. Amherst, MA, University of Massachusetts Dept. of Civil and Env. Eng.
- Ahlfeld, D.P., Barlow, P.M. and Mulligan, A.E. (2005) GWM — A ground-water management process for the U.S. Geological Survey modular ground-water model (MODFLOW-2000). U.S. Geol. Survey Open-File Rept. 2005-1072.
- Anderson, M.P. (1984) Movement of contaminants in groundwater: Groundwater transport — Advection and dispersion. In *Groundwater Contamination*, pp. 37–45, National Academy Press, Washington, D.C.
- Anderson, M.P., Ward, D.S., Lappala, E.G. and Prickett, T.A. (1992) Computer models for subsurface water. In *Handbook of Hydrology*, ed. D.R. Maidment, pp. 22.1–22.34, McGraw-Hill, Inc., New York.
- Anderson, M.P. and Woessner, W.W. (1992) *Applied Groundwater Modeling*, Academic Press, San Diego.
- Barlow, P.M. and Dickerman, D.C. (2001) Numerical-simulation and conjunctive-management models of the Hunt-Annaquatic-Pettaquamscutt stream-aquifer system, Rhode Island. U.S. Geol. Survey Prof. Paper 1636.
- Barrett, R., Berry, M., Chan, T.F., Demmel, J., Donato, J., Dongarra, J., Eijkhout, V., Pozo, R., Romine, C. and van der Vorst, H. (1994) *Templates for the Solution of Linear Systems: Building Blocks for Iterative Methods*, Soc. for Industrial and Applied Mathematics, Philadelphia.
- Bear, J. (1979) *Hydraulics of Groundwater*, McGraw-Hill, New York.
- Bear, J. and Verruijt, A. (1987) *Modeling Groundwater Flow and Pollution*, Reidel Publishing Co., Dordrecht, Holland.
- Bennett, G.D. (1976) Introduction to ground-water hydraulics: A programmed text for self-instruction, *Techniques of Water-Res. Invests. of the U.S. Geol. Survey*, Book 3, Ch. B2.
- Bower, K.M., Gable, C.W. and Zyvoloski, G.A. (2005) Grid resolution study of ground water flow and transport, *Ground Water*, 43(1): 122–132.
- Bredehoeft, J.D. and Konikow, L.F. (1993) Ground-water models: Validate or invalidate. *Ground Water*, 31: 178–179.
- Buxton, H.T. and Smolensky, D.A. (1999) Simulation of the effects of development of the ground-water flow system of Long Island, New York: U.S. Geol. Survey Water-Res. Inv. Rept. 98-4069.
- Carrera, J., Alcolea, A., Medina, A., Hidalgo, J. and Slooten, L.J. (2005) Inverse problems in hydrogeology. *Hydrogeol. J.*, 13: 206–222.
- Cooley, R.L. (1977) A method of estimating parameters and assessing reliability for models of steady state ground-water flow, 1, Theory and numerical properties. *Water Resour. Res.*, 13: 318–324.
- Cooley, R.L. (1992) A modular finite-element model (MODFE) for areal and axisymmetric ground-water flow problems, Part 2: Derivation of finite-element equations and comparisons with analytical solutions, *Techniques of Water-Res. Invests. of the U.S. Geol. Survey*, Book 6, Ch. A4.

- Domenico, P.A. and Robbins, G.A. (1984) A dispersion scale effect in model calibrations and field tracer experiments. *J. Hydrol.*, 70: 123–132.
- Domenico, P.A. and Schwartz, F.W. (1997) *Physical and Chemical Hydrogeology*, 2nd ed., John Wiley & Sons, New York.
- Franke, O.L. and Reilly, T.E. (1987) The effects of boundary conditions on the steady-state response of three hypothetical ground-water systems — Results and implications of numerical experiments, U.S. Geol. Survey Water-Supply Paper 2315.
- Franke, O.L., Reilly, T.E. and Bennett, G.D. (1987) Definition of boundary and initial conditions in the analysis of saturated ground-water flow systems — An introduction, *Techniques of Water-Res. Invests. of the U.S. Geol. Survey*, Book 3, Ch. B5.
- Franke, O.L., Reilly, T.E., Pollock, D.W. and LaBaugh, J.W. (1998) Estimating areas contributing recharge to wells: lessons from previous studies. U.S. Geol. Survey Circular 1174.
- Freyberg, D.L. (1988) An exercise in ground-water model calibration and prediction. *Ground Water*, 26: 350–360.
- Garder, A.O., Peaceman, D.W. and Pozzi, A.L. (1964) Numerical calculation of multidimensional miscible displacement by the method of characteristics. *Soc. Petroleum Eng. J.*, 4: 26–36.
- Gelhar, L.W. (1993) *Stochastic Subsurface Hydrology*, Prentice Hall, Englewood Cliffs, NJ.
- Gelhar, L.W., Welty, C. and Rehfeldt, K.R. (1992) A critical review of data on field-scale dispersion in aquifers. *Water Resour. Res.*, 28: 1955–1974.
- Gingerich, S.B. and Voss, C.I. (2005) Three-dimensional variable-density flow simulation of a coastal aquifer in southern Oahu, Hawaii, USA. *Hydrogeol. J.*, 13: 436–450.
- Goode, D.J. and Appel, C.A. (1992) Finite-difference interblock transmissivity for unconfined aquifers and for aquifers having smoothly varying transmissivity, U.S. Geol. Survey Water-Res. Inv. Rept. 92–4124.
- Goode, D.J. and Konikow, L.F. (1989) Modification of a method-of-characteristics solute-transport model to incorporate decay and equilibrium-controlled sorption or ion exchange, U.S. Geol. Survey Water-Res. Inv. Rept. 89–4030.
- Goode, D.J. and Konikow, L.F. (1990) Apparent dispersion in transient groundwater flow. *Water Resour. Res.*, 26: 2339–2351.
- Greenwald, R.M. (1998) Documentation and user's guide — MODMAN, an optimization module for MODFLOW, version 4.0: HSI GeoTrans, Freehold, NJ.
- Grove, D.B. (1976) Ion exchange reactions important in groundwater quality models. In *Advances in Groundwater Hydrology*, pp. 409–436. Am. Water Res. Assoc.
- Haitjema, H.M. (1995) *Analytic Element Modeling of Groundwater Flow*, Academic Press, San Diego, CA.
- Harbaugh, A.W. (2005) MODFLOW-2005, The U.S. Geological Survey modular ground-water model—the Ground-Water Flow Process, U.S. Geol. Survey Techniques and Methods 6-A16.
- Harbaugh, A.W., Banta, E.R., Hill, M.C. and McDonald, M.G. (2000) MODFLOW-2000, The U.S. Geological Survey modular ground-water model — User guide to modularization concepts and the ground-water flow process, U.S. Geol. Survey Open-File Rept. 00-92.
- Harbaugh, A.W. and McDonald, M.G. (1996) User's documentation for MODFLOW-96, an update to the U.S. Geological Survey modular finite-difference ground-water flow model, U.S. Geol. Survey Open-File Rept. 96–485.
- Heberton, C.I., Russell, T.F., Konikow, L.F. and Hornberger, G.Z. (2000) A three-dimensional finite-volume Eulerian-Lagrangian Localized Adjoint Method (ELLAM) for solute-transport modeling, U.S. Geol. Survey Water-Res. Inv. Rept. 00-4087.
- Hill, M.C. (1998) Methods and guidelines for effective model calibration, U.S. Geol. Survey Water-Res. Inv. Rept. 98-4005.
- Hill, M.C., Banta, E.R., Harbaugh, A.W. and Anderman, E.R. (2000) MODFLOW-2000, The U.S. Geological Survey Modular Ground-Water Model—User guide to the Observation, Sensitivity, and Parameter-Estimation Processes and three post-processing programs: U.S. Geol. Survey Open-File Rept. 00-184.

- Hsieh, P.A. and Winston, R.B. (2002) User's guide to Model Viewer, A program for three-dimensional visualization of ground-water model results, U.S. Geol. Survey Open-File Rept. 02-106.
- Huebner, K.H. (1975) *The Finite Element Method for Engineers*, John Wiley & Sons, New York.
- Hunt, C.D., Jr. (1996) Geohydrology of the island of Oahu, Hawaii, U.S. Geol. Surv. Prof. Paper 1412-B.
- Huyakorn, P.S. and Pinder, G.F. (1983) *Computational Methods in Subsurface Flow*, Academic Press, New York.
- Javandel, I., Doughty, D. and Tsang, C.-F. (1984) *Groundwater Transport: Handbook of Mathematical Models*, Am. Geophysical Union, Water Res. Monograph 10.
- Kipp, K.L., Jr. (1987) HST3D: A computer code for simulation of heat and solute transport in three-dimensional ground-water flow systems, U.S. Geol. Survey Water-Res. Inv. Rept. 86-4095.
- Konikow, L.F. (1996) Numerical models of ground-water flow and transport. In *Manual on Mathematical Models in Isotope Hydrogeology*, International Atomic Energy Agency, Vienna, Austria.
- Konikow, L.F. and Bredehoeft, J.D. (1978) Computer model of two-dimensional solute transport and dispersion in ground water, *Techniques of Water-Res. Invests. of the U.S. Geol. Survey*, Book 7, Ch. C2.
- Konikow, L.F. and Bredehoeft, J.D. (1992) Ground-water models cannot be validated. *Adv. Water Resour.*, 15: 75–83.
- Konikow, L.F., Goode, D.J. and Hornberger, G.Z. (1996) A three-dimensional method-of-characteristics solute-transport model (MOC3D), U.S. Geol. Survey Water-Res. Inv. Rept. 96-4267.
- Konikow, L.F. and Mercer, J.M. (1988) Groundwater flow and transport modeling. *J. Hydrol.*, 100: 379–409.
- Konikow, L.F., Sanford, W.E. and Campbell, P.J. (1997) Constant-concentration boundary condition: Lessons from the HYDROCOIN variable-density benchmark problem. *Water Resour. Res.*, 33: 2253–2261.
- Lambert, P.M. (1996) Numerical simulation of the movement of sulfate in ground water in southwestern Salt Lake Valley, Utah, Tech. Pub. No. 110-D, Utah Dept. of Natural Resources, Salt Lake City, UT.
- Langevin, C.D. (2001) Simulation of ground-water discharge to Biscayne Bay, Southeastern Florida: U.S. Geol. Survey Water-Res. Inv. Rept. 00-4251.
- Langevin, C.D., Shoemaker, W.B. and Guo, W. (2003) MODFLOW-2000, The U.S. Geological Survey modular ground-water model — Documentation of the SEAWAT-2000 version with the variable-density flow process (VDF) and the integrated MT3DMS transport process (IMT), U.S. Geol. Survey Open-File Rept. 03-426.
- Lefkoff, L.J. and Gorelick, S.M. (1987) AQMAN-Linear and quadratic programming matrix generator using two-dimensional ground-water flow simulation for aquifer management modeling. U.S. Geol. Survey Water-Res. Inv. Rept. 87-4061.
- Masterson, J.P. (2004) Simulated interaction between freshwater and saltwater and effects of ground-water pumping and sea-level change, lower Cape Cod Aquifer System, Massachusetts: U.S. Geol. Survey Scientific Inv. Rept. 2004-5014.
- Masterson, J.P., Walter, D.A. and Savoie, J. (1996) Use of particle tracking to improve numerical model calibration and to analyze ground-water flow and contaminant migration, Massachusetts Military Reservation, western Cape Cod, Massachusetts, U.S. Geol. Survey Open-File Rept. 96-214.
- McAda, D.P. and Barroll, P. (2002) Simulation of ground-water flow in the Middle Rio Grande Basin between Cochiti and San Acacia, New Mexico: U.S. Geol. Survey Water-Res. Inv. Rept. 02-4200.
- McDonald, M.G. and Harbaugh, A.W. (1988) A modular three-dimensional finite-difference ground-water flow model, *Techniques of Water-Res. Invests. of the U.S. Geol. Survey*, Book 6, Ch. A1.
- Mercer, J.W. and Faust, C.R. (1981) *Ground-Water Modeling*, Natl. Water Well Assoc., Worthington, OH.
- Oreskes, N., Shrader-Frechette, K. and Belitz, K. (1994) Verification, validation, and confirmation of numerical models in the earth sciences. *Science*, 263: 641–646.

- Panday, S. and Huyakorn, P.S. (2004) A fully coupled physically-based spatially-distributed model for evaluating surface/subsurface flow. *Adv. Water Resour.*, 27: 361–382.
- Peaceman, D.W. (1977) *Fundamentals of Numerical Reservoir Simulation*, Elsevier, Amsterdam.
- Peralta, R.C. (2004) Simulation/Optimization Modeling System (SOMOS). Logan, UT, Utah State University Department of Biological and Irrigation Engineering.
- Poeter, E.P. and Hill, M.C. (1998) Documentation of UCODE, A computer code for universal inverse modeling, U.S. Geol. Survey Water-Res. Inv. Rept. 98-4080.
- Poeter, E.P., Hill, M.C., Banta, E.R., Mehl, S. and Christensen, S. (2005) UCODE_2005 and six other computer codes for universal sensitivity analysis, calibration, and uncertainty evaluation: U.S. Geol. Survey Techniques and Methods 6-A11.
- Pollock, D.W. (1988) Semianalytical computation of path lines for finite-difference models. *Ground Water*, 26: 743–750.
- Pollock, D.W. (1989) Documentation of computer programs to compute and display pathlines using results from the U.S. Geological Survey modular three-dimensional finite-difference ground-water flow model, U.S. Geol. Survey Open-File Rept. 89-381.
- Pollock, D.W. (1994) User's guide for MODPATH/MODPATH-PLOT, version 3: A particle tracking post processing package for MODFLOW, the U.S. Geological Survey finite-difference ground-water flow model, U.S. Geol. Survey Open-File Rept. 94-464.
- Popper, K. (1959) *The Logic of Scientific Discovery*, Harper and Row, New York.
- Prickett, T.A., Naymik, T.G. and Lonnquist, C.G. (1981) A “random-walk” solute transport model for selected groundwater quality evaluations, Ill. State Water Survey Bulletin 65.
- Prudic, D.E. (1989) Documentation of a computer program to simulate stream-aquifer relations using a modular, finite-difference, ground-water flow model. U.S. Geol. Survey Open-File Rept. 88-729.
- Reddell, D.L. and Sunada, D.K. (1970) Numerical Simulation of Dispersion in Groundwater Aquifers, Colorado State University, Ft. Collins, Hydrology Paper 41.
- Reilly, T.E. (2001) System and boundary conceptualization in ground-water flow simulation: *Techniques of Water-Res. Invests. of the U.S. Geol. Survey*, Book 3, Ch. B8.
- Reilly, T.E., Franke, O.L., Buxton, H.T. and Bennett, G.D. (1987) A conceptual framework for ground-water solute-transport studies with emphasis on physical mechanisms of solute movement, U.S. Geol. Survey Water-Res. Inv. Rept. 87-4191.
- Reilly, T.E. and Harbaugh, A.W. (2004) Guidelines for evaluating ground-water flow models: U.S. Geol. Survey Scientific Inv. Rept. 2004-5038.
- Reilly, T.E., Plummer, L.N., Phillips, P.J. and Busenberg, E. (1994) The use of simulation and multiple environmental tracers to quantify groundwater flow in a shallow aquifer. *Water Resour. Res.*, 30: 421–433.
- Remson, I., Hornberger, G.M. and Molz, F.J. (1971) *Numerical Methods in Subsurface Hydrology*, Wiley, New York.
- Ségal, G. (1994) *Classic Groundwater Simulations: Proving and Improving Numerical Models*, PTR Prentice Hall, Englewood Cliffs.
- Smith, L. and Schwartz, F.W. (1980) Mass transport, 1, A stochastic analysis of macroscopic dispersion. *Water Resour. Res.*, 16: 303–313.
- Souza, W.R. and Voss, C.I. (1989) Assessment of potable ground water in a fresh-water lens using variable-density flow and solute transport simulation. In: Proceedings, National Water Well Association (NWWA) Conference on Solving Ground Water Problems with Models, Indianapolis, Indiana, Feb. 7–9, 1989, pp. 1023–1043.
- Swedish Nuclear Power Inspectorate (1987) The International HYDROCOIN Project — Background and Results, OECD, Paris.
- Tenbus, F.J. and Fleck, W.B. (2001) Simulation of ground-water flow and transport of chlorinated hydrocarbons at Graces Quarters, Aberdeen Proving Ground, Maryland: U.S. Geol. Survey Water-Res. Inv. Rept. 01-4106.

- Torak, L.J. (1993) A modular finite-element model (MODFE) for areal and axisymmetric ground-water flow problems, Part 1: Model description and user's manual, *Techniques of Water-Res. Invests. of the U.S. Geol. Survey*, Book 6, Ch. A3.
- Turco, M.J. and Buchmiller, R.C. (2004) Simulation of ground-water flow in the Cedar River alluvial aquifer flow system, Cedar River, Iowa: U.S. Geol. Survey Scientific Inv. Rept. 2004-5130.
- Voss, C.I. (1984) SUTRA — Saturated Unsaturated Transport — A finite-element simulation model for saturated-unsaturated fluid-density-dependent ground-water flow with energy transport or chemically-reactive single-species solute transport, U.S. Geol. Survey Water-Res. Inv. Rept. 84-4369.
- Voss, C.I. (1999) USGS SUTRA Code — History, practical use, and application in Hawaii. In J. Bear et al. (eds.) *Seawater Intrusion in Coastal Aquifers*, Kluwer Academic Publ., Dordrecht, pp. 249–313.
- Voss, C.I., Boldt, D. and Shapiro, A.M. (1997) A graphical-user interface for the U.S. Geological Survey's SUTRA code using Argus ONE (for simulation of variable-density saturated-unsaturated ground-water flow with solute or energy transport): U.S. Geol. Survey Open-File Rept. 97-421.
- Voss, C.I. and Provost, A.M. (2002) SUTRA, A model for saturated-unsaturated variable-density ground-water flow with solute or energy transport, U.S. Geol. Survey Water-Res. Inv. Rept. 02-4231.
- Voss, C.I. and Souza, W.R. (1987) Variable density flow and solute transport simulation of regional aquifers containing a narrow freshwater–saltwater transition zone. *Water Resour. Res.*, 23: 1851–1866.
- Voss, C.I. and Souza, W.R. (1998) Dynamics of a regional freshwater–saltwater transition zone in an anisotropic coastal aquifer system, U.S. Geol. Survey Open-File Rept. 98-398.
- Voss, C.I. and Wood, W.W. (1994) Synthesis of geochemical, isotopic and ground-water modeling analysis to explain regional flow in a coastal aquifer of southern Oahu, Hawaii. In *Mathematical Models and their Applications to Isotope Studies in Groundwater Hydrology*, International Atomic Energy Agency (IAEA) Vienna, Austria, IAEA-TECDOC-777, pp. 147–178.
- Wang, J.F. and Anderson, M.P. (1982) (reprinted, 1995). *Introduction to Groundwater Modeling*, Academic Press, San Diego, CA.
- Wexler, E.J. (1992) Analytical solutions for one-, two-, and three-dimensional solute transport in ground-water systems with uniform flow, *Techniques of Water-Res. Invests. of the U.S. Geol. Survey*, Book 3, Ch. B7.
- Winston, R.B. and Voss, C.I. (2003) SutraGUI: A Graphical User Interface for SUTRA, A model for ground-water flow with solute or energy transport: U.S. Geol. Survey Open-File Rept. 03-285.
- Yager, R.M. (1997) Simulated three-dimensional ground-water flow in the Lockport Group, a fractured dolomite aquifer near Niagara Falls, New York, U.S. Geol. Survey Water-Supply Paper 2487.
- Zheng, C. and Bennett, G.D. (2002) *Applied Contaminant Transport Modeling*, 2nd ed., Wiley-Interscience, New York.
- Zheng, C. and Wang, P.P. (1999) MT3DMS: A modular three-dimensional multispecies model for simulation of advection, dispersion and chemical reactions of contaminants in groundwater systems; Documentation and User's Guide, Report SERDP-99-1, U.S. Army Engineer Research and Development Center, Vicksburg, MS.
- Zheng, C. and Wang, P.P. (2002) MGO — A Modular Groundwater Optimizer incorporating MODFLOW and MT3DMS, Documentation and user's guide. The University of Alabama and Groundwater Systems Research Ltd., Tuscaloosa, AL.
- Zienkiewicz, O.C. (1971) *The Finite Element Method in Engineering Science*, McGraw-Hill, London.
- Zimmermann, S., Koumoutsakos, P. and Kinzelbach, W. (2001) Simulation of pollutant transport using a particle method. *J. Comput. Phys.*, 173: 322–347.

Further Information

Text books and examples of good reports on site-specific models are provided as a starting point for readers who would like to obtain more information or study representative applications.

Text Books

Ahlfeld and Mulligan (2000) provide an introduction to combined simulation-optimization modeling of groundwater flow.

Anderson and Woessner (1992) present an overview of applied groundwater flow and advective transport modeling.

Zheng and Bennett (2002) present an overview of the theory and practice of contaminant transport modeling.

Examples of Reports on Site-Specific Models

Comprehensive reports on site-specific models provide insight into applied groundwater simulation. A few examples from the work of the United States Geological Survey are provided below. Obviously, this list is not inclusive and many other reports could have been listed.

Regional Flow Models

Buxton and Smolensky (1999) simulate the effects of water use on the groundwater-flow system on Long Island, New York.

McAda and Barroll (2002) describe a 3D flow model of the Albuquerque Basin in New Mexico.

Local Flow Model

Turco and Buchmiller (2004) Simulated groundwater flow in the Cedar River alluvial aquifer flow system, Cedar River, Iowa.

Local Advective-Transport Model

Barlow and Dickerman (2001) examined contributing areas to public-supply wells in Rhode Island.

Solute-Transport Model

Lambert (1996) used a 3D model to simulate contaminant plume in approximately 480-km² area in Utah.

Tenbus and Fleck (2001) simulated the transport of chlorinated hydrocarbons at a site on Aberdeen Proving Ground, Maryland.

Variable-Density Flow Models

Langevin (2001) simulated groundwater discharge to Biscayne Bay, Southeastern Florida.

Masterson (2004) simulated the effects of groundwater pumping and sea-level change on the lower Cape Cod Aquifer System, Massachusetts.

Model Calibration

Masterson et al. (1996) used particle tracking and contaminant plumes to improve calibration of a 3D flow model.

Yager (1997) used a parameter-estimation model (MODFLOWP) to help calibrate a 3D flow model for a fractured dolomite aquifer system.

Internet

A number of sites on the internet provide compendiums of codes and sources of information about groundwater modeling, as well as providing links to other Web sites related to groundwater modeling. Many of these sites allow codes to be downloaded. Examples of several groundwater-oriented Home Page locations are: <http://www.thehydrogeologist.com/>, <http://hydrologyweb.pnl.gov/>, <http://www.mines.edu/igwmc/>, and <http://www.ggsd.com/>. Also, many of the U.S. Geological Survey public domain codes are available from the “Software” tab on the USGS Water Resources Information Home page at: <http://water.usgs.gov/>.

24

Groundwater Model Validation

24.1	Introduction.....	24-1
24.2	Practical Views of Validation	24-2
24.3	The Proposed Model Validation Process	24-3
24.4	Acceptance Criteria for Stochastic Model Realizations and Application to PSA Model..... Proposed Acceptance Criteria for PSA • Single Validation Target Illustration • Testing the Efficacy of P_1 for a Single Validation Target • Multiple Validation Targets Illustration • Testing the Efficacy of P_1 for Multiple Validation Targets • Testing the Efficacy of P_2 for Multiple Validation Targets	24-6
24.5	Uncertainty Reduction Using Validation Data and Application to the Amchitka Model	24-18
	Model Background • Bayesian Framework for Uncertainty Reduction • Application to the Milrow Site in Amchitka	
24.6	Summary and Conclusions	24-22
	Glossary.....	24-23
	Acknowledgments.....	24-24
	References	24-24
	Further Information	24-27

Ahmed E. Hassan
*Desert Research Institute,
Las Vegas, Nevada, and
Cairo University, Giza, Egypt*

24.1 Introduction

The term validation has been the subject of literature debate for more than a decade. Some skeptics argue the term should not be used as it implies correctness and accuracy for a groundwater model that no one can claim. Similarly, others believe it is impossible to validate a groundwater numerical model because such a claim would assert a demonstration of truth that can never be attained for our approximate solutions to subsurface problems (Oreskes et al., 1994). Anderson and Bates (2001) edited a book covering the issue of model validation in hydrological sciences. Some of the quotes they presented from the skeptics and opponents to the use of the term “validation” include:

Absolute validity of a model is never determined. (National Research Council 1990).

“What is usually done in testing the predictive capability of a model is best characterized as calibration or history matching; it is only a limited demonstration of the reliability of the model. We believe the terms validation and verification have little or no place in groundwater science; these terms lead to a

false impression of model capability. More meaningful descriptors of the process include model testing, model evaluation, model calibration, sensitivity testing, benchmarking, history matching, and parameter estimation.” (Konikow and Bredehoeft 1992).

The above views consider validation in the strictest definition of the word. That is, they refer to validation as a demonstration of the accuracy of the model in representing the true system and they warn against misconception of the public about the meaning of the term. As discussed in Hassan (2004a), most models, if not all, are not being used to reveal the truth of a system (an objective they simply cannot achieve). Models are in many cases decision-making tools. When a model successfully passes a rigorous development, calibration, and testing process, it becomes a reasonable decision-making tool, often the best available for subsurface problems that demand answers now. The model validation process can be regarded as an additional filter for independent model evaluation to assess the suitability of a model for its given purpose (likely to be decision-making). Most of the literature debate focuses on validation terminology and not on the process. No one argues the process is unimportant, unneeded, or useless and no one disagrees on the concept of using an independent data set to test the model. The disagreement is on what we call it and what the implications are for the term we use.

This lack of focus on the development of model validation processes or tools is reflected clearly in Vogel and Sankarasubramanian’s (2003) discussion about the validation of watershed models. “When one considers the wide range of watershed models and the heavy emphasis on their calibration (Duan et al., 2003), it is surprising how little attention has been given to the problem of model validation. In three recent reviews of watershed modeling (Singh 1995; Hornberger and Boyer 1995; Singh and Woolhiser 2002) and watershed model calibration (Duan et al., 2003), there was little attention given to developments in the area of model validation. Although there is rough agreement on the goal of model validation, no agreement exists on a uniform methodology for executing model validation.”

There is an urgent need for a model validation process that can be used to evaluate model performance and build confidence in a model’s suitability for making decisions. This confidence building is a long-term iterative process and it is the author’s belief that this process is what should be termed model validation. Model validation is a process, not an end result. That is, the process of model validation cannot always assure acceptable model predictions or quality of the model. Rather, it provides safeguard against faulty models or inadequately developed and tested models. If model results end up being used as the basis for decision-making, then the validation process indicates the model is valid for making decisions (not necessarily an absolute representation of truth).

Again, such a process should not be viewed as a mechanism for proving that the model is valid, but rather as a mechanism for enhancing the model, reducing its uncertainty, and improving its predictions through an iterative, long-term confidence-building process. The process should contain trigger mechanisms that will drive the model back to the characterization–conceptualization–calibration–prediction stage (i.e., back to the beginning), but with a better understanding of the modeled system.

Following this introduction, a brief review of practical definitions of the term validation is presented in Section 24.2. A description of a proposed model validation process for stochastic groundwater models (Hassan, 2004b) is presented in Section 24.3. The details of this process will not be repeated here, but rather more explanation and discussion of some of the challenging aspects of the process will be addressed as they apply to two actual field sites: the Project Shoal Area (PSA) in Nevada and the Amchitka site in Alaska. In particular, we focus on two main aspects (1) the selection of the acceptance criteria for the stochastic model realizations for PSA presented in Section 24.4, and (2) the evaluation of model input parameters and the reduction of their uncertainty at Amchitka presented in Section 24.5. A wrap up and concluding remarks are presented in Section 24.6.

24.2 Practical Views of Validation

Refsgaard (2001) defines model validation as the process of demonstrating a given site-specific model is capable of making accurate predictions for periods outside a calibration period. He also adds that a

model is said to be validated if its accuracy and predictive capability in the validation period have been proven to lie within acceptable limits or errors. de Marsily et al. (1992) argued that using a model in a predictive mode and comparing it with new data does not (and does not have to) prove the model is correct for all circumstances, it only increases the confidence in the model's value. They added "We do not want certainty; we will be satisfied with engineering confidence."

Neuman (1992) defines the validation of safety assessment models as the process of building scientific confidence in the methods used to perform such assessment. However, he recognizes that this confidence-building approach to validation is possibly open-ended, as many iterations between modelers and regulators may be needed. Eisenberg et al. (1994) support the idea of confidence building and indicate this term recognizes that full scientific validation of models of performance assessment may be impossible and the acceptance of mathematical models for regulatory purposes should be based on appropriate testing, which will lead to a reasonable assurance that the results are acceptable. Hassanzadeh (1990) differentiates between research (or analysis) model validation and safety assessment modeling (or predictive modeling) validation. The former is a tool to help understand processes, uncertainties, and the like, whereas the latter is a goal to help the decision-making process. Sargent (1990) regards validation as a process that consists of performing tests and evaluations during model development until sufficient confidence is obtained that a model can be considered valid for its intended application.

There is a general consensus that the main concern is whether or not a model is adequate for its intended use and whether or not there is sufficient evidence that the model development followed logical and scientific approaches and did not fail to account for important features and processes. It should be noted that determining the adequacy of a model or building confidence in its prediction is not ideally a one-time exercise. It can be considered an iterative process that should be viewed as part of an integral loop with trigger mechanisms or decision points that force the model back to the data collection–conceptualization–calibration–prediction phase if the model validation process indicates the need to do so. Key to this process is the use of a diverse set of tests that should be designed to evaluate a diverse set of aspects related to the model.

It is clear that often-quoted statements such as "groundwater models cannot be validated" and "groundwater models can only be invalidated" (Konikow and Bredehoeft 1992, 1993; Bredehoeft and Konikow 1992, 1993) refer to validation in only the strictest sense, responding to a concern regarding layman's possible misconceptions. Unfortunately, such statements may lead to a relaxed attitude on the part of researchers, consultants, and even regulatory agencies when it comes to evaluating model predictions. With the perception that the groundwater model will never be validated, there may be a temptation to believe good model development, building, and calibration are sufficient and nothing more can be done. All groundwater modelers agree that their models cannot be validated in the strictest sense (at least with present-day technology) but similarly agree on the importance of post-prediction testing and evaluation. By expanding our definition of validation to encompass a long-term process of confidence building, modelers and model users can develop rigorous validation processes that will ultimately improve the models and the quality of decisions based on those models.

24.3 The Proposed Model Validation Process

Hassan (2004b) proposed an approach for the validation process of stochastic numerical models. The details of the process will not be repeated here, but rather more explanation and discussion of some of the challenging aspects will be addressed as they apply to two actual field sites. However, for completeness, the overall validation approach is briefly discussed here.

The validation process is designed to account for the stochastic nature of some groundwater models. The focus of the process is centered around the following three main themes (1) testing if predictions of numerical groundwater models and the underlying conceptual models are robust and consistent with regulatory purposes, (2) improving model predictions and reducing uncertainty using data collected

through field activities designed for model validation, and (3) linking validation efforts to long-term monitoring and management of the site.

Figure 24.1, adapted from Hassan (2004b), displays the step-by-step approach for performing the validation and post-audit processes for a typical stochastic groundwater model. The proposed steps of the model validation process are described below:

Step 1: Identify the data needed for validation, the number and location of the wells, and the type of laboratory or field experiments needed. Well locations can be determined based on the existing model and should favor locations likely to encounter fast migration pathways.

Step 2: Install the wells and collect the largest amount of data possible from the wells. The data should include geophysical logging, head measurements, conductivity measurements, contaminant concentrations, and any other information that could be used to test the model structure, input, or output.

Step 3: Perform different validation tests to evaluate various components of the model. The stochastic validation approach proposed by Luis and McLaughlin (1992) is an example of an approach that can be used to test the flow model output (heads) under saturated conditions. Other goodness-of-fit tests can be used for the heads to complement this stochastic approach. The philosophy here is to test each individual realization with as many diverse tests (in terms of the statistical nature of the test and the tested aspect of the model) as possible and have a quantitative measure of the adequacy of each realization in capturing the main features of the modeled system. It is important to note that goodness-of-fit results and other statistical results for the current realization will be used after analyzing all realizations to obtain some of the acceptance criteria measures, P_1 through P_5 , which are discussed in detail in Section 24.4.

Step 4: Link the results of the calibration accuracy evaluations performed during the model-building stage and the validation tests (step 3) for all realizations and sort the realizations in terms of their adequacy and closeness to the field data. A subjective element may be invoked in this sorting based on expert judgment and hydrogeologic understanding. The objective here is to filter out the realizations that show a major deviation in many of the tested aspects and focus on those that “passed” the majority of the tests. By doing so, the range of output uncertainty is reduced and the subsequent effort can be focused on the most representative realizations and scenarios. To continue reducing the uncertainty level, a refinement of the conductivity or other input distribution can be made based on information collected from the validation wells as presented in Section 24.5 for the Amchitka site.

Step 5: Step 4 results will determine the subsequent step and guide the decision of whether there is a sufficient number of realizations with high scores or not.

5a. If the number of realizations with high scores is very small compared to the total number of model realizations (the P_1 measure discussed in Section 24.4), it could be an indication that the model has a major deficiency or conceptual problem or it could be that the input is not correct. In the latter case, the model may be conceptually good, but the input parameter distributions are skewed one way or another. Generating more realizations and keeping those fitting the validation criteria can shift the distribution to the proper position. This can be done using the existing model without conditioning or using any of the new validation data. If the model has a major deficiency or conceptual problem, generating additional realizations will not correct it and continued failure per the validation criteria will be obvious. The importance of the distinction between conceptualization problems and inappropriate input distributions has been discussed by Bredehoeft (2003). He states one should carefully ask the question of whether the mismatch between the model and the field observations is a result of poor parameter adjustments or does it suggest we rethink the conceptual model. The use of different metrics such as P_1 and P_2 discussed in Section 24.4 provides a tool for making this distinction.

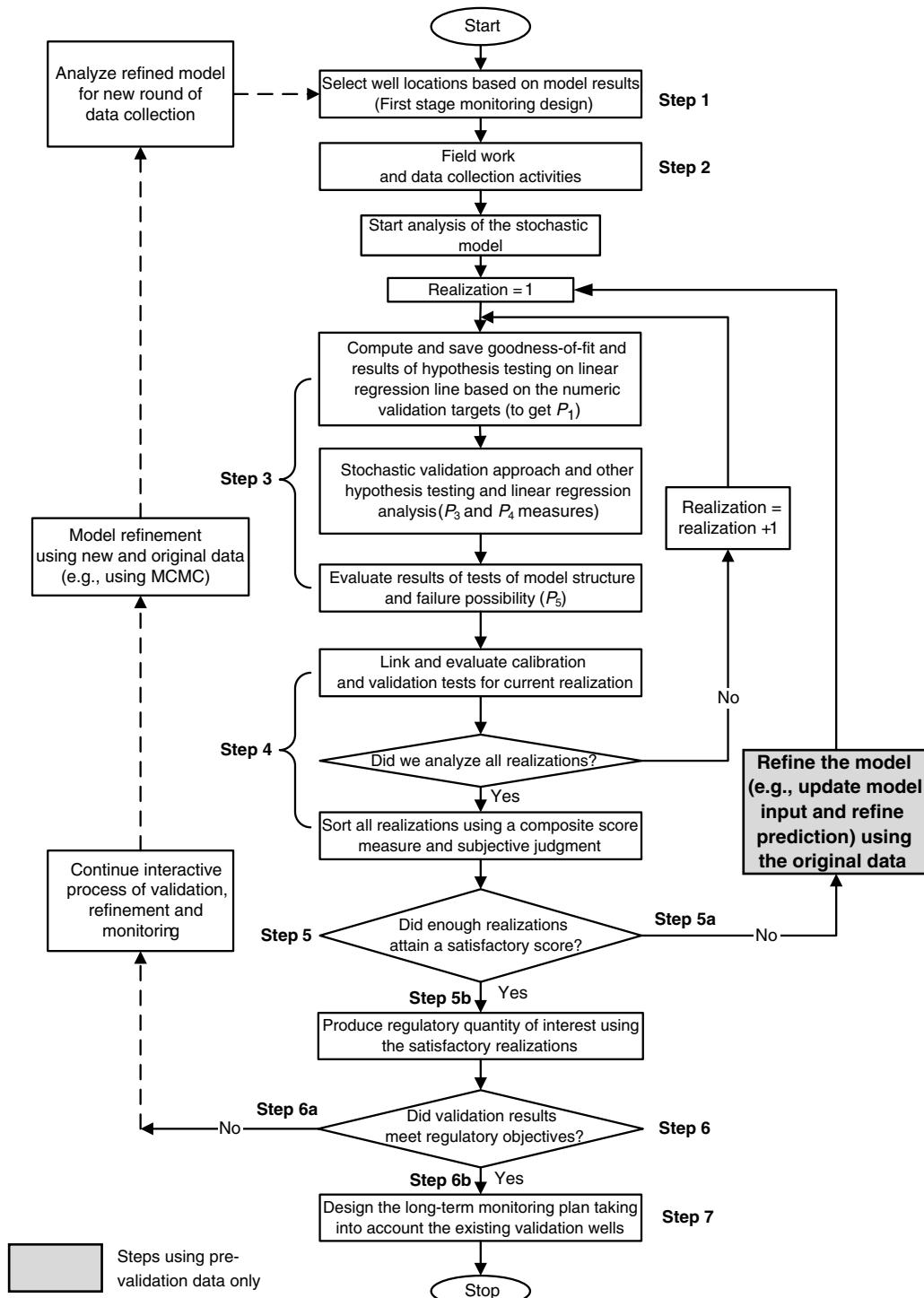


FIGURE 24.1 Details of the proposed model validation processes for stochastic groundwater models. (Reprinted from *Ground Water* with permission of the National Ground Water Association Press. Copyright 2004.)

- 5b. If the number of realizations with high scores is found sufficient, this indicates the model does not have any major deficiencies or conceptual problems. Based on the realizations retained in the analysis and deemed acceptable, the regulatory quantity of interest can be calculated and compared to the original model-predicted quantity. This comparison will be presented for reference by decision makers in step 6.

Step 6: Once the model performance has been evaluated per the acceptance criteria, the model sponsors and regulators have to answer the last question in Figure 24.1. This question will determine whether the validation results meet the regulatory objectives or not. This is the trigger point that could lead to significant revision of the original model.

- 6a. If the answer to the question is no, then the left-hand side path in Figure 24.1 begins, and a new iteration of model development will begin using the original data plus data collected for validation. Steps 1 to 6 will be eventually repeated.
- 6b. If the answer to the question is yes, validation is deemed sufficient and the model is considered adequate or robust and the process proceeds to step 7, which starts the development of the long-term monitoring program for the site.

It can be seen and expected that the process of validating a site-specific groundwater model is not an easy one. Throughout the structured process described above, there may be a desire to confirm that the work is on the right track. The way to this confirmation is the cumulative knowledge gained from the different stages of the validation process. That is, a set of independent tests and evaluations will provide knowledge about the model performance, and the test results will provide some incremental, but additive, pieces of information that will be of importance. While there are no guarantees of success (attaining a conclusive outcome about model performance), the combined presence of these different results and evaluations sharply improves the odds that one can make a good decision about the model performance.

Two aspects of the validation process are explained in detail in the following subsections. These are (1) the criteria for determining the sufficiency of the number of acceptable realizations, which are demonstrated for the Project Shoal Model (PSA) model, and (2) the reduction of uncertainty of model input parameters, which is demonstrated for the Amchitka model. Other aspects of this validation approach need further analysis and development. The work on developing a model validation process is still at its infancy and the hope is that more attention will be devoted to the development of rigorous approaches for conducting groundwater model validation processes.

24.4 Acceptance Criteria for Stochastic Model Realizations and Application to PSA Model

The Project Shoal Area (PSA), about 50 km southeast of Fallon, Nevada, is the location of the Shoal underground nuclear test. The nuclear device was emplaced 367 m below ground surface, about 65 m below the water table, in fractured granite. Details of the geology and hydrogeology of the site are provided in Pohll et al. (1999).

Ongoing environmental remediation efforts at the PSA have successfully progressed through numerous technical challenges related to the substantial uncertainties present when characterizing a heterogeneous subsurface environment. The original Corrective Action Investigation Plan (CAIP) for the PSA described a plan to drill and test four characterization wells, followed by flow and transport modeling. The data analysis and the first modeling effort are described in Pohll et al. (1999). After evaluating the modeling effort, the model sponsor determined the degree of uncertainty in transport predictions for PSA remained unacceptably large. Thus, a second CAIP was developed prescribing a rigorous analysis of uncertainty in the PSA model and quantification of methods for reducing uncertainty by collecting additional data. This analysis formed the basis for a second major characterization effort, where four additional characterization wells were drilled during summer and fall of 1999. A key component of the second field program was a tracer test conducted between two of the new four wells, which is described in Reimus et al. (2003).

Results of the tracer test and new characterization efforts produced a new groundwater flow and radionuclide transport model that was approved by the model sponsor and regulators (Pohlmann et al., 2004). The next step in the environmental management of the PSA is to start the validation process and develop the long-term monitoring network. This section focuses on the selection of acceptance criteria and how they can be applied for PSA model. An attempt is made to develop and test some acceptance criteria to be used through the model validation process. It is important to note that this analysis is still in the development stage and the application here is made for hypothetical cases (i.e., hypothetical validation data). Field data will be collected for the validation process of the PSA model at the time when this book will be in press.

24.4.1 Proposed Acceptance Criteria for PSA

According to the validation approach shown in Figure 24.1, the first set of analyses using the field data collected for validation purposes will yield results that will be evaluated to determine the path forward. The first “if” statement in the validation approach pertains to whether a sufficient number of realizations attained satisfactory scores on how they represent the field data used for calibration (old) and for validation (new). The determination of whether a sufficient number exists will be based on five criteria with the decision made in a hierarchical manner as will be discussed later. The five criteria are summarized below:

1. Individual realization scores ($S_j, j = 1, \dots, \text{number of realizations}$), obtained based on how well each realization fits the validation data, will be evaluated. The first criterion then becomes the percentage of these scores, P_1 , that exceeds a certain reference value.
2. The number of validation targets where field data fits within the inner 95% of the probability density function (pdf) of these targets as used in the model is the second criterion, P_2 .
3. The results of hypothesis testing to be conducted using the stochastic perturbation approach of Luis and McLaughlin (1992) as described in detail in Hassan (2004b), P_3 , represent the third criterion.
4. The results of linear regression analysis and other hypothesis testing (e.g., testing error variance based on calibration data and based on validation data) that could be feasible (depending on the size of data set obtained in the field), P_4 , are considered the fourth criterion.
5. The results of the correlation analysis where the log-conductivity variance is plotted against the head variance for the targeted locations and the resulting plot for the model is compared against the field validation data, P_5 .

The hierarchical approach to make the above determination is described by a decision tree. This decision tree for the acceptance of the realizations and for passing the first decision point on the validation approach is shown in Figure 24.2. First S_j is evaluated to determine whether the percentage of realizations with scores above the reference value, P_1 , is more than 40%, between 30% and 40%, or less than 30%. If the number is more than 40%, it is deemed sufficient. If it is between 30% and 40% or less than 30%, the second criterion, P_2 , is used as shown in Figure 24.2.

The second criterion represents the number of validation targets where the field data lie within the inner 95% of the pdf for that target as used in (input) or produced by (output) the model. Then if P_1 is between 30% and 40% and P_2 is between 40% and 50% or if P_1 is less than 30% but P_2 is greater than 50%, the number of realizations is deemed sufficient. If P_1 is less than 30% and P_2 is less than 40%, then the remaining three measures, P_3 , P_4 , and P_5 , are used to determine whether the model needs revision or whether more realizations can be generated to replace some of the current realizations. In this latter case, the model may be conceptually good but the input parameter distribution is skewed or inappropriate and by generating more realizations and keeping the ones fitting the above criteria, the distribution attains the proper position. This can be done using the existing model without conditioning or using any of the new validation data (i.e., no additional calibration). The rationale for selecting the above thresholds (30–40% for P_1 and 40–50% for P_2) is described through an example and when these metrics are evaluated with statistical hypothesis testing later in this section.

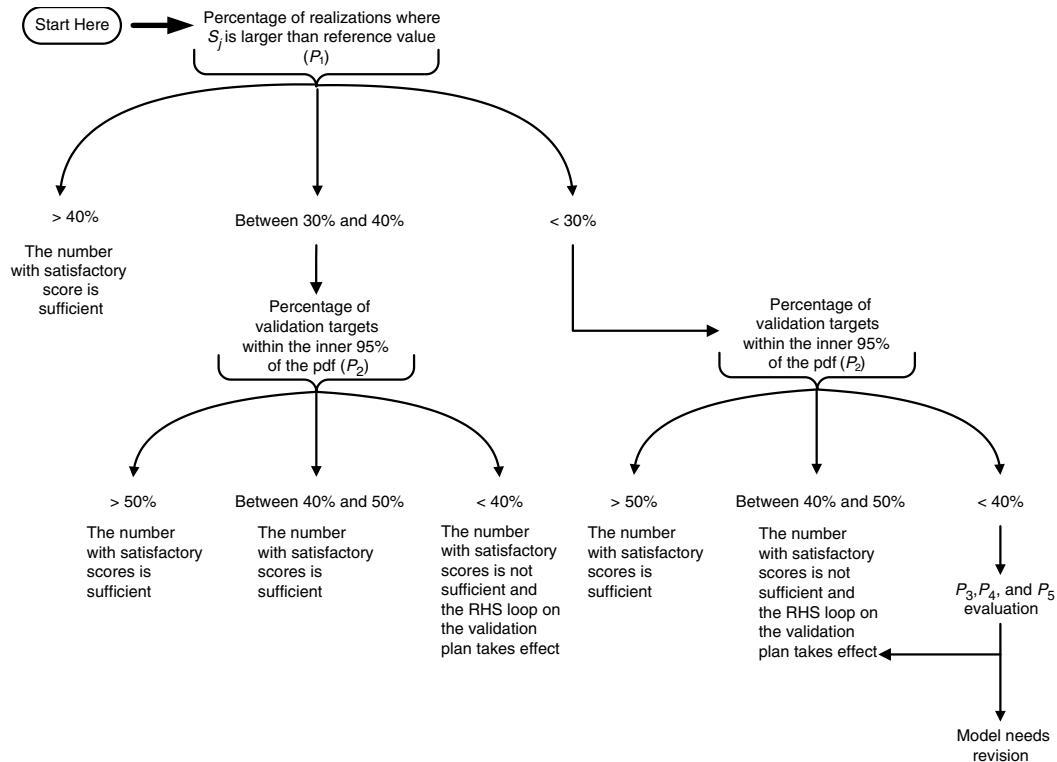


FIGURE 24.2 A decision tree chart showing how the first decision (step 6) in the validation process will be made and the criteria for determining the sufficiency of the number of acceptable realizations.

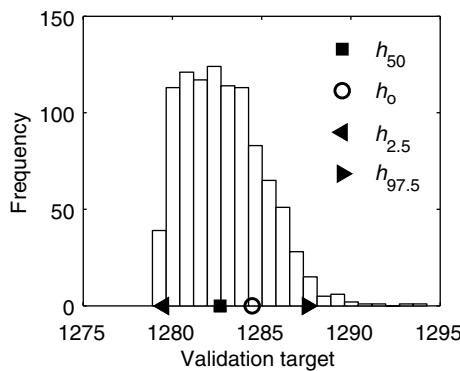


FIGURE 24.3 The head distribution as obtained from the PSA model with the 2.5th, 50th, and 97.5th percentiles shown with the *solid symbols* and the hypothesized field data shown by the *open circle*.

24.4.2 Single Validation Target Illustration

The first criterion is to compute the number of realizations with scores S_j above a reference value. To demonstrate how this reference value is computed, assume we only have one validation target (e.g., the head measurement in one interval in a well). Figure 24.3 shows the head distribution as produced by the stochastic PSA model where the solid symbols represent the 2.5th, 50th, and 97.5th percentiles and the open circle indicates a hypothesized field measurement, h_0 . The reference value and the score for any

individual realization for this simple case can be computed as

$$\text{Reference value}(RV) = \begin{cases} \exp\left[-\frac{(h_o - h_{2.5})^2}{(h_{97.5} - h_{2.5})^2}\right] & \text{for } h_o < h_{50} \\ \exp\left[-\frac{(h_o - h_{97.5})^2}{(h_{97.5} - h_{2.5})^2}\right] & \text{for } h_o > h_{50} \end{cases} \quad (24.1)$$

$$\text{Realization score } (S_j) = \exp\left[-\frac{(h_o - h_j)^2}{(h_{97.5} - h_{2.5})^2}\right] \text{ for } j = 1, \dots, \text{NMC} \quad (24.2)$$

$$P_1 = \frac{\# \text{ of Realizations where } S_j > RV}{\text{NMC}} \quad (24.3)$$

where j is the realization index and it varies from 1 to NMC (number of Monte Carlo realizations) with NMC being 1000 realizations for the PSA model. This leads to all realizations with absolute errors smaller than $\min\{|(h_o - h_{2.5})|, |(h_o - h_{97.5})|\}$, attaining a score higher than the reference value. Figure 24.4 shows the resulting scores and how they compare to the reference value, RV as obtained from the above equations.

It can be seen from Equations 24.1–24.3 that the maximum value RV or S_j can attain is 1.0. Thus if the observed value, h_o , is equivalent to the 2.5th or the 97.5th value, P_1 becomes zero because RV becomes 1.0 and all S_j values will be less than 1.0. Also, if the observed value is found to be less than $h_{2.5}$ or greater than $h_{97.5}$, P_1 will be automatically set to zero. In such cases, one may conclude the model output is skewed toward higher or lower values than indicated by field data. However, this does not necessarily indicate conceptual problems and it may be an indication of incorrect input parameter distributions. The other tests and evaluations can help identify the reasons for this output skewness. When the measured value coincides with the 50th percentile of the target output, h_{50} , then P_1 will approximately be 95% indicating that 95% of the realizations attained scores higher than RV .

24.4.3 Testing the Efficacy of P_1 for a Single Validation Target

To investigate the P_1 metric for the case of a single validation target, a distribution form for the model output is assumed. For simplicity, it is assumed that the model predictions follow a standard normal distribution with zero mean and unit variance, so $h_{50} = 0.0$, $h_{2.5} = -1.96$, and $h_{97.5} = 1.96$. The performance of this metric is tested for a range of measurement values (hypothesized values for the single field data point) between -10.0 and $+10.0$. For each one of these hypothesized values, the RV can be obtained according to Equation 24.1 and the results are shown in Figure 24.5. The RV metric decreases rapidly as the observation value approaches the median, h_{50} . When the measured value lies outside the middle 95% of the output distribution (i.e., outside the range $[-1.96, 1.96]$), RV is not computed as P_1 becomes zero. Also, as shown in the figure, when h_o equals -1.96 ($h_{2.5}$) or 1.96 ($h_{97.5}$), RV equals 1.0. Due to the exponential form in Equation 24.2, all S_j values will be less than 1.0 resulting in a zero value for P_1 when h_o is at the 2.5th or 97.5th percentile.

The next step is to calculate the S_j score for each Monte Carlo realization, with S_j being a similar measure to the RV , but using individual realization predictions for the head. The S_j score is compared to the RV score and the relative number of S_j values that exceed the RV are tallied to obtain P_1 . The S_j values and the corresponding P_1 value were tallied for a range of single observation values in the range $[-10, 10]$ as shown in Figure 24.6.

Figure 24.6 also compares the P_1 metric to the t -distribution with one degree of freedom. The t -distribution is commonly used to test the statistical differences among means when the variance of the distribution is not known. The distribution plotted with the solid line in the figure simply shows the

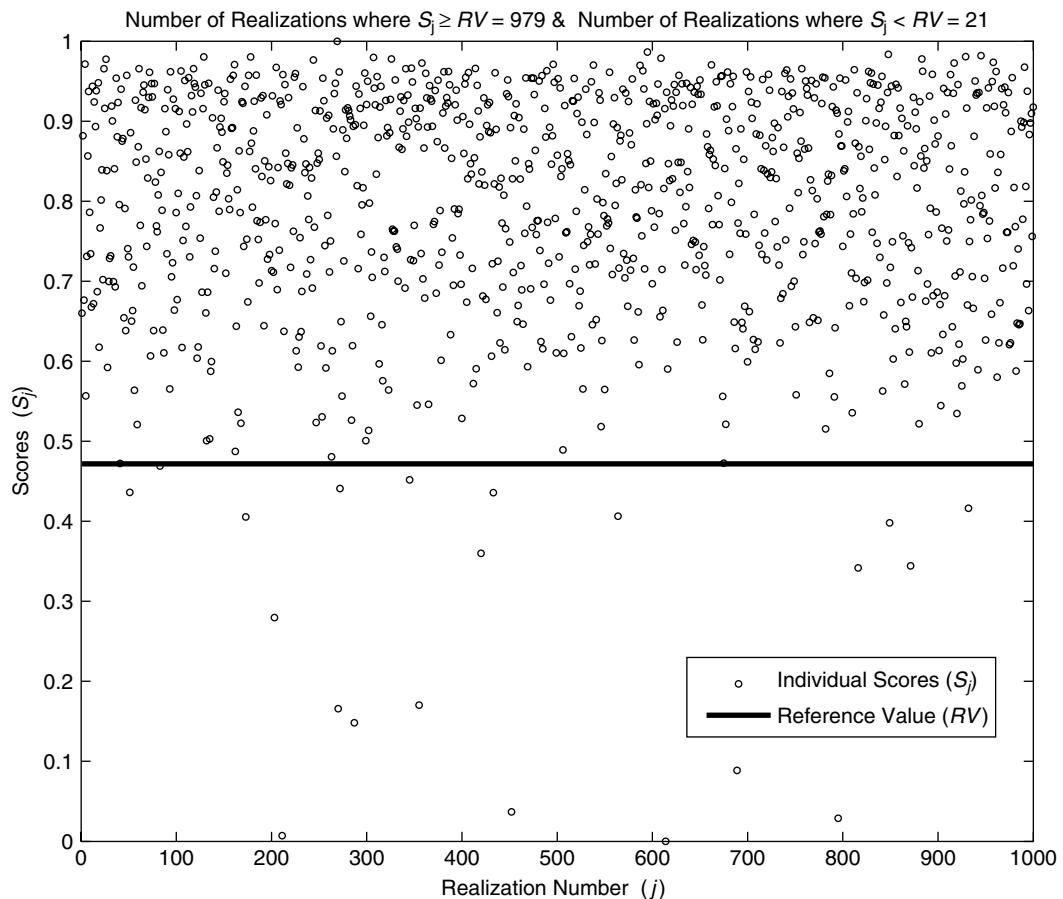


FIGURE 24.4 Realizations scores, S_j , relative to the reference value, RV for the single validation target case. The P_1 value here is 97.9% ($= 979/1000$).

value of the significance level, α , at which each observation on the range $[-10, 10]$ would be rejected in a hypothesis testing evaluating the statistical difference between the mean of the model output (assumed standard normal distribution) and each observed value (assuming each observed value represents a distribution with only one [$n = 1$] sample). The one-degree of freedom used in this plot is not exactly correct as the degrees of freedom are actually $n - 1 = 0$.

To avoid this limitation, we employ the Z -test that is commonly used for the same purpose, but it assumes the variances of the distributions are known. It is assumed that each observation is a mean of a normal distribution and each output realization represents a mean of a normal distribution. For each observation value, we then test the following hypothesis:

$$\begin{aligned} H_0: h_j = h_o &\quad \text{for } j = 1, \dots, \text{NMC} \\ H_1: h_j \neq h_o &\quad \text{for } j = 1, \dots, \text{NMC} \end{aligned} \tag{24.4}$$

Then the proportion of Monte Carlo realizations where the null hypothesis, H_0 , above is not rejected is plotted against each observation value as shown with the dashed line in Figure 24.6.

The plots in Figure 24.6 provide an indication of how the P_1 test compares against standard statistical tests. According to the figure, one would accept all model realizations for any of the observed values $[-10, 10]$ based on the student t -test. In other words, if the t -test is used, one would not reject any of the

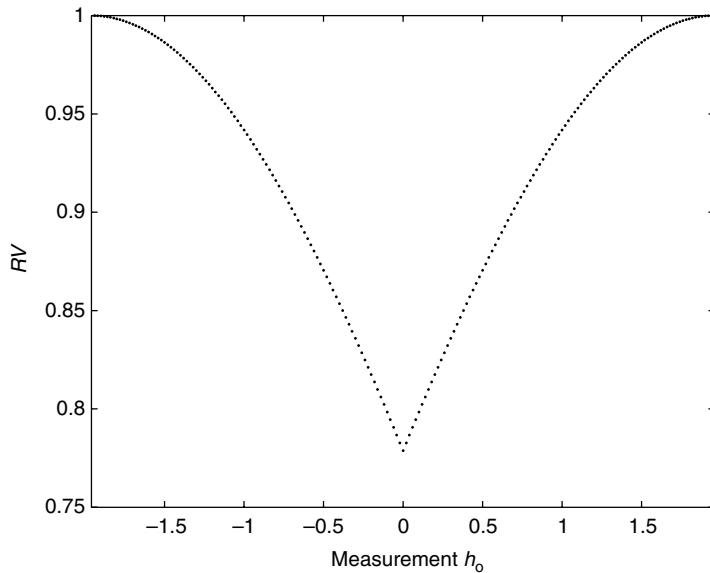


FIGURE 24.5 The reference value, RV for the single validation target case as a function of the measured field value.

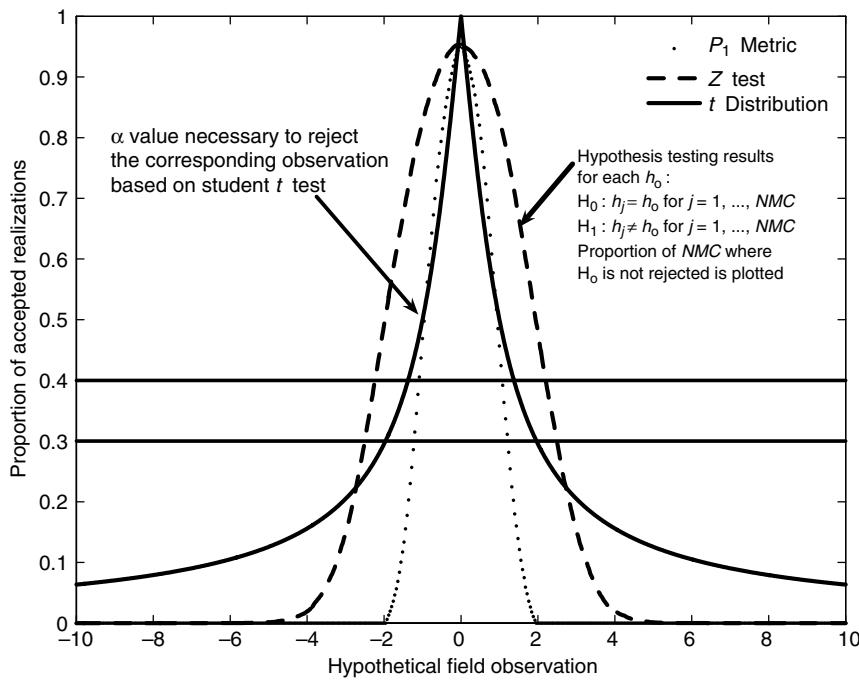


FIGURE 24.6 The P_1 metric, student t distribution, and the results of hypothesis testing using the Z -test.

model realizations until approximately the absolute value of the observation is well above 10 (at the 95% confidence level). On the other hand, the P_1 measure and the Z -test both indicate decreasing proportions of acceptable realizations as one deviates from the median of the model output distribution which is zero in this test case. At the 5% significance level and if the observed value coincides with the median of the model output, only 95% of the realizations are deemed acceptable using the P_1 measure and the Z -test.

When the observed value deviates from the median, the proportion of acceptable realizations drops faster using the P_1 measure compared to the Z -test. For example, 40% or more of the model realizations would be accepted using the Z -test for any observation value in the range $[-2.22, 2.22]$, whereas the P_1 measure gives this level of acceptance for a narrower range of observation values $[-1.07, 1.07]$.

At first glance it appears that the two methods (the P_1 measure versus the Z -test or the t -test) are in large disagreement. But Type I error (rejecting a model realization when in fact it is a good one) versus Type II error (accepting a poor model realization) must be considered. The P_1 metric is essentially reducing Type II error at the expense of Type I error. As discussed by Sargent (1990), the probability of Type I error is called model builder's risk, whereas the probability of Type II error is called model user's risk, and in model validation, model user's risk is extremely important and must be kept small. As a result, it is believed that the restrictiveness of the P_1 measure helps minimize Type II error and thus reduce the model user's risk (both model sponsor and regulators) at the expense of increasing model builder's risk (supposedly the research team constructing the model).

24.4.4 Multiple Validation Targets Illustration

For the general case of having N validation targets, the above equations should be modified to account for these different validation targets. In this case, the RV and the individual scores, S_j , will depend on the sum of squared deviations between each observation, h_o , and the corresponding $h_{2.5}$ or $h_{97.5}$. The equations thus become

$$RV = \exp \left(- \sum_{i=1}^N \min[(h_{o_i} - h_{2.5_i})^2, (h_{o_i} - h_{97.5_i})^2] \Big/ \sum_{i=1}^N [h_{97.5_i} - h_{2.5_i}]^2 \right) \quad (24.5)$$

$$S_j = \exp \left(- \sum_{i=1}^N [h_{o_i} - h_j]^2 \Big/ \sum_{i=1}^N [h_{97.5_i} - h_{2.5_i}]^2 \right) \quad \text{for } j = 1, \dots, \text{NMC} \quad (24.6)$$

For demonstration purposes and as an example, assume the hypothetical case that data are collected on 18 validation targets. These, for example, could be conductivity data in three wells, three measurements each (i.e., nine intervals) and head data for the same intervals. For each one of these targets, the current stochastic PSA model provides a distribution of values. It is then assumed that the values of the field data are known (we pick at random one realization to provide an example observation for all targets). Figure 24.7 and Figure 24.8 show the results of this example (Example 24.1) where P_1 is found to be about 76.7%. In this case, we do not check for P_2 and accept the sufficiency of the number of realizations having acceptable scores. Note, however, that if we were to check P_2 , it would be about 94% ($= 17/18$) because 17 data points lie between the 2.5th and the 97.5th percentiles for the corresponding targets.

Using another set of random values to hypothesize the field data, a different result is obtained as shown in Figure 24.9 and Figure 24.10, for Example 24.2. In this case, both P_1 and P_2 are less than 40% (as the number of validation targets where the open circle is between the 2.5th and the 97.5th percentiles is only 2–11%). In this case, the additional hypothesis tests and linear regression evaluations will be performed to assert whether the model needs to be revised or if the parameter distributions need to be modified.

In Example 24.1, the field data values are hypothesized to be equivalent to one of the model realizations. That is, the values of the 18 validation targets are obtained from one single realization and assumed to represent field data collected for the validation analysis. In spite of assuming field values exactly matching one of the model realizations, the P_1 metric was found to be about 76.7%. This value is obviously dependent on which realization is selected. Therefore, we repeated the above example 1000 times with each of the model realizations assumed to represent the field data in one of those times (similar to a jackknife method). The P_1 metric is obtained for these 1000 experiments and its mean value was found to be about 43%. Given the actual field data to be collected for the validation analysis are very unlikely

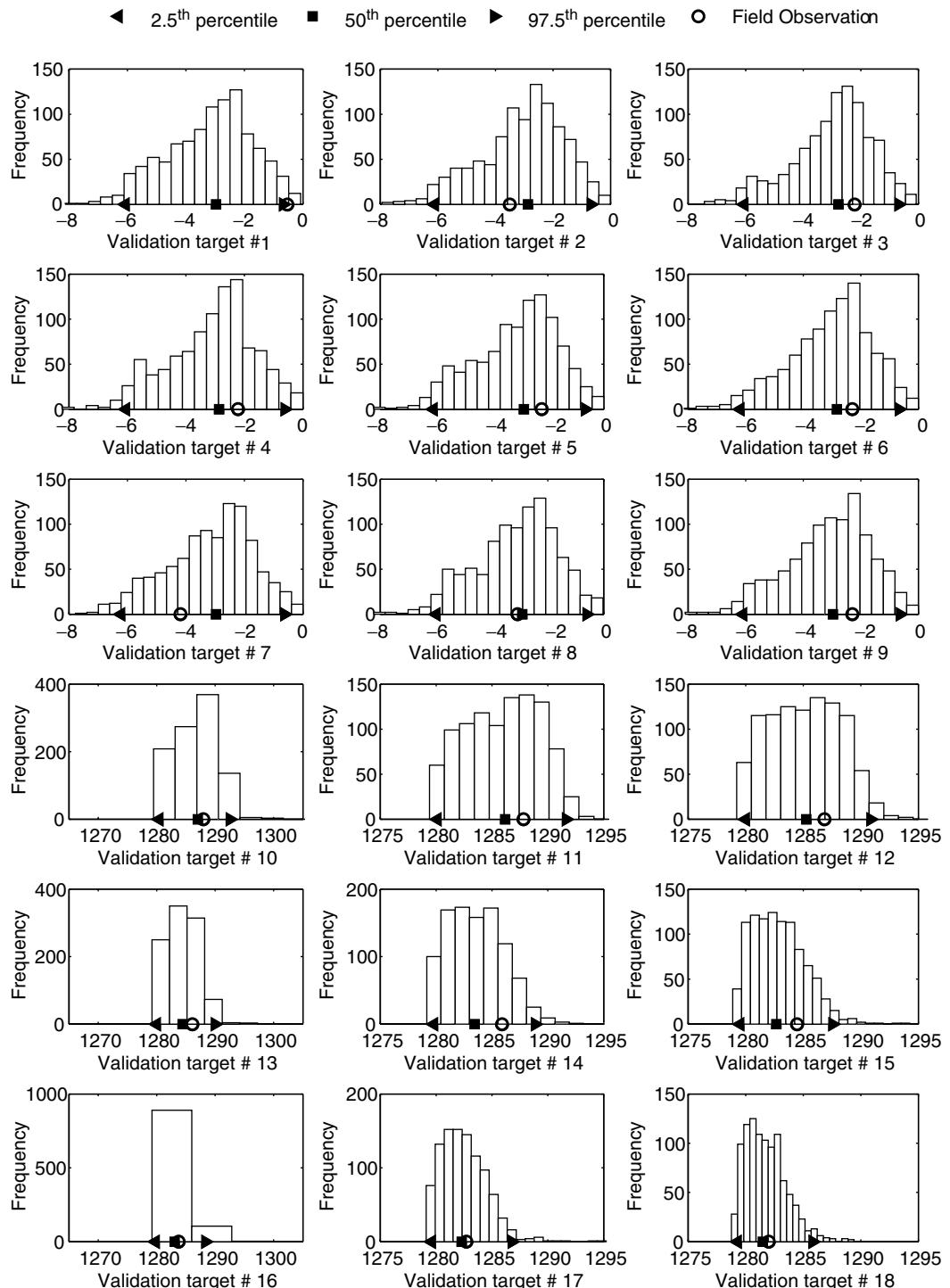


FIGURE 24.7 Example 24.1 showing the distributions for validation targets 1 through 18 with the 2.5th, 50th, and 97.5th percentiles shown with the *solid symbols* and the hypothesized field data shown by the *open circles*.

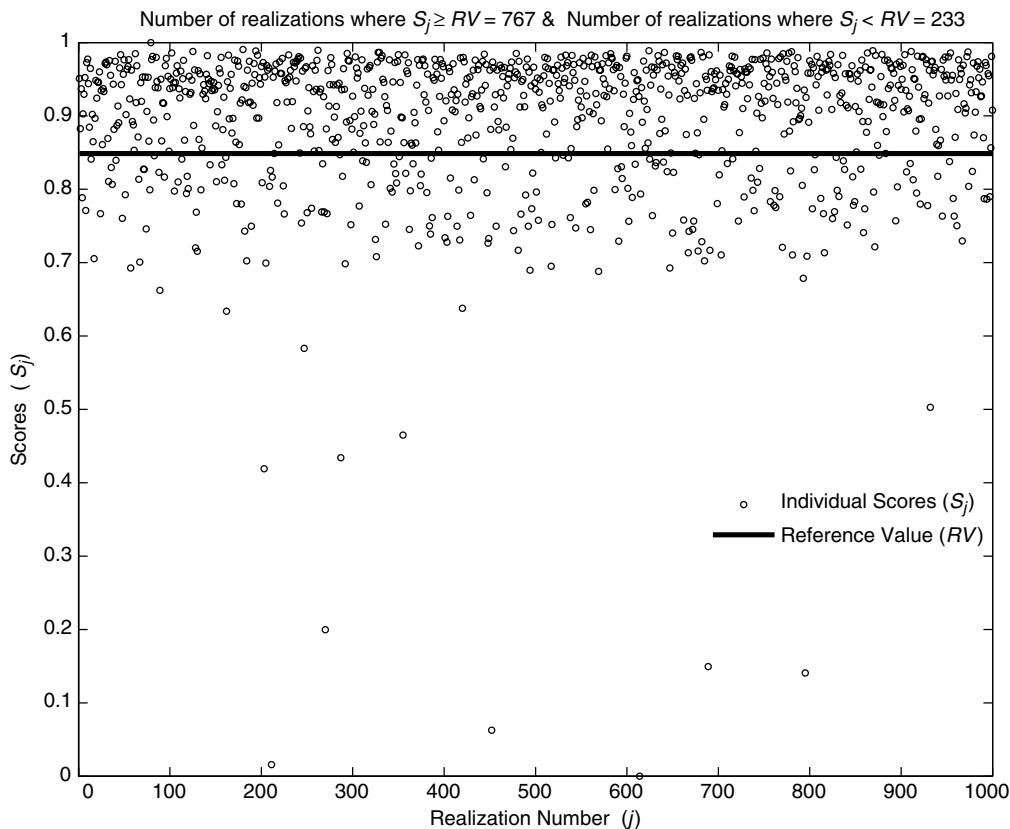


FIGURE 24.8 Example 24.1 showing individual realization scores, S_j , relative to the reference value, RV ; the P_1 value here is about 76.7% (= 767/1000).

to exactly match any of the PSA model realizations, the 30–40% threshold for P_1 is considered realistic. In other words, if one, on average, obtains 43% for P_1 when one of the model realizations is assumed to match real field conditions, one can safely assume the model conceptually valid if P_1 is between 30% and 40% when using the actual validation data.

24.4.5 Testing the Efficacy of P_1 for Multiple Validation Targets

A numerical experiment is performed to evaluate the P_1 metric for the case of multiple validation targets. The experiment is run as follows:

1. A model is assumed to produce multiple outputs, each following a standard normal distribution with zero mean and unit variance.
2. To test the sensitivity of the P_1 metric, 30 observations are randomly selected, with the mean value of the observation set being constant. A range of observation set means is used to determine at what point the model will be rejected. The mean of each observation set is tested over the range -4.0 to 4.0 (i.e., $-4.0, -3.9, \dots, 4.0$).
3. For each mean value, 30 observations are randomly drawn from a normal distribution with the mean equal to the current mean value (i.e., $-4.0, -3.9, \dots, 4.0$) and a standard deviation = 1.0.
4. The RV value for the 30 validation targets is computed using Equation 24.5.
5. For each observation mean, the scores S_j for 10,000 realizations of a model (model is assumed to be standard normal) are computed and the metric P_1 is obtained according to Equation 24.3.
6. Steps 3–5 are then repeated for each observation mean in the range $[-4.0, 4.0]$.

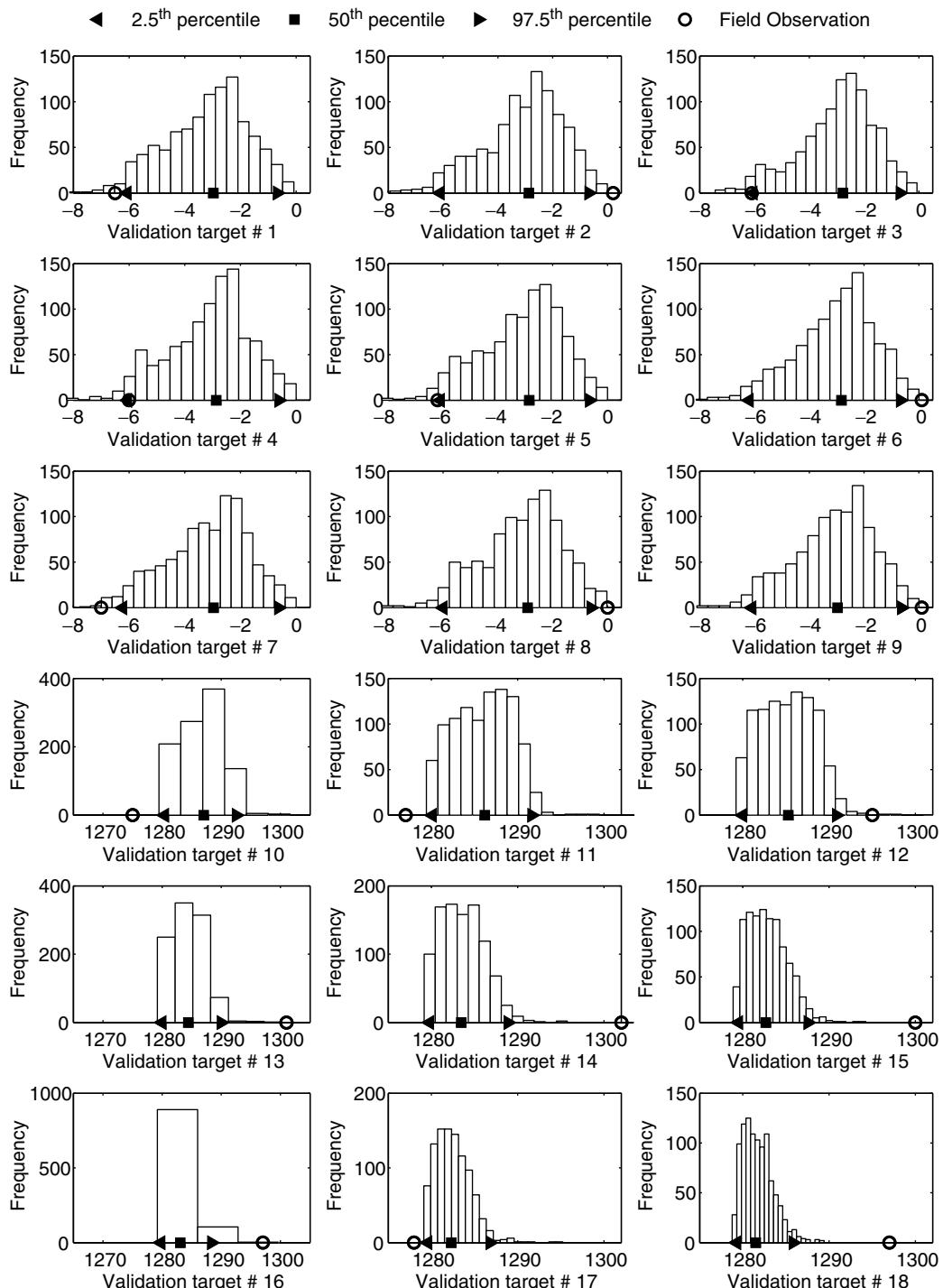


FIGURE 24.9 Example 24.2 showing the distributions for validation targets 1 through 18 with the 2.5th, 50th, and 97.5th percentiles shown with the solid symbols and the hypothesized field data shown by the open circles.

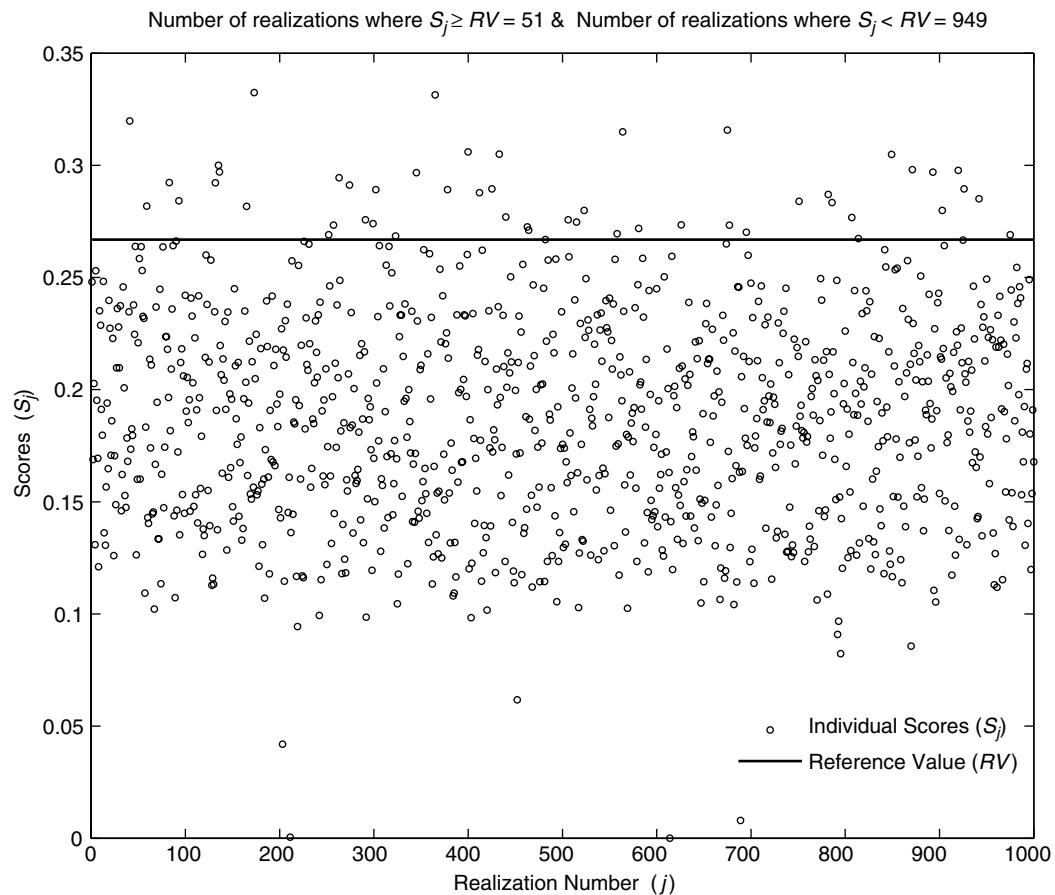


FIGURE 24.10 Example 24.2 showing individual realization scores, S_j , relative to the reference value, RV . The P_1 value here is about 5.1% ($= 51/1000$).

The purpose of this experiment is to determine the point at which a model will be considered invalid. Each observation set represents data that are either close to the model predictions (i.e., mean values close to zero), or poor fitting data with mean values far away from zero. This experiment allows us to compare the rejection region for using a simple hypothesis test (i.e., Z -test) versus the P_1 measure.

Due to the random nature of the distributions generated in the above procedure, we repeated the above experiment 100 times and the average results are shown in Figure 24.11. The dots in the figure represent the results for the P_1 metric, the dashed line shows the results of the Z -test that is similar to the test conducted for the single validation target case, the thin solid line represents the mean value (of 100 values) of the P_1 metric at each observation mean value, and the thick solid line represents a normal distribution that best fits the P_1 results.

For the Z -test, we assume each output realization represent a mean of a normal distribution. For each observation mean value, we then test the following hypothesis:

$$\begin{aligned} H_0: h_j = h_o &\quad \text{for } j = 1, \dots, \text{NMC} \\ H_1: h_j \neq h_o &\quad \text{for } j = 1, \dots, \text{NMC} \end{aligned} \tag{24.7}$$

Then the proportion of Monte Carlo realizations (assumed 10,000 in this experiment) where the null hypothesis, H_0 , above is not rejected is plotted against each observation mean as shown with the thin solid

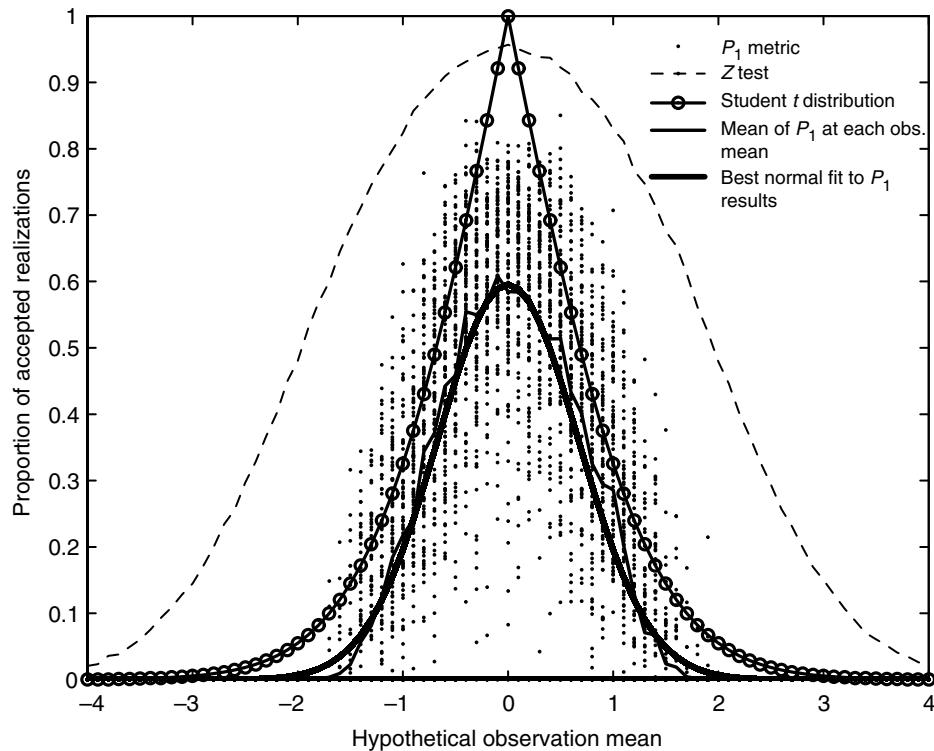


FIGURE 24.11 The P_1 metric (dots), its mean (thin solid line), its best fit normal distribution (thick solid line), student t distribution (line with circles), and the results of hypothesis testing using the Z-test (dashed line) for the multiple validation targets case.

line in Figure 24.11. According to the figure, the t -test would suggest accepting all model realizations if the mean value of the observations is inside the range $[-2.2, 2.2]$ at 95%. The P_1 criterion has a narrower acceptance region ($[-1.6, 1.6]$ according to the thick or thin solid line) again suggesting the P_1 metric is overemphasizing (i.e., trying to reduce) Type II error. Therefore, the P_1 criterion is more stringent than typical hypothesis tests and provides a useful method to test multiple validation targets, which is a more difficult task with standard hypothesis test procedures.

It is important to note that according to P_1 and the Z -test, decreasing proportions of acceptable realizations are obtained as one deviates from the median of the model output distribution (zero in this test case). At a 5% significance level and if the observed mean value coincides with the median of the model output, 95% of the realizations are deemed acceptable using the Z -test, whereas only 60% of the model realizations are deemed acceptable using the P_1 measure. Therefore, a rejection region of less than 30% for the P_1 criteria is very stringent and should not be confused with the 95% confidence interval used for presenting the output uncertainty.

24.4.6 Testing the Efficacy of P_2 for Multiple Validation Targets

A numerical experiment is constructed to test the efficacy of the P_2 metric as follows:

1. A model is assumed to produce output according to a standard normal distribution.
2. Observations are assumed to follow a normal distribution with mean μ and unit variance. The numerical experiment chooses mean values μ from an observation distribution range -4.0 to 4.0 (i.e., $-4.0, -3.9, \dots, 4.0$).

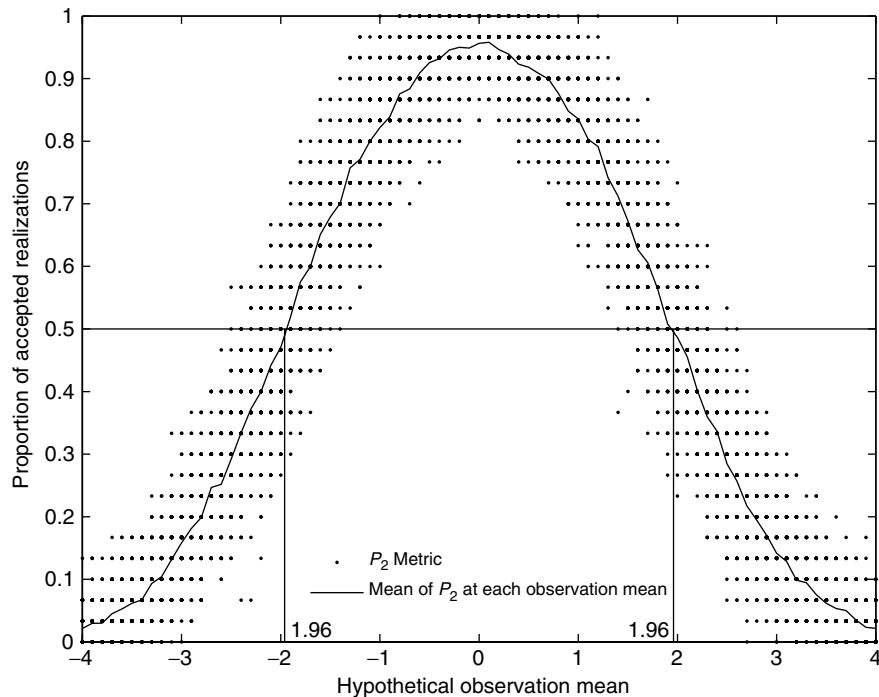


FIGURE 24.12 The P_2 metric (dots) and its mean (solid line) for the multiple validation targets case. The horizontal and vertical lines show that at the 50% threshold, the acceptance region is $[-1.96, 1.96]$ which is the same acceptance region for a standard t test at the 95% confidence level.

3. For each mean value, a random sample of 30 observations is drawn from a normal distribution with the mean equal to the current mean value (i.e., $-4.0, -3.9, \dots, 4.0$) and a standard deviation equal to 1.0.
4. Each of the 30 observations is then compared to the model's distribution $N(0, 1)$ to determine what percentage falls outside of the 95% confidence interval (i.e., -1.96 to 1.96).
5. The process is repeated for all observation means $[-4.0, 4.0]$.

Due to the random nature of the distributions generated in the above procedure, we repeated the above experiment about 100 times and the results are shown in Figure 24.12. The figure shows that if 50% is chosen as the rejection threshold for the P_2 metric, then the model would be accepted for $\mu = [-1.96, 1.96]$. This is a very interesting result as one might initially think 95% should be the acceptance threshold, but 50% yields the same acceptance region as a standard t -test at a 95% confidence level. This warrants further analysis, and as stated earlier, the different aspects of the validation process need more attention for the ultimate goal of arriving at a rigorous set of steps for conducting a model validation process.

24.5 Uncertainty Reduction Using Validation Data and Application to the Amchitka Model

The use of validation data to evaluate input parameter distributions and reduce their uncertainty level is demonstrated for the Amchitka site as a case study. Amchitka is the southernmost island of the Rat Island Group of the Aleutian Island chain extending southwestward from mainland Alaska. The focus here is on one of three underground nuclear tests conducted on the island, a test known as Milrow.

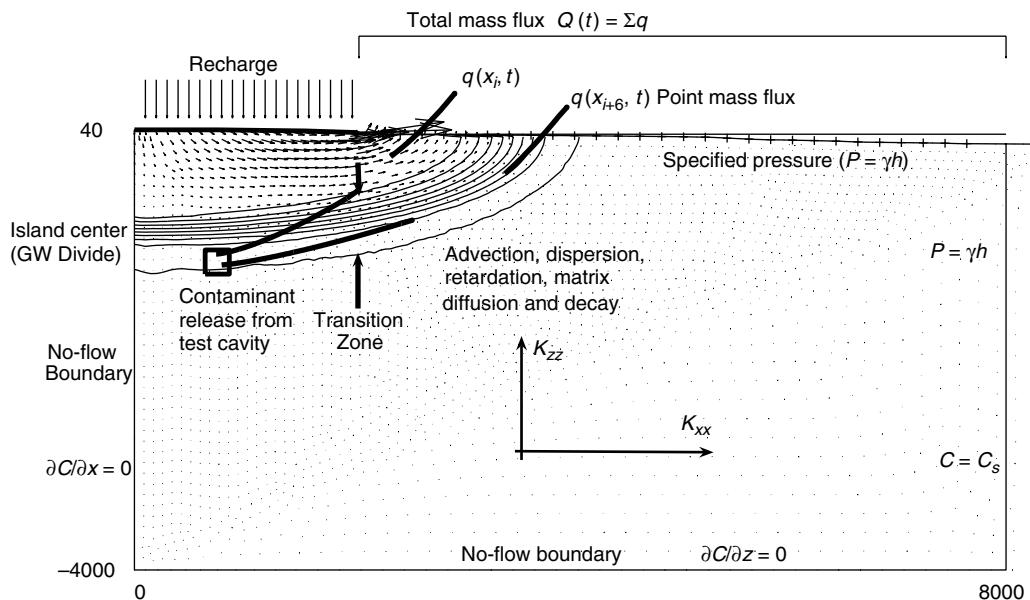


FIGURE 24.13 Milrow model domain showing boundary conditions, simulated processes, test cavity transition zone location and resulting velocity vectors as obtained from one realization.

24.5.1 Model Background

The groundwater flow and radionuclide transport at the Milrow test was modeled using two-dimensional numerical simulations as described in Hassan et al. (2002). The Milrow model involved solving a density-driven flow problem (seawater intrusion problem) to obtain the salinity distribution and the groundwater velocities followed by solving the transport problem to predict radionuclide mass fluxes leaving the groundwater system and discharging into the sea. A multi-parameter uncertainty analysis was adapted and used to address the effects of the uncertainties associated with the definition of the modeled processes and the values of the parameters governing these processes. Details of the modeling efforts at Amchitka can be found in Hassan et al. (2001, 2002) and Pohlmann et al. (2002).

With uncertain parameters, the output of the seawater intrusion problem (i.e., the location of the transition zone between freshwater and saltwater in the groundwater system) varies between model realizations with significant impacts on the solution of the transport equation. This location is mainly determined by the recharge-conductivity ratio. This ratio varies dramatically in the model due to the large parametric uncertainty built into the conductivity and recharge distributions. Figure 24.13 shows the model domain, boundary conditions of the density-driven flow problem, the radionuclide source, transport processes considered and an example transition zone location and velocity vectors as produced by one realization.

Although the analysis conducted in Hassan et al. (2002) conservatively accounted for parametric uncertainty, the need remained for the validation of the numerical groundwater model with an independent data set. The Consortium for Risk Evaluation with Stakeholder Participation II (CRESP II), which represents an Organization of the Institute for Responsible Management, initiated a field expedition in the summer of 2004 for the purpose of collecting data as part of the Amchitka Independent Assessment Science Plan.¹ One of the objectives of this data collection campaign was to provide data to reduce uncertainty about the risk of radionuclide release through the groundwater system to the marine environment.

Of great importance to the groundwater model are the magnetotelluric (MT) measurements for determining the subsurface salinity and porosity structure of the Island, which were recently released in a report

¹<http://www.cresp.org/>

by Unsworth et al. (2005). MT data were collected on profiles that passed through the contaminant source at Milrow. The data presented in Unsworth et al. (2005) showed the presence of a pattern of increasing, decreasing, and increasing resistivity with increasing depth at the site. The depth at which there is an inflection point of the resistivity profile where resistivity begins to decrease is interpreted by Unsworth et al. (2005) as corresponding to the top of the transition zone (TZ) as the salinity increases. The deeper inflection point (increase in resistivity after its decrease with depth) is interpreted as corresponding to the base of the transition zone, as salinity remains constant and the decreasing porosity causes a rise in resistivity (Unsworth et al., 2005).

This independent set of data provides an opportunity for applying the validation process to the groundwater flow model at the Milrow site and reducing the uncertainty in the model parameters and subsequently the model output. To accomplish this, data pertaining to the location of the freshwater lens (from the interpretation of the MT data) could be compared to the groundwater model input distributions as part of the validation process, and then the distributions tightened around the new data for uncertainty reduction.

24.5.2 Bayesian Framework for Uncertainty Reduction

In hydrologic modeling, the uncertainty of parameter estimation needs to be accounted for and the impact on the model output uncertainty needs to be quantified. Bayesian inference provides a framework within which these issues can be addressed with the end product of models being a probability distribution known as the posterior distribution of the model parameters (input) and predictions (output) quantifying uncertainty after data have been collected and incorporated. Although a number of recent studies used this framework for rainfall-runoff modeling and parameter estimation (e.g., Kuczera, 1983; Freer et al., 1996; Kuczera and Parent, 1998; Kuczera and Mroczkowski, 1998; Bates and Campbell, 2001; Marshall et al., 2004), its application to other areas has been limited due to computational difficulties. The advent of Markov Chain Monte Carlo (MCMC) methods has helped address some of these computational difficulties (Marshall et al., 2004).

For certain simple analyses, the posterior distributions (on which inferences are usually made) can be derived analytically, which means they can be written down in standard statistical notation. When this is not possible, a Bayesian solution can still be obtained through the use of simulation methods, such as the MCMC methods. Some details about the Bayesian framework and the MCMC method are presented next in relation to the Milrow model.

For the stochastic groundwater flow model at Milrow, it is assumed that the newly collected data are expressed by the data vector \mathbf{D} . One of the elements in this data vector would be the groundwater salinity profile beneath the island, C , which also represents the steady-state output of the model. One can express the model as

$$C(\mathbf{x}) = M(\mathbf{G}, \mathbf{x}; \boldsymbol{\Theta}) + \varepsilon(\mathbf{x}) \quad (24.8)$$

where $C(\mathbf{x})$ is the observed data at location \mathbf{x} , $M(\mathbf{G}; \boldsymbol{\Theta})$ is the model output for location \mathbf{x} , \mathbf{G} is the set of model input describing domain geometry, boundary conditions, and discretizations, $\boldsymbol{\Theta}$ is the vector of unknown model parameters to be estimated from the data (e.g., hydraulic conductivity, recharge, porosity), and $\varepsilon(\mathbf{x})$ is an error term that is assumed to be a normally distributed random variable having zero mean and variance σ_ε^2 . At this point it is also assumed that the error terms are all mutually independent.

The set of model input \mathbf{G} includes those aspects that do not change from one realization to another. The set of model parameters $\boldsymbol{\Theta}$ includes all the uncertain parameters needed to run the model and that change from one model realization to another. The vector $\boldsymbol{\Theta}$ is treated as a random variable distributed according to a probability density function. This density function expresses our uncertainty about $\boldsymbol{\Theta}$. Before considering the newly collected data, the knowledge about the model parameter set can be summarized in a distribution $P(\boldsymbol{\Theta})$ called the prior distribution.

The posterior distribution of the parameter set, $P(\Theta|C)$ can be obtained through the application of Bayes' theorem

$$P(\Theta|C) = \frac{P(\Theta) P(C|\Theta)}{P(C)} \quad (24.9)$$

where $P(C|\Theta)$ is the likelihood function that summarizes the model relation to the collected data given the parameters used in the model and $P(C)$ is a proportionality constant required so that $P(\Theta|C)$ is a proper density function. It is important to note that the posterior distribution assumes a shape similar to the prior when available data are limited. But when there are large data sets, the posterior distribution will be influenced more by the data than by the assumed prior distribution. Also, the information in the new sample data will dominate the posterior if the prior distribution is selected to represent vague prior knowledge. $P(\Theta|C)$ thus contains all of the available information about Θ coming from both prior knowledge and collected data. Bayesian inference, therefore, reduces to summarizing the posterior density $P(\Theta|C)$.

MCMC sampling explores the posterior distribution by generating a random process (a Markov process) that eventually converges to the stationary, posterior distribution of the parameters. While many different MCMC sampling algorithms may exist, the Metropolis–Hastings algorithm is the most commonly used. Details of this algorithms are beyond the scope of this chapter, but can be found in Bates and Campbell (2001), Campbell et al. (1999), Marshall et al. (2004), Kuczera and Parent (1998), to name a few sources.

24.5.3 Application to the Milrow Site in Amchitka

The validation data at Milrow are obtained from the MT data and the interpretation of the measured MT signals. Specifically, this provided identification of the top and bottom bounds of the transition zone from freshwater to seawater beneath the island. This set of data is used with MCMC to develop the posterior distributions for conductivity, recharge, and conductivity–recharge ratio. The development of the posterior distribution serves to evaluate the original (prior) distribution selection and to reduce the uncertainty in the parameter distributions, both of which are objectives of the validation process.

In the original Milrow model (Hassan et al., 2002), lognormal distributions were assumed for both recharge and conductivity and they were assumed to be uncorrelated. These prior distributions were based on calibrating the model to a set of salinity data that was available from one borehole at the site. This data set clearly defines the increase in salinity with depth to near-seawater concentrations. Therefore, the location of the transition zone separating the shallow freshwater from the deep saltwater provided by this data set guided the selection of the distributions used in the original model. These distributions are used here as the prior distributions for the MCMC analysis and with the help of the MT data, posterior distributions are developed.

Figure 24.14 shows the potential of the MCMC approach to help in the validation process and uncertainty reduction by conditioning on field data. The figure compares the prior distributions for the recharge, conductivity, and the recharge–conductivity ratio to the posterior distributions for these parameters obtained utilizing the prior knowledge and the new MT-related data. It is clear that the posterior distributions are very different than the prior ones, especially for the recharge–conductivity ratio. The interpretation of the MT data indicated a deeper transition zone than what was used in the model. This translated into the posterior distribution of conductivity being skewed to lower values relative to the prior and that of recharge being skewed to the higher values. The end result is a higher recharge–conductivity ratio for which the posterior distribution has a sharp peak and a much smaller range compared to the prior indicating a dramatic reduction in uncertainty of this ratio.

Although the posterior distributions of model input parameters are different than the prior, no parameter value in the posterior range is outside the ranges implemented in the original model. The MCMC tool has the potential to be a very useful tool in the validation process and in particular for uncertainty reduction of model input parameters and consequently model output.

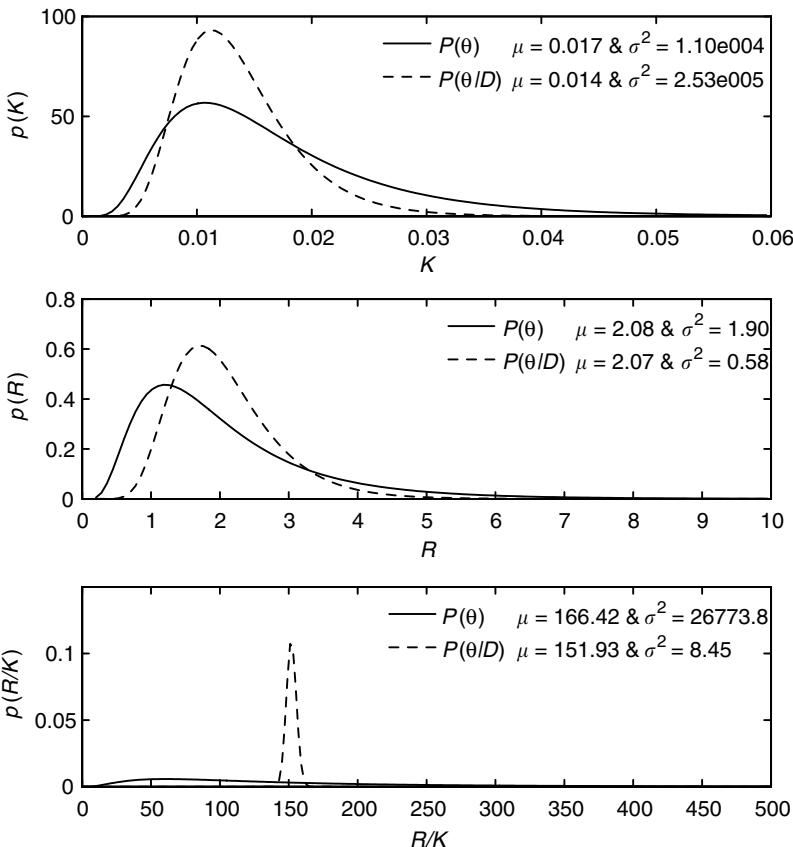


FIGURE 24.14 Comparison between the prior (solid lines) and posterior (dashed lines) distributions obtained using validation data and the MCMC approach for conductivity, K , recharge, R , and recharge-conductivity ratio.

24.6 Summary and Conclusions

Models have an inherent uncertainty. The difficulty in fully characterizing the subsurface environment makes uncertainty an integral component of groundwater flow and transport models, which dictates the need for continuous monitoring and improvement. Building and sustaining confidence in closure decisions and monitoring networks based on models of subsurface conditions require developing confidence in the models through an iterative process.

The definition of model validation is postulated as a confidence-building and long-term iterative process (Hassan, 2004a). Model validation should be viewed as a process, and not an end result.

Following Hassan (2004b), an approach is proposed for the validation process of stochastic groundwater models. The approach is briefly summarized here and detailed analyses of acceptance criteria for stochastic realizations and of using validation data to reduce input parameter uncertainty are presented and applied to two case studies.

During the validation process for stochastic models, a question arises as to the sufficiency of the number of acceptable model realizations (in terms of conformity with validation data). We propose to use a hierarchical approach to make this determination. This approach is based on computing five measures or metrics and following a decision tree to determine if a sufficient number of realizations attain satisfactory scores regarding how they represent the field data used for calibration (old) and for validation (new).

The first two of these measures are applied to hypothetical scenarios using the first case study and assuming field data are consistent with the model or significantly different than the model results. In

both cases it is shown how the two measures would lead to the appropriate decision about the model performance. Standard statistical tests are used to evaluate these measures with the results indicating they are appropriate measures for evaluating model realizations.

The use of validation data to constrain model input parameters is shown for the second case study using a Bayesian approach known as Markov Chain Monte Carlo. The approach shows a great potential to be helpful in the validation process and in incorporating prior knowledge with new field data to derive posterior distributions for both model input and output.

Glossary

Model: An abstraction or a simple representation of a real system or process.

Conceptual Model: A hypothesis for how a system or a process operates.

Mathematical Model: A quantitative expression of the system processes, forces, and events.

Computer Code: An algorithm to implement the mathematical model and perform the model computations.

Generic models: The computer codes that are used to solve the mathematical flow and transport equations.

Site-Specific Models: The computer codes combined with the conceptual models (model structure), input data and boundary conditions for a particular geographical area.

Research or Analysis Models: Models used for studying and understanding different phenomena in the subsurface and they usually rely on hypothetical domains or very well characterized field sites.

Predictive or Decision-making Models: Models that are mainly used to support and aid a regulatory decision regarding a subsurface issue.

Model Calibration: The process of tuning the model to identify the independent input parameters by fitting the model results to some field data or experimental data, which usually represent the dependent system parameters.

History Matching: Using historical field data during the model building and calibration process and trying to match them before using the model to predict future system response.

Computer Code Verification: Verification of a mathematical model or its computer code; it is obtained when it is shown that the model behaves as intended and that the equations are correctly encoded and solved.

Model Verification: A process aimed at establishing a greater confidence in the model by using a set of calibrated parameter values and stresses to reproduce a second set of field data that is available during the model-building process.

Model Validation: A process, not an end result, aimed at building confidence in the model predictions through a structured set of tests and evaluations with trigger mechanisms that forces the process back to the model-building stage if model is not consistent with field data collected for validation.

Research Model Validation: Validation of a research model that helps one understand processes, uncertainties, and the like.

Predictive Validation: Performing a validation process for a site-specific, predictive model for the goal of helping the decision-making process.

Validation Target: The model output for which field data will be collected for validation purposes.

Bayesian Inference: An alternative to the classical approach to statistical analysis that benefits from prior knowledge as well as newly collected data to quantify parameter uncertainty.

Markov Chain Monte Carlo (MCMC): A form of Bayesian inference that can be used to simulate the posterior distribution of model parameters.

Prior Distribution: The parameter distribution that summarizes all the knowledge about the parameter before collecting new data.

Posterior Distribution: The parameter distribution that is obtained through Bayesian inference and it relies on both prior knowledge and newly collected data.

Likelihood Function: A probability function describing the probability of seeing the observed data (newly collected data) given that a certain model or assumption is true.

Density-driven Flow: A groundwater flow problem in which the flow depends on (or is driven by) the variations in groundwater density (e.g., due to thermal effects or salinity effects).

Transition Zone: The varying salinity zone separating shallow freshwater (mainly from recharge) from deep saltwater (from the ocean or the sea).

Acknowledgments

The author would like to express his special thanks to Greg Pohll, Jenny Chapman, and Karl Pohlmann for very helpful comments and stimulating discussions and also for the help on the statistical evaluation of the acceptance criteria. Appreciation is extended to Hesham Bekhit for his help on the Metropolis Algorithm. This work was performed under the auspices of the U.S. Department of Energy by the Desert Research Institute under contract number DE-AC08-00NV13609.

References

- Anderson, M.G. and Bates P.D. (2001) *Model Validation: Perspectives in Hydrological Science*. New York, NY: John Wiley & Sons, Ltd.
- Andersson, K., Grundfelt, B., Larsson, A. and Nicolson, T. (1989) INTRAVAL as an integrated international effort for geosphere model validation — A status report. *Proceedings of Symposium on Safety Assessment of Radioactive Waste Repositories*, Paris, October 9–13, OECD Nuclear Energy Agency.
- AWR (1992a) Special issue: Validation of Geo-Hydrological Models — Part 1. *Advances in Water Resources* 15, 1–87.
- AWR (1992b) Special issue: Validation of Geo-Hydrological Models — Part 2. *Advances in Water Resources* 15, 153–217.
- Balci, O. (1988) Credibility assessment of simulation results: The state of the art, methodology and validation. *Simulation Series, The Society for Computer Simulation* 19, 19–25.
- Balci, O. (1989) How to assess the acceptability and credibility of simulation results. In *Proceedings of the 1989 Winter Simulation Conference*, Washington, DC. Edited by E. MacNair, K. Musselman, and P. Heidelberger.
- Balci, O. and Sargent, R.G. (1981) A methodology for cost-risk analysis in the statistical validation of simulation models. *Communication of the ACM* 24, 190–197.
- Balci, O. and Sargent, R.G. (1982) Validation of multivariate response simulation models by using Hotelling's two-sample T^2 test. *Simulation* 39, 185–192.
- Balci, O. and Sargent, R.G. (1984) A bibliography on the credibility, assessment and validation of simulation and mathematical models. *Simuletter* 15, 15–27.
- Bates, B. C. and Campbell, E. P. (2001) A Markov chain Monte Carlo scheme for parameter estimation and inference in conceptual rainfall runoff modeling. *Water Resources Research* 37, 937–947.
- Beljin, M.S. (1988) Testing and validation of models for solute transport in groundwater: Code intercomparison and evaluation of validation methodology. International Ground Water Modeling Center, Holcomb Institute, Butler Univ., Indianapolis, Indiana. Report GWMI 88–11.
- Bredehoeft, J.D. and Konikow, L.F. (1992) Reply to comment. *Advances in Water Resources* 15, 371–372.
- Bredehoeft, J.D. and Konikow, L.F. (1993) Ground water models: Validate or invalidate. *Ground Water* 31, 178–179.
- Bredehoeft, J.D. (2003) From models to performance assessment: The conceptualization problem. *Ground Water* 41, 571–577.
- Broyd, T., Read, D. and Come, B. (1990) The CHEMVAL Project: An international study aimed at the verification and validation of equilibrium speciation and chemical transport computer programs.

- In *Proceedings of GEOVAL90 Symposium*, Stockholm, May 14–17, 1990, Swedish Nuclear Power Inspectorate (SKI), Stockholm.
- Campbell, E. P., Fox, D. R. and Bates, B. C. (1999) A Bayesian approach to parameter estimation and pooling in nonlinear flood event models. *Water Resources Research* 35, 211–220.
- de Marsily, G., Combes, P. and Goblet, P. (1992) Comment on ‘Groundwater models cannot be validated,’ by L.F. Konikow and J. D. Bredehoeft. *Advances in Water Resources* 15, 367–369.
- Duan, Q., Gupta, H.V., Sorooshian, S., Rousseau, A.N. and Turcotte, R. (Eds.) (2003) Calibration of watershed models, *Water Sci. Appl. Ser.*, Vol. 6, AGU, Washington, D.C.
- Eisenberg, N., Federline, M., Sagar, B., Wittmeyer, G., Andersson, J. and Wingefors, S. (1994) Model validation from a regulatory perspective: A summary. In *GEOVAL 94, Validation Through Model Testing*, Proceedings of an NEA/SKI Symposium, Paris, France, 11–14 October, pp. 421–434.
- Freer, J., Beven, K. and Ambroise, B. (1996) Bayesian estimation of uncertainty in runoff prediction and the value of data: An application of the GLUE approach. *Water Resources Research* 32, 2161–2173.
- Gass, S.I. (1983) Decision-aiding models: Validation, assessment, and related issues for policy analysis. *Operations Research* 31, 601–663.
- Gass, S.I. and Thompson, B.W. (1980) Guidelines for model evaluation: An abridged version of the U.S. General Accounting Office exposure draft. *Operations Research* 28, 431–479.
- GEOVAL87 (1987) *Proceedings of Symposium on Verification and Validation of Geosphere Performance Assessment Models*, organized by Swedish Nuclear Power Inspectorate, Stockholm, Sweden, April 7–9.
- GEOVAL90 (1990) *Proceedings of Symposium on Validation Geosphere Flow and Transport Models*, organized by Swedish Nuclear Power Inspectorate, Stockholm, Sweden, May 14–17.
- GEOVAL94 (1994) *Proceedings of Symposium on Verification through Model Testing*, organized by OECD Nuclear Energy Agency and the Swedish Nuclear Power Inspectorate, Paris, France, October 11–14.
- Grundfelt, B., Lindbom, B., Larsson, A. and Andersson, K. (1990) HYDROCOIN level 3 — Testing methods for sensitivity/uncertainty analysis. Proceedings of GEOVAL90 Symposium, Stockholm, May 14–17, 1990. Swedish Nuclear Power Inspectorate, (SKI), Stockholm.
- Grundfelt, B. (1987) The HYDROCOIN Project — Overview and results from level one. Proceedings of International GEOVAL-87 Symposium, Swedish Nuclear Power Inspectorate, (SKI), Stockholm, April 7–9, 1987.
- Hassan, A.E. (2004a) Validation of numerical groundwater models used to guide decision making. *Ground Water* 42, 277–290.
- Hassan, A.E. (2004b) A methodology for validating numerical groundwater models. *Ground Water* 42, 347–362.
- Hassan, A.E., Pohlmann, K., and Chapman, J. (2001) Uncertainty analysis of radionuclide transport in a fractured coastal aquifer with geothermal effects. *Transport in Porous Media* 43, 107–136.
- Hassan, A., Pohlmann, K. and Chapman, J. (2002) Modeling Groundwater Flow and Transport of Radionuclides at Amchitka Island’s Underground Nuclear Tests: Milrow, Long Shot, and Cannikin. Desert Research Institute, Division of Hydrologic Sciences, Publication No. 45172, DOE/NV/11508–51.
- Hassanzadeh, S.M. 1990. Panel discussion. In *GEOVAL-90, Symposium on Validation of Geosphere Performance Assessment Models*, Stockholm, Sweden, 14–17 May, pp. 631–658.
- Herbert, A., Dershowitz, W., Long, J. and Hodgkinson, D. (1990) Validation of fracture flow models in the Stripa project. Proceedings of GEOVAL90 Symposium, Stockholm, May 14–17, Swedish Nuclear Power Inspectorate (SKI).
- Huyakorn, P.S., Kretschek, A.G., Broome, R.W., Mercer, J.W. and Lester, B.H. (1984) Testing and validation of models for simulating solute transport in groundwater. International Ground Water Modeling Center, Holcomb Research Institute, Butler University, Indianapolis, IN. Report GWMI 84–13.
- Hornberger, G.M. and Boyer, E.W. (1995) Recent advances in watershed modeling. *U.S. Natl. Rep. Int. Union Geod. Geophys. 1991–1994, Rev. Geophys.*, 33, 949–957.
- INTRACOIN (1984) Final report level 1, Code verification. Report SKI 84:3, Swedish Nuclear Power Inspectorate, Stockholm, Sweden.

- INTRACOIN (1986) Final report levels 2 and 3, Model validation and uncertainty analysis. Report SKI 86:2, Swedish Nuclear Power Inspectorate, Stockholm, Sweden.
- Konikow, L.F. and Bredehoeft, J.D. (1992) Groundwater models cannot be validated. *Advances in Water Resources* 15, 75–83.
- Konikow, L.F. and Bredehoeft, J.D. (1993) The myth of validation in groundwater modeling. In *Proceedings 1993 Groundwater Modeling Conference*, IGWMC, Golden, CO, pp. A-4.
- Kuczera, G. (1983) Improved parameter inference in catchment models: 1. Evaluating parameter uncertainty. *Water Resources Research* 19, 1151–1162.
- Kuczera, G. and Mroczkowski, M. (1998) Assessment of hydrologic parameter uncertainty and the worth of multiresponse data. *Water Resources Research* 34, 1481–1489.
- Kuczera, G. and Parent, E. (1998) Monte Carlo assessment of parameter uncertainty in conceptual catchment models: The Metropolis algorithm. *Journal of Hydrology* 211, 69–85.
- Luis, S.J. and McLaughlin, D. (1992) A stochastic approach to model validation. *Advances in Water Resources* 15, 15–32.
- Marshall, L., Nott, D. and Sharma, A. (2004) A comparative study of Markov chain Monte Carlo methods for conceptual rainfall-runoff modeling. *Water Resources Research* 40, W02501, doi:10.1029/2003WR002378.
- Miller, R.E. and van der Heijde, P.K.M. (1988) A groundwater research data center for model validation. *Proceedings of the Indiana Water Resources Assoc. Symposium*, June 8–10, Greencastle, IN.
- Moran, M.S. and Mezgar, L.J. (1982) Evaluation factors for verification and validation of low-level waste disposal site models. Oak Ridge National Laboratory, Oak Ridge, TN. Report DOE/OR/21400-T119.
- National Research Council (1990) *Groundwater Models; Scientific and Regulatory Applications*. National Academy Press, Washington, D.C., 303 pp.
- Neuman, S.P. (1992) Validation of safety assessment models as a process of scientific and public confidence building. In *Proceedings of HLWM Conference*, Vol. 2, Las Vegas.
- Nicholson, T.J. (1990) Recent accomplishments in the INTRALVAL Project — A status report on validation efforts. Proceedings of GEOVAL90 Symposium, Stockholm, Sweden, May 14–17, Swedish Nuclear Power Inspectorate (SKI).
- Oren, T. (1981) Concepts and criteria to assess acceptability of simulation studies. A Frame of Reference. *Communication of the ACM* 24, 180–189.
- Oreskes, N., Shrader-Frechette, K., and Belits, K. (1994) Verification, validation, and confirmation of numerical models in the earth sciences. *Science* 264, 641–646.
- Pohll, G., Hassan, A.E., Chapman, J.B., Papelis, C. and Andricevic, R. (1999) Modeling groundwater flow and radioactive transport in a fractured aquifer. *Ground Water* 37, 770–784.
- Pohlmann, K., Hassan, A.E. and Chapman, J. (2002) Modeling density-driven flow and radionuclide transport at an underground nuclear test: Uncertainty analysis and effect of parameter correlation. *Water Resources Research* 38, 10.1029/2001WR001047.
- Pohlmann, K., Pohll, G., Chapman, J., Hassan, A.E., Carroll, R. and Shirley, C. (2004) Modeling to support groundwater contaminant boundaries for the Shoal underground nuclear test. Division of Hydrologic Sciences, Desert Research Institute, Publication No. 45184-revised, pp. 197.
- Refsgaard, J.C. (2001) Discussion of model validation in relation to the regional and global scale. In *Model Validation: Perspectives in Hydrological Sciences*, M.G. Anderson and P.D. Bates (Eds.) John Wiley & Sons, New York.
- Reimus, P., Pohll, G., Mihevc, T., Chapman, J., Haga, M., Lyles, B., Kosinski, S., Niswonger, R. and Sanders, P. (2003) Testing and parameterizing a conceptual model for solute transport in a fractured granite using multiple tracers in a forced-gradient test. *Water Resources Research* 39, 1356–1370.
- Sargent, R.G. (1984) Simulation Model Validation. In *Simulation and Model-based Methodologies: An Integrative View*. Edited by Oren *et al.*, Springer-Verlag.
- Sargent, R.G. (1988) A tutorial on validation and verification of simulation models. *Proceedings of 1988 Winter Simulation Conference*. Edited by M. Abrams, P. Haigh, and J. Comfort, pp. 33–39.

- Sargent, R.G. (1990) Validation of mathematical models. In *GEOVAL-90, Symposium on Validation of Geosphere Performance Assessment Models*, Stockholm, Sweden, 14–17 May, pp. 571–579.
- Schruben, L.W. (1980) *Establishing the Credibility of Simulations — The Art and the Science*. Prentice-Hall.
- Singh, V.P. (1995) *Computer Models of Watershed Hydrology*, 1129 pp, Water Resour. Publ., Highlands Ranch, Colorado.
- Singh, V.P. and Woolhiser, D.A. (2002) Mathematical modeling of watershed hydrology. *Journal of Hydrol. Eng.* 7, 270–292.
- SSI (Swedish National Institute of Radiation Protection) (1990) *Proceedings of Symposium and Workshop on the Validity of Environmental Transfer Models (BIOMOVS)*, October 8–12, 1990. Swedish National Institute of Radiation Protection (SSI), Stockholm.
- Swedish Nuclear Power Inspectorate (1987) The International HYDROCOIN Project — Background and Results. OECD, Paris.
- Swedish Nuclear Power Inspectorate (1990) The International INTRAVAL Project — Background and Results. OECD, Paris.
- Tsang, C.F. (1991) The modeling process and model validation. *Ground Water* 29, 825–831.
- van der Heijde, P.K.M., Huyakorn, P.S. and Mercer, J.W. (1985) Testing and validation of groundwater models. Proceedings of the NWWA/IGWMC Conference on “Practical Applications of Ground Water Models.” August 19–20, Columbus, OH.
- van der Heijde, P.K.M. (1987) Quality assurance in computer simulations of groundwater contamination. *Environmental Software* 2.
- van der Heijde, P.K.M., Elderhorst, W.I.M., Miller, R.A. and Trehan, M.F. (1989) The establishment of a groundwater research data center for validation of subsurface flow and transport models. International Ground Water Modeling Center, Holcomb Research Institute, Butler Univ., Indianapolis, IN. Report GWMI 89-01.
- Vogel, R.M., Sankarasubramanian (2003) Validation of a watershed model without calibration, *Water Resources Research* 39, 1292, doi:10.1029/2002WR001940.
- Unsworth, M., Soyer, W. and Tuncer, V. (2005) Magnetotelluric measurements for determining the subsurface salinity and porosity structure of Amchitka Island, Alaska. In *Biological and Geophysical Aspects of Potential Radionuclide Exposure in the Amchitka Marine Environment*, edited by Powers et al. Consortium for Risk Evaluation with Stakeholder Participation, Institute for Responsible Management, Piscataway, New Jersey.
- Zeigler, B.P. (1976) *Theory of Modeling and Simulation*. John Wiley & Sons, Inc., New York.

Further Information

This chapter attempts to introduce a newly proposed validation methodology for numerical groundwater models cast in a stochastic framework. Research on model validation is extensively reported in the literature, but mostly focuses on process models and definitions of the term. In the area of toxic waste management, a number of authors (e.g., Moran and Mezgar, 1982; Huyakorn et al., 1984; van der Heijde et al., 1985; Van der Heijde, 1987; Beljin, 1988) have considered the question of whether a model used in a safety assessment program is valid in making appropriate long-term predictions. During the late 1980s, an effort was made to establish a groundwater research data center for the validation of subsurface flow and transport models (Miller and Van der Heijde, 1988; Van der Heijde et al., 1989).

In the area of nuclear waste management, the need to validate groundwater models has received increased emphasis. This has led to institutionalized and publicized programs for validation of hydrogeological models. A number of international cooperative projects such as INTRACOIN (1984, 1986), HYDROCOIN (Grundfelt, 1987; Grundfelt et al., 1990), INTRAVAL (Andersson et al., 1989; Nicholson, 1990), STRIPA (Herbert et al., 1990), CHEMVAL (Broyd et al., 1990), BIOMOVS (SSI, 1990) were devoted to the validation of models. Model validation was also extensively discussed in symposia including GEOVAL87 (1987), GEOVAL90 (1990), and GEOVAL94 (1994). The journal *Advances in Water Resources*

dedicated two special issues to the topic of model validation (AWR, 1992a, 1992b). Additionally, a wealth of literature has been published on validation in the field of systems engineering and operations research (Tsang, 1991), some of which may be useful for subsurface model validation. Examples cited by Tsang (1991) include Balci (1988, 1989), Balci and Sargent (1981, 1982, 1984), Gass (1983), Gass and Thompson (1980), Oren (1981), Sargent (1984, 1988), Schruben (1980), and Zeigler (1976).

The Swedish Nuclear Power Inspectorate, SKI, initiated and completed three international cooperation projects to increase the understanding and credibility of models describing groundwater flow and radionuclide transport. The INTRACOIN project is the first of these, and it focused on verification and validation of transport models. The HYDROCOIN study was the second study and represented an international cooperative project for testing groundwater-modeling strategies for performance assessment of nuclear waste disposal. The SKI initiated the study in 1984, and the technical work was finalized in 1987 (Swedish Nuclear Power Inspectorate, 1987). The participating organizations were regulatory authorities as well as implementing organizations in ten countries. The study was devoted to testing of groundwater flow models and was performed at three levels: computer code verification, model validation, and sensitivity/uncertainty analysis.

Based upon lessons learned from INTRACOIN and HYDROCOIN, international consensus grew prior to and during the GEOVAL Symposium in Stockholm in April 1987 to begin a new project dealing with the validation of geosphere transport models. This new international cooperative project, named INTRAVAL, began in October 1987. As with the preceding projects, INTRAVAL was organized and managed by the SKI.

The INTRAVAL project was established to evaluate the validity of mathematical models for predicting the potential transport of radioactive substances in the geosphere (Swedish Nuclear Power Inspectorate, 1990). The unique aspect of INTRAVAL was the interaction between the experimentalists and modelers simulating the selected test cases for examining model validation issues. The test cases selected consisted of laboratory and field transport experiments and natural analogue studies that incorporate hydrogeologic and geochemical processes relevant to safety assessments of radioactive waste repositories.

25

Scale Issues

Keith Loague
Stanford University

Dennis L. Corwin
USDA-ARS

25.1	Introduction.....	25-1
	Scale, Definition, and Issues • A Brief Background • A Rich Literature	
25.2	Measurement and Interpretation	25-5
25.3	Upscaling and Downscaling.....	25-7
	Upscaling • Downscaling	
25.4	Examples	25-10
	Alternative Characterizations of Spatial Variability at the Catchment Scale • Extrapolation from Site to Region • Querying the Effect of Upscaling	
25.5	Epilogue	25-14
	Glossary.....	25-15
	Acknowledgments.....	25-16
	References	25-16

Anything you build on a large scale or with intense passion invites chaos.

Francis Ford Coppola

25.1 Introduction

25.1.1 Scale, Definition, and Issues

One of the many definitions for *scale* in Webster's dictionary is *a system of grouping or classifying in a series of steps or degrees according to a standard of relative size, amount, importance, perfection, etc.* The jargon used to describe scales in earth science includes, for example, continuum, correlation, integral, microscopic, macroscopic, local spatial average, REA (representative elementary area), and REV (representative elementary volume). The objective of this chapter is to introduce the subject of scale as it relates to hydrology, the entirety of which could easily fill several volumes.

Dooge's (1989) classifications for length and time scales in hydrology are given in Table 25.1. Relative to Dooge's scheme, it is worth pointing out that some scale classifications include a mega class. A relationship between scale and model type, relative to non-point source pollution models, is given in Figure 25.1. An important consideration for conceptualization of a given problem is for the model to account for the predominant processes occurring at the spatial and temporal scales of interest (Corwin and Loague, 2005).

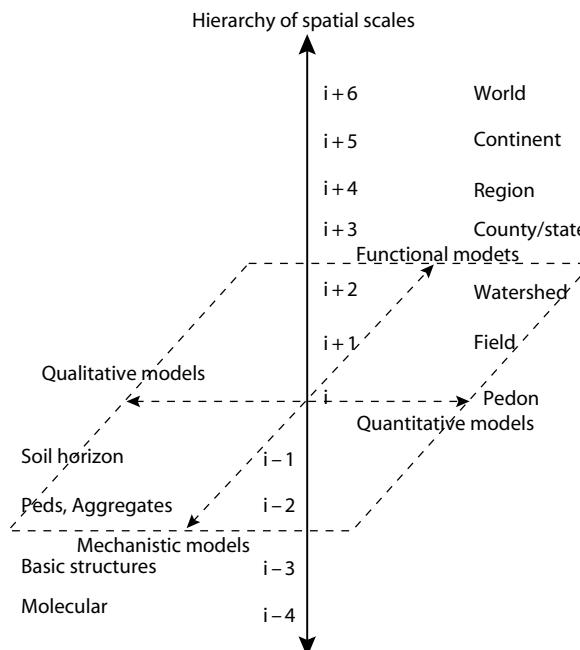
The issue of scale (both spatial and temporal) is ever present in modern hydrology. Whether one is engaged in laboratory or field experiments, parameter estimation or modeling, scale is an important

TABLE 25.1 Spatial and Temporal Scales in Hydrology

Spatial Scale			Temporal Scale		
Class	System	Length (m) ^a	Class	System	Time (s) ^b
Micro	Molecular cluster	1×10^{-8}	Micro	Water cluster	1×10^{-13}
Micro	Continuum point	1×10^{-5}	Micro	Continuum point	1×10^{-6}
Micro	REV ^c	1×10^{-2}	Micro	Experimental plot	1
Meso	Module	1×10^2	Meso	Earth rotation	1×10^4
Meso	Sub-catchment	1×10^3	Meso	Moon orbit	1×10^6
Meso	Small catchment	1×10^4	Meso	Earth orbit	1×10^7
Macro	Large catchment	1×10^5	Macro	Sun spots	1×10^8
Macro	Continental	1×10^6	Macro	Orbital eccentricity	1×10^{12}
Macro	Planetary	1×10^7	Macro	Stellar evolution	1×10^{16}

^a Typical.^b Order of magnitude.^c REV, representative elementary volume.

Source: After Dooge (1989).

**FIGURE 25.1** Organization hierarchy of spatial scales pertinent to non-point source pollution models. (Redrawn from Wagenet, R.J. and Hutson, J.L., *Journal of Environmental Quality*, 25, 499, 1996.)

consideration. At the surface the focus can be on the land phase, the channel phase, or the combined land and channel phase. In the subsurface the focus can range from an individual soil profile through an entire continent. For coupled atmospheric–surface and surface–subsurface problems the spatial and temporal scales can be significantly different. The water balance estimates in Table 25.2 illustrate the considerable range in volumes and residence times for the various components of the hydrologic cycle. One must also recognize that the underlying spatial and temporal scales can be different for water movement than they are for solute–sediment transport. For example, Figure 25.2 illustrates the tremendous range in spatio-temporal scales for hydrologically driven soil erosion processes.

TABLE 25.2 World Water Balance Estimates

System	Surface Area (km ² × 10 ⁶)	Volume (km ³ × 10 ⁶)	Volume (%)	Equivalent Depth (m) ^a	Residence Time (sec)
Oceans and seas	361	1,370	94	2,500	$\sim 1.3 \times 10^{11}$
Lakes and reservoirs	1.55	0.13	<0.01	0.25	$\sim 3.2 \times 10^8$
Swamps	<0.1	<0.01	<0.01	0.007	$3.2 \times 10^7 \Rightarrow 3.2 \times 10^8$
Rivers and channels	<0.1	<0.01	<0.01	0.003	$\sim 1.2 \times 10^6$
Soil moisture	130	0.07	<0.01	0.13	$1.2 \times 10^6 \Rightarrow 3.2 \times 10^7$
Groundwater	130	60	4	120	$1.2 \times 10^6 \Rightarrow 3.2 \times 10^{11}$
Icecaps and glaciers	17.8	30	2	60	$3.2 \times 10^8 \Rightarrow 3.2 \times 10^{10}$
Atmospheric water	504	0.01	<0.01	0.025	$\sim 8.6 \times 10^5$
Biopsheric water	<0.1	<0.01	<0.01	0.001	$\sim 6.1 \times 10^5$

^a Computed as though storage were uniformly distributed over the entire surface of the earth.

Source: Freeze, R.A. and Cherry, J.A., *Groundwater*, Prentice-Hall, Englewood Cliffs, NJ, 1979.

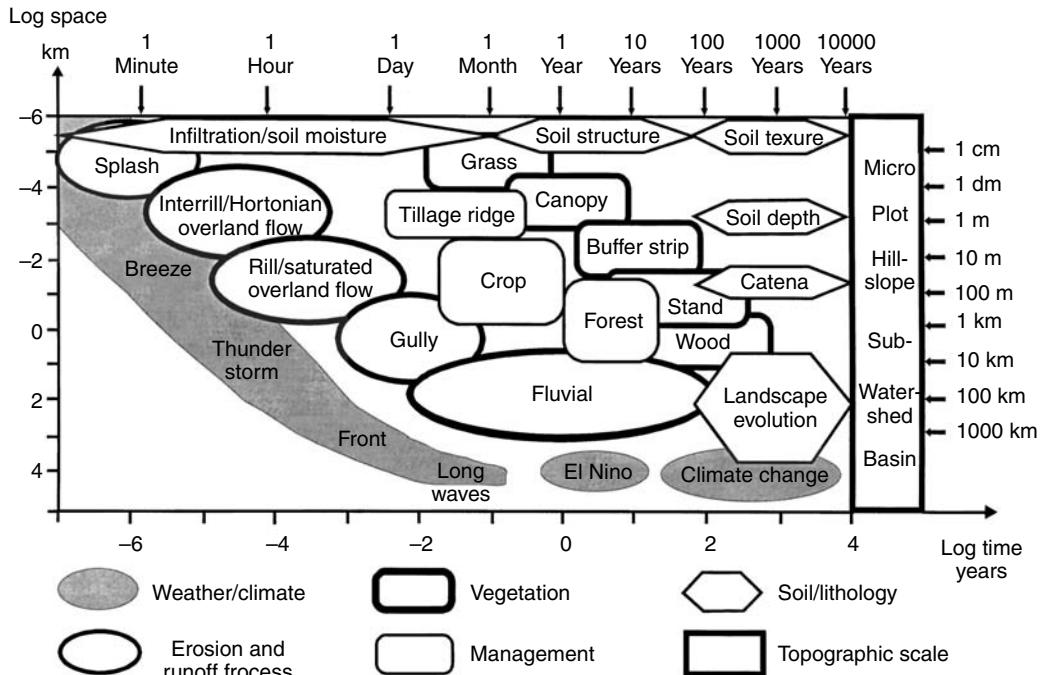


FIGURE 25.2 Temporal and spatial extent of hydrological processes and natural phenomena with special consideration for soil erosion. (Redrawn from Renschler, C.S., Strategies for implementing natural resource management tool — A geographical information science perspective on water and sediment balance assessment at different scales, PhD dissertation, University of Bonn, Germany, 2000.)

The spatial variability of surface, near surface, and subsurface hydrologic parameters is well known (e.g., de Marsily et al., 2005). Perception plays a major role in the characterization of spatial variability and, subsequently, scale. The two-part Figure 25.3 cleverly illustrates that what appears at first blush to be the visible variability at an assumed scale is in fact the variability at a much larger scale. It is also critically important to recognize that not all hydrologic parameters are time invariant. The development of a grid or mesh for process-based simulations of hydrologic response requires a level of discretization sufficient to represent spatial variations in hydrologic parameters. For transient hydrologic-response simulation

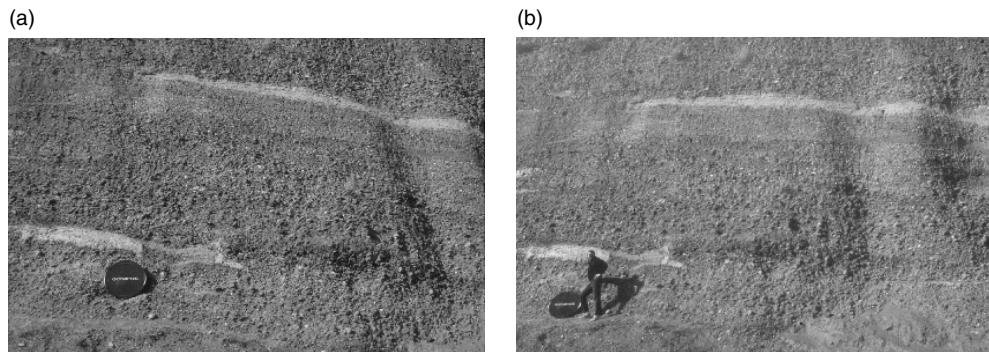


FIGURE 25.3 (a) Photograph of a section of heterogeneous porous media, with an Olympus camera lens cap for scale. (b) Photograph of the same section of heterogeneous porous media, showing that the diameter of the camera lens cap is much larger (i.e., 1 m) than was most likely assumed for the initial inspection of (a). Scott Tyler shot the two photographs [Stephen Wheatcraft provides the correct length scale in (b)]. Note, Tyler and Wheatcraft (1990) show alternative versions of these photos.

(e.g., event-based versus long-term) the proper selection of a minimum and maximum time step is critical.

Notwithstanding the type of model, all hydrologic response simulation efforts are faced with the daunting task of effectively representing the spatial and temporal variability for the parameters associated with the problem. Most often these parameters are measured for a scale that is much smaller (with a characteristic variability) than the level of discretization to be employed for the simulation, requiring an upscaling protocol. When the objective is to reconstruct small-scale variability from a larger scale average, a downscaling protocol is required.

25.1.2 A Brief Background

Much of the backbone of modern hydrology was developed before scale was the center stage issue it is today (e.g., Darcy, 1856; Buckingham, 1907; Terzaghi, 1925; Richards, 1931; Hubbert, 1940; Biot, 1941; Horton, 1945; Penman, 1948). For nearly half a century, high-speed digital computers have facilitated the development of physics-based hydrologic response models (e.g., Freeze, 1971; Harbaugh and McDonald, 1996; Zheng and Wang, 1998; VanderKwaak, 1999). The simulation scale in hydrology ranges from the soil profile (e.g., Freeze, 1969) to the continental scale (e.g., Garven, 1995). Segol (1994) and Beven (2000) review many of the classic groundwater and rainfall-runoff simulations. There is a large spectrum of modeling approaches that have been used in hydrology (e.g., continuum, stochastic continuum, pore-scale network, effective domain network, percolation theory, stochastic stream-tube, transfer function), each of which is impacted by the problems of scale.

Earlier, it was assumed that sophisticated deterministic-conceptual simulations, employed in parallel to exhaustive plot- or well-scale information, could easily be scaled up to simulate process-based hydrologic response at larger scales. This assumption, grounded in the fact that plots make up hillslopes, hillslopes make up catchments, and catchments make up watersheds, was shortsighted as, to date, no data set has allowed for the complete characterization, with high levels of confidence in both space and time, at the larger scales. In the context of scale-dependent deterministic-conceptual simulation, equifinality is a well-known problem (Beven, 1993; Savenije, 2001; Beven, 2005). Freeze (1975, 1980) was a pioneer in the use of stochastic-conceptual simulation for unraveling scale issues in both groundwater flow and rainfall-runoff processes, considering the variability and uncertainty in both inputs and outputs (Gelhar, 1974).

Five significant scale-related contributions to our improved understanding of hydrologic response are (i) the Miller and Miller (1956) theory for similar media (i.e., similitude analysis), (ii) the Pilgrim

(1983) identification of the problems associated with transferring hydrological relationships between regions, (iii) the Klemes (1983) argument that the level of scale at which meaningful conceptualizations of physical processes are possible are not arbitrary and their ranges are not continuous, (iv) the Cushman (1986) discussion of the relationships between measurements, scale, and scaling, and (v) the Woolhiser and Goodrich (1988) work showing the effects of storm rainfall intensity patterns on surface runoff. Obviously, the five contributions listed here are just the tip of the proverbial iceberg.

25.1.3 A Rich Literature

Four of the earliest workshops that focused entirely on the problems of scale in hydrology were the 1982 meeting in Caracas, Venezuela (Rodriguez-Iturbe and Gupta, 1983), the 1984 meeting at Princeton University (Gupta et al., 1986), the 1993 meeting in Robertson, Australia (Kalma and Sivapalan, 1995; Sivapalan and Kalma, 1995), and the 1996 meeting at Krumbach, Austria (Bloschl et al., 1997). Other meetings have also resulted in special issues focused on scale and scaling in hydrology (e.g., Sivapalan et al., 2003a, 2004). There are several excellent books dealing with scale-related issues in hydrology and related fields (e.g., Cushman, 1990; Hillel and Elrick, 1990; Feddes, 1995; Stewart et al., 1996; Rodriguez-Iturbe and Rinaldo, 1997; Sposito, 1998; Bierkens, et al., 2000; Grayson and Bloschl, 2000; Pachepsky et al., 2003). There are also several outstanding review and commentary papers on scale issues in hydrology and related fields (e.g., Church and Mark, 1980; Wood et al., 1990; de Boer, 1992; Gelhar et al., 1992; Koltermann and Gorelick, 1996; Shuttleworth et al., 1997; Mayer et al., 1999; Dodds and Rothman, 2000; Beven, 2001; Bloschl, 2001a,b; Cushman et al., 2002; Farmer, 2002; Grayson et al., 2002; Western et al., 2002; Lin, 2003; Neuman, 2003; Skoien et al., 2003; Sivapalan, 2003; Sivapalan et al., 2003b; Shaman et al., 2004; Bloschl, 2005; Noetinger et al., 2005).

25.2 Measurement and Interpretation

More than 20 years ago Gupta et al. (1984) reported an amazingly detailed large-scale saturated subsurface interpretation for groundwater flow simulations within a multiaquifer system. The 3D finite-element representation of Long Island, New York by Gupta et al. was based upon six hydrogeologic units that were gleaned from stratigraphic maps. To the best of our knowledge, the finite-element mesh in Gupta et al. is the only *foldout* ever to be published in *Water Resources Research*.

Characterization of the unseen subsurface, with the information that is typically available, is difficult at best. Even an accurate representation of surface topography (e.g., identifying size from digital elevation models) can be troublesome (Montgomery and Foufoula-Georgiou, 1993). Notwithstanding the information problem, the underlying foundation to the success of any hydrologic-response simulation is the best possible conceptualization of the problem. Bredehoeft (2005) defines *surprise* as new data that renders a prior conceptualization invalid. More data is generally better; however, the \$64,000 question of when enough information is really enough (James and Gorelick, 1994) is important for the problems of scale.

The size of the *area of influence* for a given hydrogeologic parameter measurement is often related to the scaling problem. For example, saturated hydraulic conductivity measurements can be made at the column, point, plot, and local scales. Figure 25.4 illustrated that data from different measurement scales can have different correlation lengths. One of the implications of invasive measurement techniques is that the system is changed, which can result in measurement error. Data collected from small areas are often employed in hydrologic response simulations to represent larger areas (Hopmans et al., 2002). Soil texture is often used as a surrogate for soil-hydraulic properties when no data are available (Rawls et al., 1991). Pedotransfer functions, remote sensing, and other non-invasive techniques are also used to estimate near-surface hydraulic and hydrogeologic parameters at larger scales (Corwin et al., 1999). Water balance and hydrologic-response simulations are often used (in a residual or inverse mode) to estimate hydrologic parameters at regional scales. The use of electrical resistivity and electromagnetic induction methods

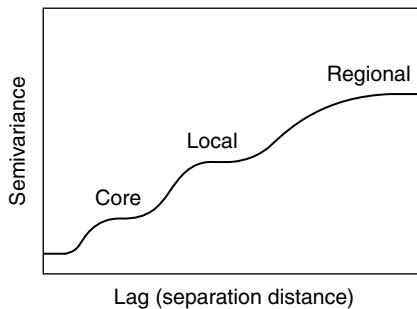


FIGURE 25.4 Hypothetical variogram for scale-dependent hydraulic conductivity. (Redrawn from Gelhar, L.W., *Water Resources Research*, 22, 1355, 1986.)

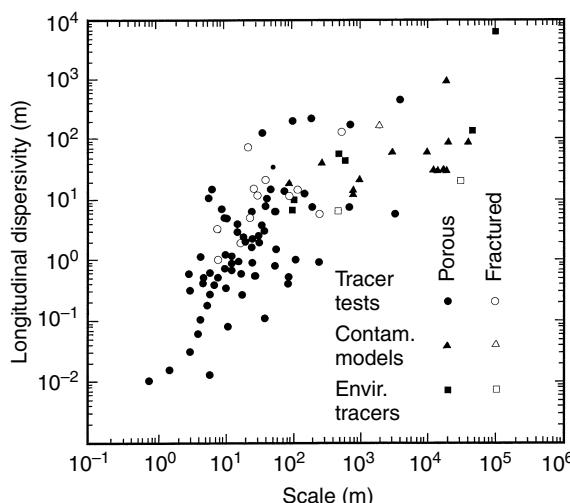


FIGURE 25.5 Longitudinal dispersivity as dependent on an overall displacement scale for different types of observations and media. (Redrawn from Gelhar, L.W., Welty, C., and Rehfeldt, K.R., *Water Resource Research*, 28, 1955, 1992.)

to interpret between widely spaced conventional hydrogeologic measurements is extremely promising (Corwin and Plant, 2005; Day Lewis et al., 2006).

Davis (1969) was one of the first to thoroughly summarize porosity and permeability values for earth materials. Nielsen et al. (1973) and Philip (1980) were among the first to discuss spatial variability for near-surface soil-hydraulic properties. A big dilemma in employing data from different measurement scales for hydrologic-response simulations is the quantification of reliability for the mixed information (e.g., hydraulic conductivity correlated to grain size analysis and hydraulic conductivity from pump test results). An important problem for simulating hydrologic response is the identification and characterization of small-scale preferential flow paths located within larger systems (e.g., Clothier, 2002; Zheng and Gorelick, 2003). An often ignored obstacle related to effectively exciting physics-based models is the poor assumption of time invariance for many near-surface hydraulic and hydrogeologic parameters. For example, Loague and VanderKwaak (2004) report permeability values for a single location where the measurement was repeated and the difference was greater than 700%.

Figure 25.5 and Figure 25.6, from the pioneering work of Gelhar et al. (1992), clearly illustrate scale impacts for longitudinal (along the flow path) dispersivity for both porous and fractured media. Some researchers take dispersivity as a characteristic property used to estimate the magnitude of hydrodynamic

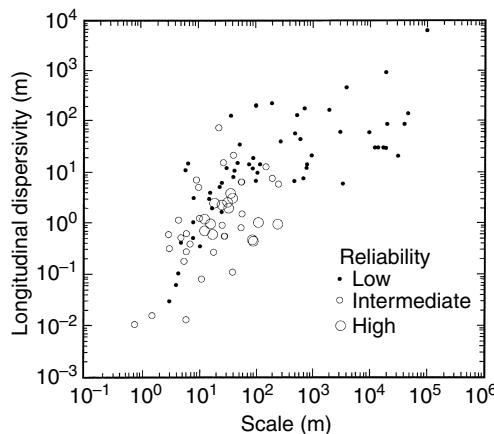


FIGURE 25.6 Field longitudinal dispersivity data classified according to reliability. (Redrawn from Gelhar, L.W., Welty, C., and Rehfeldt, K.R., *Water Resource Research*, 28, 1955, 1992.)

dispersion in solute transport simulations; others argue that it is a calibration parameter employed because there is insufficient information to describe the flow regime. The dispersivity values shown in Figure 25.5 were estimated from a mixed set of tracer experiments (smaller scales), environmental tracers (larger scales), and numerical simulations. The estimated dispersivities in Figure 25.5 range from 0.01 m to 10 km. The dispersivity estimates shown in Figure 25.5 are presented again in Figure 25.6, now classified based upon their reliability. Figure 25.6 shows that most of the reliable information falls within a scale of less than approximately 100 m and that most of the least reliable data is from scales over a kilometer.

Geostatistics and geographic information system (GIS) technology are well-established tools for the characterization/representation of spatial variability relative to assessing hydrologic response at different scales (e.g., Burrough and McDonnell, 1998; Deutsch and Journel, 1998; Webster and Oliver, 2001; Longley et al., 2005). The Netherland of geostatistical theory [e.g., use of training images (Caers, 2003)] and the array of GIS methods are well beyond the scope of this chapter. Fractal mathematics (e.g., Rodriguez-Iturbe and Rinaldo, 1997), fuzzy rules (e.g., Bardossy et al., 2002), neural networks (e.g., Tsegaye et al., 2003), and wavelet analysis (e.g., Chu et al., 1998), as they are related to scale problems in hydrology, are not covered in this chapter.

The term “scaling”, to many, is veiled in a nimbus of exciting mystery. At a basic level, part of the mystery simply comes from confusion of two connotations of the word — meaning either scale invariance (i.e., processes behaving similarly at small scales) or upscaling/downscaling (i.e., aggregating/disaggregating data).

Bloschl (2001b)

25.3 Upscaling and Downscaling

The material in this section is gleaned in large part from the lucid contributions of Bierkens et al. (2000) and Bloschl (2005). From a generic prospective there are two types of upscaling or downscaling methods: (i) how model equations/parameters change with scale, and (ii) how to best represent variability (space and time) at various scales. Bierkens et al. (2000) use the term scale for the concept of support, where support is defined as the largest area [L^2], volume [L^3], or time interval [T] for which the property of interest is considered homogeneous. The scale triplet (spacing, extent, support), defined by Bloschl and Sivapalan (1995), is illustrated in Figure 25.7 and Figure 25.8. The scale triplet applies to both

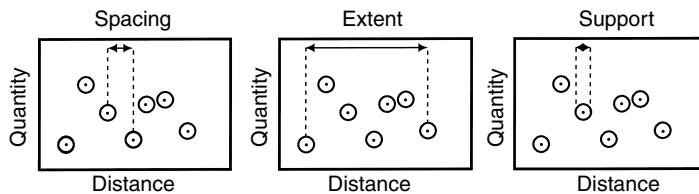


FIGURE 25.7 Schematic illustration of the scale triplet. (Redrawn from Bloschl, G. and Grayson, R., *Spatial Patterns in Catchment Hydrology: Observations and Modeling*, Cambridge University Press, Cambridge, United Kingdom, 2000, Chap. 2.)

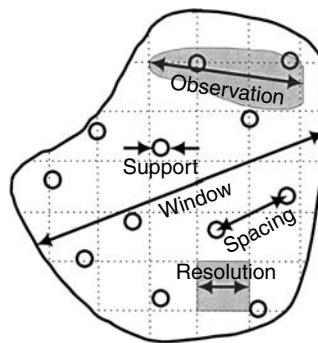


FIGURE 25.8 Scales affecting hydrogeologic variables. (Redrawn from Neuman, S.P. and Di Federico, V., *Review of Geophysics*, 41, 4–1, 2003.)

measurements and simulations (Bloschl and Grayson, 2000). Discussing the scale triplet Bloschl (2005) writes:

In dedicated studies, soil hydraulic conductivity measurements, for example, may have spacings of decimeters, while rain gauges in a region are typically spaced at tens of kilometers. The extent, which is the overall size of the domain sampled, may again range from meters to hundreds of kilometers in hydrological application. The support is the integration volume or area of the samples ranging from, say 1 dm^2 in the case of time domain reflectometry *in situ* probes, to hectares in the case of groundwater pumping tests (Anderson, 1997) or micrometeorological studies of the atmosphere (Schmid, 2002), and square kilometers in the case of remotely sensed data (Western et al., 2002). In hydrology, the catchment area can also be thought of as a support scale (Bierkens et al., 2000). In case of a model, the notion of a scale triplet is similar. For example, for a spatially distributed hydrologic model the scale triplet may have typical values of, say 25 m spacing (i.e., the grid spacing), 1 km extent (i.e., the size of the catchment or aquifer to be modeled), and 25 m support (the cell size). Analogue scales apply to the temporal domain.

In terms of extent, upscaling is usually referred to as extrapolation (with the opposite being singled out). In terms of spacing, downscaling is usually referred to as interpolation (with the opposite being singled out). In terms of support, both upscaling and downscaling are referred to as aggregation and disaggregation, respectively. Skoien and Bloschl (2006) performed Monte Carlo experiments to examine the nuances associated with the scale triplet.

The typical upscaling-downscaling exercise involves the following four steps (Bloschl, 2005): (i) analyzing the local data and scrutinizing the literature to decide on the model type, (ii) estimating the parameters from the data, (iii) verifying the upscaling-downscaling model against an independent data set, and (iv) performing the actual upscaling-downscaling step. Bloschl (2005) discusses upscaling-downscaling for

six important cases: (i) upscaling point rainfall to catchments, (ii) temporal disaggregation of rainfall, (iii) statistical downscaling of the output of global circulation models, (iv) flood frequency as a function of catchment scale, (v) upscaling and downscaling soil moisture, and (vi) subsurface media characterization and generation. Overviews of both upscaling and downscaling methods are provided in the following subsections.

25.3.1 Upscaling

Upscaling methods can be divided into four major classes (Bierkens et al., 2000): (i) averaging observations or model outputs, (ii) finding representative parameters or input variables, (iii) averaging model equations, and (iv) model simplification. The different classes are based upon five criteria (Bierkens et al., 2000): (i) if a model is involved, (ii) if the model is linear in its input variables and parameters, (iii) if the model can be employed at many locations or time steps, (iv) if the form of the model is the same at the two scales involved, and (v) if the larger scale model can be analytically derived from the smaller scale model. When no model is involved the upscaling problem is trivial, with observed information at a given scale (S_1) used to represent the larger scale (S_2). When the model is linear the upscaling problem is relatively simple [i.e., (i) conduct many simulations at S_1 then average the S_1 outputs to S_2 or (ii) average the S_1 model inputs to S_2 , then conduct a single simulation at S_2]. When the model can be applied at many locations or time steps, simulations are conducted for all the scenarios where information at S_1 is available; the S_1 outputs are averaged for S_2 (i.e., calculate first, average later). If the information coverage at S_1 is insufficient (i.e., the data crisis) or if the model is computational inefficient (i.e., too much time or storage) it is preferable to conduct simulations at S_2 directly. When the model has the same form for both scales, representative values for the input information are needed (i.e., the representative value for the support unit at S_1 produces output at S_2). When the model does not have the same form for both scales, a relationship between the S_1 and S_2 models is needed (e.g., the form of the S_2 model depending on the S_1 information, the S_1 model, and the support ratio S_2/S_1). When the larger scale model can be analytically derived from the smaller scale model it usually involves averaging the equations, with supplementary assumptions regarding the input information within the support units at S_2 . If linear, both models will have the same form. When the larger scale model cannot be analytically derived from the smaller scale model the upscaling must be accomplished in another manner [e.g., (i) input information for S_1 support units are determined, (ii) the S_1 model is employed for the support units with the outputs averaged for S_2 , (iii) a simplified model is postulated for S_2 , and (iv) outputs from the simplified model are compared to outputs from the S_1 model for the support units; if the comparisons are unacceptable the simplified model can either be calibrated or reformulated].

Approaches for averaging non-exhaustive (i.e., the trivial case) information include auxiliary information, deterministic functions, and geostatistical methods (Bierkens et al., 2000). Auxiliary information can be either quantitative or qualitative (note, both deterministic functions and geostatistical methods can sometimes be applied to auxiliary information). Deterministic functions usually involve interpolation where the weights depend solely on the sample configuration (e.g., Thiessen polygons). Geostatistical methods estimate variability as a realization from a stochastic function, where the weights depend upon both the sample configuration and a model of spatial-temporal structure estimated from the data (e.g., block Kriging). Approaches for finding non-exhaustive representative information include deterministic functions and stochastic methods (Bierkens et al., 2000). Deterministic functions provide full coverage with a method of interpolation. Stochastic methods, with statistics estimated from known information, describe unknown variations with conditional realizations from a stochastic function, ultimately yielding a single probability distribution. Methods for averaging model equations (i.e., temporal/volume and ensemble averaging) and model simplification (i.e., lumped conceptual and meta modeling) are discussed by Bierkens et al. (2000).

It is worth pointing out that Bierkens et al. (2000) provide excellent schematic diagrams for all the upscaling methods described herein. As well, there are many examples of upscaling in the hydrology

related literature (e.g., Desbarats, 1987; Rubin and Gomez-Hernandez, 1990; Indelman and Dagan, 1993a, 1993b; Sanchez-Villa et al., 1995; Christie, 1996; Dolman and Blyth, 1997).

25.3.2 Downscaling

Downscaling is a more homogeneous problem than the upscaling problem. Essentially each downscaling problem consists of reconstructing (often a non-unique solution) the variation of a given property at some smaller scale (S_1) from information at a larger scale (S_2), given only the value (arithmetic average) at the larger scale. Downscaling problems can be classified as deterministic, conditional stochastic, and unconditional stochastic (Bierkens et al., 2000). For the deterministic downscaling problem the average value of the property is known exactly at scale S_2 . The task with this method is to find a single deterministic function that describes the variation (temporal or spatial) of the property at S_1 with the condition that the S_2 average value of the function is the same as the known average. For the conditional stochastic downscaling problem the average value of the property is known exactly (as in deterministic downscaling) at scale S_2 . The task with this method is to find a set of equally probable functions (an ensemble) describing the variation of the property at S_1 with the condition that the S_2 average value for each function (a realization) are the same as the known average (i.e., conditional). It is often necessary to constrain the functions to prevent negative values (e.g., precipitation, soil-water content). For the unconditional stochastic downscaling problem the average value of the property is not known exactly at scale S_2 , rather a probability distribution function (PDF) for the property is known. The task with this method is to find a set of equally probable functions describing the variation of the property at S_1 with the condition that the PDF of the S_2 averages of these functions is the same as the known PDF of the average (i.e., each individual realization is not required to have the same average for S_2). The type of functions used for deterministic or stochastic downscaling can be classified as auxiliary information, mechanistic, or empirical (Bierkens et al., 2000).

Deterministic and stochastic functions can both be employed with auxiliary information [i.e., a property (e.g., topography) strongly related to the target property (e.g., soil-water content) at S_1] available to explain the variation of the property at S_1 within S_2 . For example, deterministic functions can be employed when the average value of the property is known exactly for S_2 and the variation of the property at S_1 within S_2 is described with a single deterministic function. When auxiliary information is not available both deterministic and stochastic functions can be employed when there is a mechanistic model (e.g., groundwater flow equation) describing the unknown variation of the property at S_1 within S_2 . For example, deterministic functions can be employed when the average value of the property is known exactly for S_2 and the variation of the property at S_1 within S_2 is described with a single deterministic function. Deterministic and stochastic empirical functions, describing the unknown variation of a property at S_1 within S_2 , can be used when neither auxiliary information nor mechanistic models are available. For example, deterministic empirical function (e.g., spline, linear, additive) can be employed when the average value of the property is known exactly for S_2 and the variation of the property at S_1 within S_2 is described with a single deterministic function.

It is worth pointing out that Bierkens et al. (2000) provide excellent schematic diagrams for all the downscaling methods described herein. As well, there are many examples of downscaling in the hydrology related literature (e.g., Rodriguez-Iturbe, 1986; Robinson and Sivapalan, 1997; Wilby et al., 1998; Bishop et al., 1999; Prudhomme et al., 2002; Wood et al., 2004).

25.4 Examples

25.4.1 Alternative Characterizations of Spatial Variability at the Catchment Scale

The 0.1-km² rangeland catchment (known as R-5) located near Chickasha, Oklahoma has been host to considerable work focused on characterizing near-surface hydrologic response via simulation. A critical

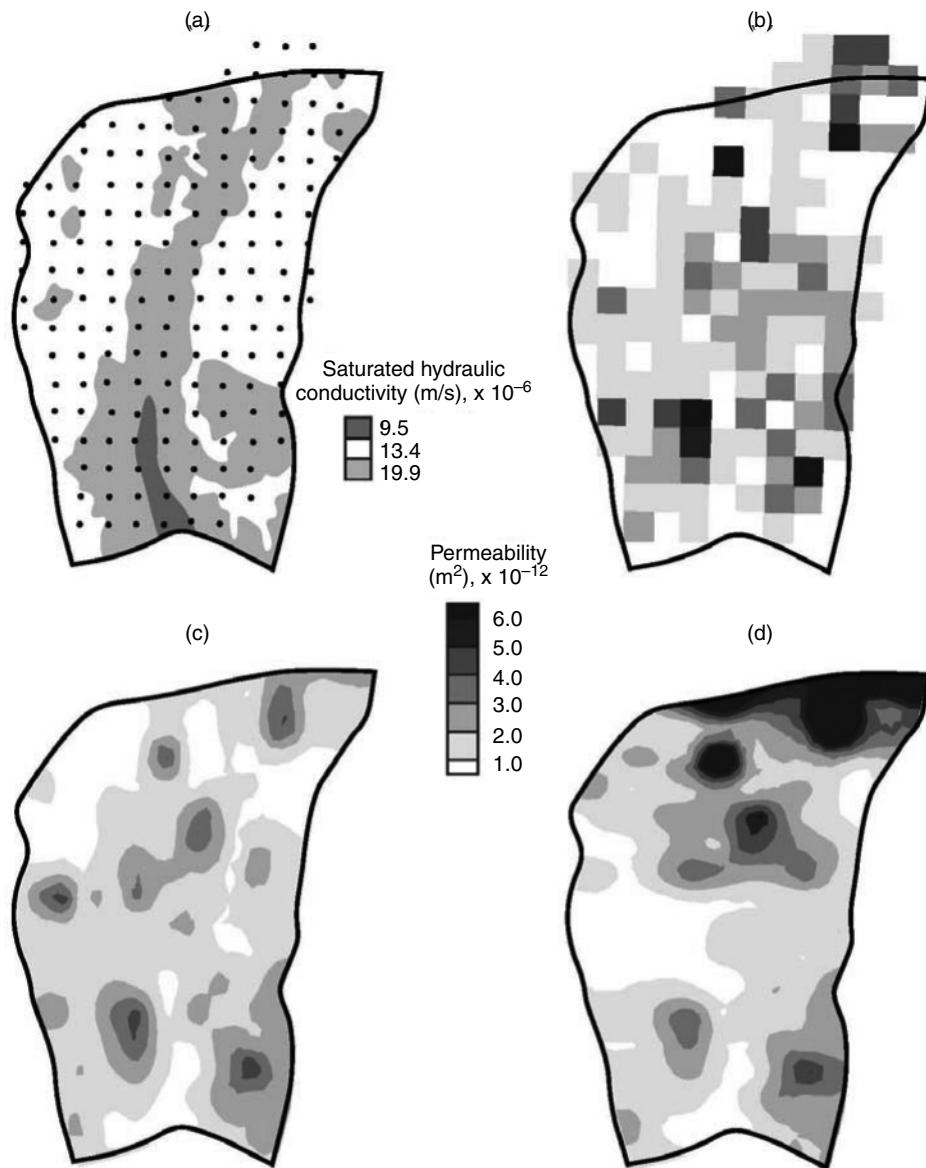


FIGURE 25.9 Alternative characterizations of spatially variable near-surface permeability based upon the same 157 steady-state infiltration experiments (Loague and Gander, 1990) across the 0.1-km^2 R-5 catchment. (a) Average saturated hydraulic conductivity values for three soil series (Loague, K., *Water Resources Research*, 26, 973, 1990), (b) Permeability (corrected for temperature) tapestry based upon the $\Delta x \approx \Delta y \approx 25\text{ m}$ measurement grid (Loague, K. and Kyriakidis, P.C., *Water Resources Research*, 33, 2883, 1997), (c) Permeability (corrected for temperature) estimated with geostatistical methods (VanderKwaak, J.E. and Loague, K., *Water Resources Research*, 37, 999, 2001), and (d) Permeability (corrected for temperature and scaled by the measurement chronology) estimated with geostatistical methods (Heppner, C.S. and Loague, K., *Earth Surface Processes and Landforms*, 2006).

piece of information in these simulations has been the spatial variability of surface permeability. Loague and Gander (1990) reported the results from 157 steady-state infiltration experiments (over a 51-day period in 1984) across a relatively uniform R-5 grid (i.e., $\Delta x \approx \Delta y \approx 25\text{ m}$; see Figure 25.9a). As illustrated in Figure 25.9 these data have been interpreted in several different ways over the past 20 years.

The data were used to (i) estimate average saturated hydraulic conductivity values for the three R-5 soil series (see Figure 25.9a), (ii) prepare saturated hydraulic conductivity and permeability (corrected for temperature) tapestries based upon the measurement grid (see Figure 25.9b), and (iii) estimate the spatial structure of permeability (corrected for temperature) with geostatistical methods (see Figure 25.9c). Most recently, Heppner and Loague (2006) used the data to estimate the spatial structure of permeability (corrected for temperature and scaled by the measurement chronology) with geostatistical methods (see Figure 25.9d). The four scenarios in Figure 25.9, albeit all defendable and well within the same overall range of values, are strikingly different in their spatial representation. The scenarios shown in Figure 25.9a to 25.9c, were each used for rainfall-runoff simulations with a relatively simple Horton type model (see summary in Loague et al., 2000). More recently, the scenarios shown in Figure 25.9c and Figure 25.9d were each used for comprehensive hydrologic-response simulations with a physics-based model (VanderKwaak and Loague, 2001; Loague and VanderKwaak, 2002; Loague et al., 2005; Heppner et al., 2006). When used to simulate hydrologic response, with either the Horton or comprehensive models, each of the four scenarios produced different results. It is worth pointing out that an order of magnitude estimate for permeability, which is often taken as sufficient for groundwater simulations, is not good enough for the catchment-scale near-surface simulations, whose success is based upon capturing the nuances of infiltration and exfiltration, especially when the rainfall rates and the surface permeabilities are similar in magnitude. There is no easy method for upscaling soil-hydraulic property data in the near surface.

25.4.2 Extrapolation from Site to Region

One method of upscaling is extrapolation by means of soil taxonomy. Soil taxonomy (Soil Survey Staff 1999) is a hierarchical classification system that establishes several successive categories of soils. Taxonomic levels in the U.S. system include: order, suborder, great group, subgroup, family, and series. The variability of a given soil property within a taxa increases at higher taxonomic categories. The classification was designed to bring out natural relationships among soils and enhance prediction about the behavior of a given kind of soil in relation to other kinds of soil for which information is available. Classification criteria include a wide range of morphologic, chemical, and physical characteristics. Average values of a given soil property are determined for a specific category and the average values are used wherever the taxonomic category is found in the region of interest. Upscaling by soil taxonomy is based on an approximation that soils of similar taxonomy have similar values for the soil property. Figure 25.10 shows an illustrative example of upscaling by soil taxonomy as discussed by Loague et al. (1989). One consideration in the use of such approximations is that, as more soils are grouped into larger categories, the variation within the group becomes larger. It is desirable to form as few groups as possible while not increasing the within-group variation too rapidly. The relation between the number of groups and the within-group variance can be represented by plotting the within-group variance versus the number of categories one obtains in the hierarchy.

25.4.3 Querying the Effect of Upscaling

Between the late 1950s and the time of its statewide cancellation in August of 1977, there was widespread use of DBCP (1, 2-dibromo-3-chloropropane) throughout California's San Joaquin Valley. Almost three decades after its cancellation, DBCP-contaminated groundwater persists as a problem in the San Joaquin Valley. The objective of the Fresno Case Study (FCS) was to address, from a simulation perspective, if label-recommended NPS applications were the principal source of DBCP-contaminated groundwater for an area approximately coincident with east-central Fresno County (Loague et al., 1998a,b). The FCS consisted of two 35-yr (1960–1994) simulation phases. In Phase I, the potential fate and transport of DBCP between the surface and the water table for multiple applications was quantitatively estimated with 1172 separate 1D simulations (Loague et al., 1998a). In Phase II, the aggregate of the DBCP concentrations loaded to the water table (from Phase I) were input to 3D transient saturated subsurface transport simulations (Loague et al., 1998b). Loague et al. (1998a) identify upscaling as a potential source of uncertainty in the FCS.

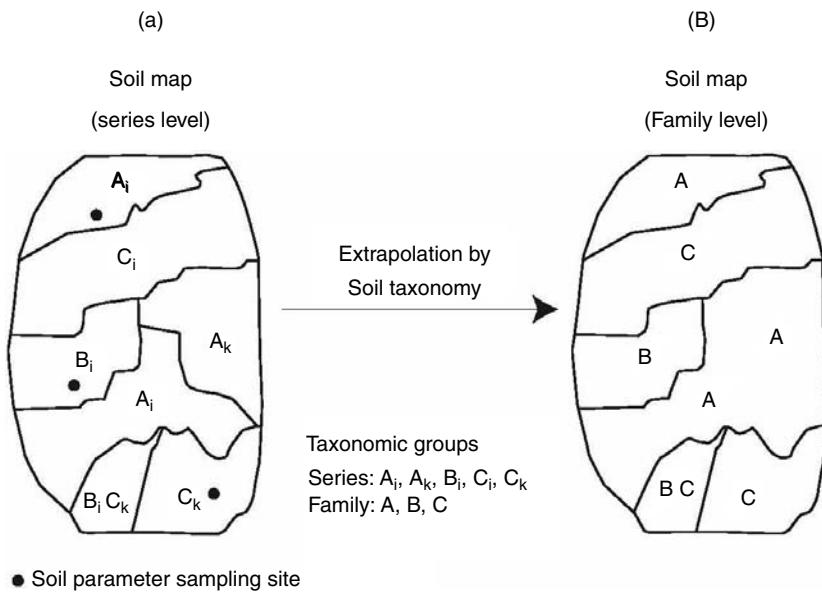


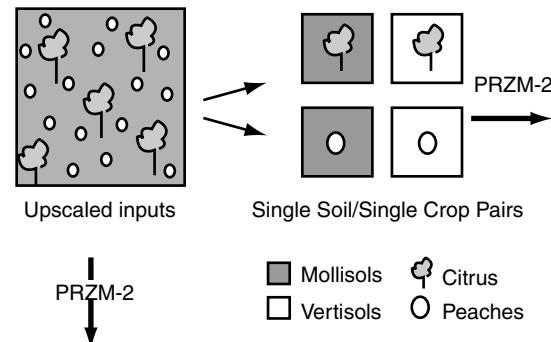
FIGURE 25.10 Schematic illustration of how soil parameter information is extrapolated from a lower taxonomic category to a higher taxonomic category (Loague, K.M., Yost, R.S., and Liang, T.C., *Journal of Contaminant Hydrology*, 4, 139, 1989). In (a) there are five soil series within three soil families. Each soil family has one site at which soil properties (e.g., saturated hydraulic conductivity) are measured. In (b) the measured soil properties are assumed to represent the soil series from which they were sampled and are then extrapolated to other soil series from the same soil family where no measurements were made.

Each of the near-surface FCS simulations represented a 1 km² area and required estimates of four soil properties (changing with depth): (i) soil-water content at the wilting point, (ii) soil-water content at field capacity, (iii) fraction of soil organic carbon, and (iv) bulk density. The FCS soil database was developed by averaging individual properties from the soil orders covering at least one-third of any 1-km² element. Records of changing crop cover were used (as a surrogate) to approximate multiple and variable DBCP applications in the FCS area. DBCP applications were estimated from the label-recommended rates for a specific crop weighted by the fractional area of an element containing that particular crop at a given time. With the upscaled soil and crop information, DBCP-concentration profiles were estimated via PRZM-2 (Mullins et al., 1993) simulations. The year-end mosaics of water-table concentrations indicated that DBCP had leached to groundwater (Loague et al., 1998a).

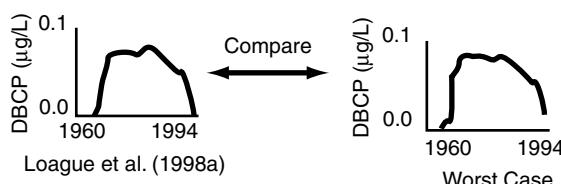
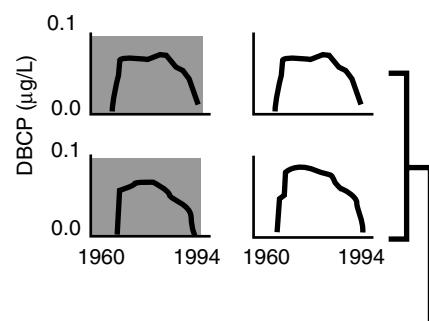
The FCS upscaling approach was limited by averaging both multiple soil orders and different application rates. Soutter and Loague (2000) critically examined the FCS water-table loading concentrations by comparison with new scenarios developed from simulations that did not employ the same upscaling. Figure 25.11 is a schematic illustration of the steps used by Soutter and Loague (2000) for simulating the worst-case DBCP water-table loading concentration scenarios without upscaling. The Soutter and Loague (2000) results indicated that (i) the areas most sensitive to upscaling were small relative to the total study area, (ii) the new concentration estimates were within the range of concentrations reported by Loague et al. (1998a), and (iii) the impacts from upscaling soil information exceeded the impacts from upscaling crop information. In as much as the Soutter and Loague (2000) study demonstrated that a somewhat coarse grid may be suitable for certain locations, thoughtful planning for variations in the grid size should help to target where upscaling is a potential problem for future regional-scale NPS simulation efforts.

There are two recently reported efforts related to the FCS. First, Stewart and Loague (2003, 2004) reported on the development (and San Joaquin Valley testing) of a type transfer function approach for regional-scale NPS groundwater vulnerability assessments. Second, Loague and Soutter (2006) reported

Step 1. Deconstruct the upscaled information from Loague et al. (1998a) into the unique single soil/single crop pairs.



Step 1. Simulate DBCP water-table loading concentrations with PRZM-2 for each single soil/single crop pair from Step 1.



Step 3. Combine the Step 2 results to construct new DBCP water-table loading concentration scenarios.

FIGURE 25.11 Schematic illustration (for a single 1 km^2 element) of the three steps used for simulating the worst-case DBCP water-table loading concentration scenarios without upscaling for comparison with the DBCP water-table loading concentration scenario with upscaling from the original Fresno case study. (Souter, L.A. and Loague, K., *Journal of Environmental Quality*, 29, 1794, 2000.)

on the cause for small-scale high-concentration DBCP hotspots within regional-scale (San Joaquin Valley like) groundwater plumes.

25.5 Epilogue

First and foremost, scale does matter in the earth sciences. For example, agricultural engineers, biometeorologists, climatologists, ecologists, geomorphologists, hydrologists, hydrogeologists, petroleum engineers, and soil physicists who attempt to understand process-based hydrologic response at the surface, in near surface, and within the saturated subsurface (for applications such as climate change, flooding, landscape evolution, surface and groundwater contamination) are all faced with scale-related problems. Relative to process-based simulation of hydrologic response, computers, and to a lesser extent theory, are not the delimiters they were just a few decades ago. However, insufficient data remains a major obstacle in hydrologic modeling. For example, it is unlikely that anytime soon there will be enough information available to rigorously quantify the impacts of small-scale preferential flow in the unsaturated and saturated subsurface flow at regional scales. Bloschl (2001b) points out that

Maybe, instead of trying to capture everything when upscaling we should be developing methods to identify dominant processes that control hydrological response in different environments and at different scales, and then develop models to focus on these dominant processes, a notion we might call the 'Dominant Processes Concept' (DPC).

Bredehoeft (2005) describes the conceptualization problem as the thorniest issue in modeling today and suggests that the best tack (today) is to consider alternative scenarios, rejecting the ones whose

estimates are unacceptable when compared to observed information. Bredehoeft (2005), faced with the problem of conceptualization (which is obviously scale related), answers *of course* to the question of whether hydrogeologists should continue to model. For example, deterministic-conceptual simulation facilitates (i) concept development, (ii) hypothesis testing, (iii) evaluation of existing (and future) data sets, (iv) design of new field experiments, (v) creation of synthetic data sets, (vi) hosting of Monte-Carlo experiments, (vii) driving physics-based solute transport and landscape evolution models, and (viii) the development and testing of simpler models (Loague and VanderKwaak, 2004).

The measure and model approach, pioneered by Robert E. Horton (Beven, 2004a, 2004b), is without question the best protocol in hydrology. Heeding the call for cooperative work between the Cains and Ables of hydrology (Dunne, 1983) is critical in the future if progress is to be made in wrestling the scale behemoth closer to the ground. The various calls for detailed long-duration field monitoring programs at selected experimental locations (e.g., Dunne, 1998; Entekhabi et al., 1999) are renewed here.

Glossary

Conceptual model A model is conceptual if their functional form is derived from consideration of physical processes.

Deterministic model A model is deterministic if all the variables are (viewed as) free from random variations.

Dispersion The spread of solute due to heterogeneities in the pore sizes and shapes (mechanical dispersion) and heterogeneities in the aquifer (macrodispersion).

Empirical model A model is empirical if their functional form is not derived from consideration of physical processes.

Equifinality Refers to more than one parameter set (for deterministic-conceptual simulation) providing an equally good representation of overall hydrologic response.

Geographic information system An organized collection of computer hardware, software, and geographic data, to efficiently capture, store, update, manipulate, analyze, and display geographically reference information.

Geostatistics A method of estimation for spatial variables, based upon the theory of random functions.

Heterogeneity The occurrence of variations in the value of a field property at different points in space.

Kriging Denotes a geostatistical method of interpolation or averaging.

Model A representation of a real system or process. A mathematical model is a set of equations (with variables, constants, and coefficients) that represents relevant processes.

Non-point source pollution Contaminants entering the environment over an extensive area, which are difficult (at best) or impossible to trace to a specific location and have the potential for maintaining a relatively long active presence on entire ecosystems.

Upscaling Representing the heterogeneous domain for modeling purposes through elements larger than the scales of heterogeneity but not large enough to warrant the use of effective properties. The effects of subelement heterogeneity are represented using coefficients that depend on the element size.

Pedotransfer function Transfer function that use particle-size distribution, bulk density, or soil organic-carbon content to yield soil-water retention or unsaturated hydraulic conductivity functions.

Preferential flow Local subsurface pathways, such as structural voids or biological channels, which can carry water at velocities much greater than those of the surrounding matrix.

Representative elementary area The scale at which a statistical treatment of spatial variability can replace a deterministic description, or where small-scale complexity gives way to larger scale simplicity.

Representative elementary volume The exact size is not determined, but it is assumed that the size is much larger than the pore scale and much smaller than the scale of the porous medium.

Scale of observation Window in space or time through which the environment or some property of the environment is viewed or measured.

Stochastic model A model is stochastic if any of the variables are described by a probability distribution.

Water balance An accounting for the inflows and outflows for a fluid control volume.

Acknowledgments

We are very grateful to Professor Gunter Bloschl for generously sharing his *in press* papers/book on scale issues. We also thank Scott Tyler and Stephen Wheatcraft for the use of the photos in Figure 25.3 and Chris Heppner for his help in preparing Figure 25.9.

References

- Anderson, M.P., Characterization of geological heterogeneity, in *Subsurface Flow and Transport: A Stochastic Approach*, Dagan, G. and Neuman, S.P., (eds.), Cambridge University Press, Cambridge, United Kingdom, 1997, chap. 1.
- Bardossy, A., Stehlík, J., and Caspary, H.J., Automated objective classification of daily circulation patterns for precipitation and temperature downscaling based on optimized fuzzy rules, *Climate Research*, 23, 11, 2002.
- Beven, K., Prophecy, reality and uncertainty in distributed hydrological modeling, *Advances in Water Resources*, 16, 41, 1993.
- Beven, K.J., *Rainfall-Runoff Modelling: The Primer*, John Wiley and Sons, Chichester, England, 2000.
- Beven, K., On landscape space to model space mapping, *Hydrological Processes*, 15, 323, 2001.
- Beven, K., Infiltration excess at the Horton hydrology laboratory (or not?), *Journal of Hydrology*, 293, 219, 2004a.
- Beven, K. Robert E., Horton's perceptual model of infiltration processes, *Hydrological Processes*, 18, 3447, 2004b.
- Beven, K., A manifesto for the equifinality thesis, *Journal of Hydrology*, 320, 18, 2006.
- Bierkens, F.P., Finke, P.A., and de Willigen, P., *Upscaling and Downscaling Methods for Environmental Research*, Kluwer Academic Publishers, Dordrecht, the Netherlands, 2000.
- Biot, M.A., General theory of three-dimensional consolidation, *Journal of Physics*, 12, 155, 1941.
- Bishop, T.F.A., McBratney, A.B., and Laslett, G.M., Modelling soil attribute depth functions with equal-area quadratic smoothing splines, *Geoderma*, 91, 27, 1999.
- Bloschl, G., Scaling issues in snow hydrology, *Hydrological Processes*, 13, 2149, 2001a.
- Bloschl, G., Scaling in hydrology, *Hydrological Processes*, 15, 709, 2001b.
- Bloschl, G., Statistical upscaling and downscaling in hydrology, in *Encyclopedia of Hydrological Sciences* (Part 1. Theory, Organization and Scale), Anderson, M., (ed.), John Wiley and Sons, Chichester, England, pp. 135–154, 2005.
- Bloschl, G. and Grayson, R., Spatial observations and interpolation, in *Spatial Patterns in Catchment Hydrology: Observations and Modeling*, Grayson, R. and Bloschl, G., (eds.), Cambridge University Press, Cambridge, United Kingdom, 2000, chap. 2.
- Bloschl, G. and Sivapalan, M., Scale issues in hydrological modelling — a review, *Hydrological Processes*, 9, 251, 1995.
- Bloschl, G., Sivapalan, M., Gupta, V., and Beven, K., Preface to the special section on scale problems in hydrology, *Water Resources Research*, 33, 2881, 1997.
- Buckingham, E., *Studies on the movement of soil moisture*, Bulletin 38, U.S. Department of Agriculture Bureau of Soils, Washington D.C., 1907.

- Burrough, P.A. and McDonnell, R.A., *Principles of Geographical Information Systems*, Oxford University Press, New York, 1998.
- Bredehoeft, J., The conceptualization model problem — surprise, *Hydrogeology Journal*, 13, 37, 2005.
- Caers, J., History matching under training-image-based geological model constraints, *Society of Petroleum Engineers Journal*, 8, 218, 2003.
- Christie, M.A., Upscaling for reservoir simulation, *Journal of Petroleum Technology*, 48, 1004, 1996.
- Chu, L.F., Schatzinger, R.A., and Tham, M.K., Application of wavelet analysis to upscaling of rock properties, *SPE Reservoir Evaluation and Engineering*, 1, 75, 1998.
- Church, M. and Mark, D.M., On size and scale in geomorphology, *Progress in Physical Geography*, 4, 342, 1980.
- Clothier, B.E., Rapid and far-reaching transport through structured soils, *Hydrological Processes*, 16, 1321, 2002.
- Corwin, D.L. and Loague, K., Multidisciplinary approach for assessing subsurface non-point source pollution, in *Soil-Water-Solute Process Characterization: An Integrated Approach*, Alvarez-Benedi, J. and R. Munoz-Carpeta, R., (eds.), CRC Press, Boca Raton, FL, 2005, chap. 1.
- Corwin, D.L. and Plant, R.E., Applications of apparent soil electrical conductivity in precision agriculture, *Computers and Electronics in Agriculture*, 46, 1, 2005.
- Corwin, D.L., Loague, K., and Ellsworth, T.R., (eds.), *Assessment of Non-Point Source Pollution in the Vadose Zone*, Geophysical Monograph 108, American Geophysical Union Press, Washington, D.C., 1999.
- Cushman, J.H., On measurement, scale, and scaling, *Water Resources Research*, 22, 129, 1986.
- Cushman, J.H., (ed.), *Dynamics of Fluids in Hierarchical Porous Media*, Academic Press, San Diego, CA, 1990.
- Cushman, J.H., Bennethum, L.S., and Ju, B.H., A primer on upscaling tools for porous media, *Advances in Water Resources*, 25, 1043, 2002.
- Darcy, H., *Les fontaines publiques de la ville de Dijon*, Victor Dalmont, Paris, France, 1856.
- Davis, S.N., Porosity and permeability of natural materials, in *Flow Through Porous Media*, DeWiest, R.J.M., (ed.) Academic Press, New York, 1969, chap. 2.
- Day-Lewis, F.D., Lane, J.W., and Gorelick, S.M., Combined interpretation of radar, hydraulic, and tracer data from a fractured-rock aquifer near Mirror Lake, New Hampshire, USA, *Hydrogeology Journal*, 14, 1, 2006.
- de Boer, D.H., Hierarchies and spatial scale in process geomorphology: A review, *Geomorphology*, 4, 303, 1992.
- de Marsily, Gh., Delay, F., Goncalves, Ph., Teles, V., and Violette, S., Dealing with spatial heterogeneity, *Hydrogeology Journal*, 161, 2005.
- Desbarats, A.J., Numerical estimation of effective permeability in sand-shale sequences, *Water Resources Research*, 23, 273, 1987.
- Deutsch, C.V. and Journel, A.G., *GSLIB: Geostatistical Software Library and User's Guide*, 2nd ed., Oxford University Press, New York, 1998.
- Dodds, P.S. and Rothman, D.H., Scaling, universality, and geomorphology, *Annual Reviews of Earth and Planetary Sciences*, 28, 571, 2000.
- Dolman, A.J. and Blyth, E.A., Patch scale aggregation of heterogeneous land surface cover for mesoscale meteorological model, *Journal of Hydrology*, 190, 253, 1997.
- Dooge, J.C.I., Scale problems in hydrology, *Fifth Memorial Chester C. Kisiel Lecture*, University of Arizona, Tucson, 1989.
- Dunne, T., Relation of field studies and modeling in the prediction of storm runoff, *Journal of Hydrology*, 65, 25, 1983.
- Dunne, T., Wolman lecture: Hydrologic science . . . in landscapes . . . on a planet . . . in the future, in *Hydrologic Sciences: Taking Stock and Looking Ahead*, National Academy Press, Washington, D.C., 1998, chap. 1.

- Entekhabi, D., Asrar, G.R., Betts, A.K., Beven, K.J., Bras, R.L., Duffy, C.J., Dunne, T., Koster, R.D., Lettenmaier, D.P., McLaughlin, D.B., Shuttleworth, W.J., van Genuchten, M.T., Wei, M.-Y., and Wood, E.F., An agenda for land surface hydrology research and a call for the second international hydrological decade, *Bulletin of the American Meteorological Society*, 80, 2043, 1999.
- Farmer, C.L., Upscaling: A review, *International Journal for Numerical Methods in Fluids*, 40, 63, 2002.
- Feddes, R.A., (ed.), *Space and Time Scale Variability and Interdependencies in Hydrological Processes*, Cambridge University Press, Cambridge, United Kingdom, 1995.
- Freeze, R.A., The mechanism of natural groundwater recharge and discharge: 1. One-dimensional, vertical unsteady, unsaturated flow above a recharging or discharging groundwater flow system, *Water Resources Research*, 5, 153, 1969.
- Freeze, R.A., Three-dimensional, transient, saturated-unsaturated flow in a groundwater basin, *Water Resources Research*, 7, 347, 1971.
- Freeze, R.A., A stochastic-conceptual analysis of one-dimensional groundwater flow in non-uniform homogeneous media, *Water Resources Research*, 11, 725, 1975.
- Freeze, R.A., A stochastic-conceptual analysis of rainfall-runoff processes on a hillslope, *Water Resources Research*, 16, 391, 1980.
- Freeze, R.A. and Cherry, J.A., *Groundwater*, Prentice-Hall, Englewood Cliffs, NJ, 1979.
- Garven, G., Continental scale groundwater flow and geologic processes, *Annual Reviews of Earth and Planetary Sciences*, 23, 89, 1995.
- Gelhar, L.W., Stochastic analysis of phreatic aquifers, *Water Resources Research*, 10, 539, 1974.
- Gelhar, L.W., Stochastic subsurface hydrology from theory to applications, *Water Resources Research*, 22, 135S, 1986.
- Gelhar, L.W., Welty, C., and Rehfeldt, K.R., A critical review of data on field-scale dispersion in aquifers, *Water Resources Research*, 28, 1955, 1992.
- Grayson, R. and Bloschl, G., (eds.), *Spatial Patterns in Catchment Hydrology: Observations and Modelling*, Cambridge University Press, Cambridge, United Kingdom, 2000.
- Grayson, R.B., Bloschl, G., Western, A.W., and McMahon, T.A., Advances in the use of observed spatial patterns of catchment hydrologic response, *Advances in Water Resources*, 25, 1313, 2002.
- Gupta, S.K., Cole, C.R., and Pinder, G.F., A finite-element three-dimensional groundwater (FE3DGW) model for a multiaquifer system, *Water Resources Research*, 20, 553, 1984.
- Gupta, V.K., Rodriguez-Iturbe, I., and Wood, E.F., (eds.), *Scale Problems in Hydrology: Runoff Generation and Basin Response*, D. Reidel Publishing Company, Dordrecht, Holland, 1986.
- Harbaugh, A.W. and McDonald, M.G., Programmer's documentation for MODFLOW-96, an update to the U.S. Geological Survey modular finite-difference ground-water flow model, U.S. Geological Survey Open-File Report 96-486, U.S. Geological Survey, Federal Center, Denver, CO, 1996.
- Hillel, D. and Elrick, D.E., (eds.), *Scaling in Soil Physics: Principles and Applications*, Soil Science Society of America, Madison, WI, 1990.
- Heppner, C.S. and Loague, K., Hydrologic response and sediment transport for the R-5 catchment: 1. Analysis of long-term observed data, *Earth Surface Processes and Landforms*, 2006 (in review).
- Heppner, C.S., Loague, K., and VanderKwaak, J.E., Hydrologic response and sediment transport for the R-5 catchment: 2. Long-term simulations with InHM, *Earth Surface Processes and Landforms*, 2006 (in review).
- Hopmans, J.W., Nielsen, D.R., and Bristow, K.L., How useful are small-scale soil hydraulic property measurements for large-scale vadose zone modeling? in *Environmental Mechanics: Water, Mass, Energy Transfer in the Biosphere*, Raats, P.A., Smiles, D., and Warrick, A.W., (eds.), Geophysical Monograph 129, p. 247, American Geophysical Union Press, Washington, D.C., 2002.
- Horton, R.E., Erosional development of streams and their drainage basins: Hydrophysical approach to quantitative morphology, *Bulletin of the Geological Society of America*, 56, 275, 1945.
- Hubbert, M.K., The theory of groundwater motion, *Journal of Geology*, 48, 785, 1940.
- Indelman, I. and Dagan, G., Upscaling of permeability of anisotropic heterogeneous formations, 1: The general framework, *Water Resources Research*, 29, 917, 1993a.

- Indelman, I. and Dagan, G., Upscaling of permeability of anisotropic heterogeneous formations, 2: General structure and small perturbation analysis, *Water Resources Research*, 29, 925, 1993b.
- James, B. and Gorelick, S.M., When enough is enough: The worth of monitoring data in aquifer remediation design, *Water Resources Research*, 30, 3499, 1994.
- Kalma, J.D. and Sivapalan, M., (eds.) *Scale Issues in Hydrological Modelling*, John Wiley and Sons, Chichester, England, 1995.
- Klemes, V., Conceptualization and scale in hydrology, *Journal of Hydrology*, 65, 1, 1983.
- Koltermann, C.E. and Gorelick, S.M., Heterogeneity in sedimentary deposits: A review of structure-imitating process-imitating and descriptive approaches, *Water Resources Research*, 32, 2617, 1996.
- Lin, H., Hydropedology: Bridging disciplines, scales, and data, *Vadose Zone Journal*, 2, 1, 2003.
- Loague, K., R-5 revisited: 2. Reevaluation of a quasi-physically based rainfall-runoff model with supplemental information, *Water Resources Research*, 26, 973, 1990.
- Loague, K. and Gander, G.A., R-5 revisited: 1. Spatial variability of infiltration on a small rangeland catchment, *Water Resources Research*, 26, 957, 1990.
- Loague, K. and Kyriakidis, P.C., Spatial and temporal variability in the R-5 infiltration data set: Deja vu and rainfall-runoff simulations, *Water Resources Research*, 33, 2883, 1997.
- Loague, K. and Souter, L.A., Desperately seeking a cause for hotspots in regional-scale groundwater plumes resulting from non-point source pesticide applications, *Vadose Zone Journal*, 5, 204, 2006.
- Loague, K. and VanderKwaak, J.E., Simulating hydrologic response for the R-5 catchment: Comparison of two models and the impact of the roads, *Hydrological Processes* 16, 1015, 2002.
- Loague, K. and VanderKwaak, J.E., Physics-based hydrologic response simulation: Platinum bridge, 1958 Edsel, or useful tool, *Hydrological Processes* 18, 2949, 2004.
- Loague, K.M., Yost, R.S., Green, R.E., and Liang, T.C., Uncertainty in a pesticide leaching assessment for Hawaii, *Journal of Contaminant Hydrology*, 4, 139, 1989.
- Loague, K., Lloyd, D., Nguyen, A., Davis, S.N., and Abrams, R.H., A case study simulation of DBCP groundwater contamination in Fresno County, California: 1. Leaching through the unsaturated subsurface, *Journal of Contaminant Hydrology*, 29, 109, 1998a.
- Loague, K., Abrams, R.H., Davis, S.N., Nguyen, A., and Stewart, I.T., A case study simulation of DBCP groundwater contamination in Fresno County, California: 2. Transport in the saturated subsurface, *Journal of Contaminant Hydrology*, 29, 137, 1998b.
- Loague, K., Gander, G.A., VanderKwaak, J.E., Abrams, R.H., and Kyriakidis, P.C., Simulating hydrologic response for the R-5 catchment: A never ending story, *Floodplain Management*, 1, 57, 2000.
- Loague, K., Heppner, C.S., Abrams, R.H., VanderKwaak, J.E., Carr, A.E., and Ebel, B.A., Further testing of the Integrated Hydrology Model (InHM): Event-based simulations for a small rangeland catchment located near Chickasha, Oklahoma, *Hydrological Processes*, 19, 1373, 2005.
- Longley, P.A., Goodchild, M.F., Maguire, D.J., and Rhind, D.W., *Geographic Information Systems and Science*, 2nd ed., John Wiley & Sons, New York, 2005.
- Mayer, S., Ellsworth, T.R., Corwin, D.L., and Loague, K., Estimating transport parameters for non-point pollution transport models, in *Assessment of Non-Point Source Pollution in the Vadose Zone*, Corwin, D.L., Loague, K., and Ellsworth, T.R., (eds.), Geophysical Monograph 108, p. 119, American Geophysical Union Press, Washington, D.C., 1999.
- Miller, E.E. and Miller R.D., Physical theory for capillary flow phenomena, *Journal of Applied Physics*, 4, 324, 1956.
- Montgomery, D.R. and Foufoula-Georgiou, E., Channel network source representation using digital elevation models, *Water Resources Research*, 29, 3925, 1993.
- Mullins, J.A., Carsel, R.F., Scarbough, J.E., and Ivory, A.M., PRZM-2, a model for predicting pesticide fate in the crop root and unsaturated soil zones: Users manual for release 2, EPA-600/R-93/046, U.S. Environmental Protection Agency Environmental Research Laboratory, Athens, GA, 1993.
- Neuman, S.P. and Di Federico, V., Multifaceted nature of hydrogeologic scaling and its interpretation, *Reviews of Geophysics*, 41, 4–1, 2003.

- Nielsen, D.R., Biggar, J.W., and Em, K.T., Spatial variability of field-measured soil–water properties, *Hilgardia*, 42, 215, 1973.
- Noetinger, B., Artus, V., and Zargar, G., The future of stochastic and upscaling methods in hydrology, *Hydrogeology Journal*, 13, 184, 2005.
- Pachepsky, Y., Radcliffe, D.E., and Selim, H.M., (eds.), *Scaling Methods in Soil Physics*, CRC Press, Boca Raton, FL, 2003.
- Penman, H.L., Natural evaporation from open water, bare soil, and grass, *Proceedings of the Royal Society London*, A, 193, 120, 1948.
- Philip, J.R., Field heterogeneity: Some basic issues, *Water Resources Research*, 16, 443, 1980.
- Pilgrim, D.H., Some problems in transferring hydrological relationships between small and large drainage basins and between regions, *Journal of Hydrology*, 65, 49, 1983.
- Prudhomme, C., Reynard, N., and Crooks, S., Downscaling of global climate models for flood frequency analysis: Where are we now? *Hydrological Processes*, 16, 1137, 2002.
- Rawls, W.J., Gish, T.J., and Brakensiek, D.L., Estimating soil water retention from soil physical properties and characteristics, *Advances in Soil Science*, 16, 213, 1991.
- Renschler, C.S., Strategies for implementing natural resource management tools — A geographical information science perspective on water and sediment balance assessment at different scales, PhD dissertation, University of Bonn, Germany, 2000.
- Richards, L.A., Capillary conduction of liquids through porous mediums, *Physics*, 1, 318, 1931.
- Robinson, J.S. and Sivapalan, M., Temporal scales and hydrological regimes: Implications for flood frequency scaling, *Water Resources Research*, 33, 2981, 1997.
- Rodriguez-Iturbe, I., Scale of fluctuation of rainfall models, *Water Resources Research*, 22, 15S, 1986.
- Rodriguez-Iturbe, I. and Gupta, V.K., Introduction, *Journal of Hydrology*, 65, vi, 1983.
- Rodriguez-Iturbe, I. and Rinaldo, A., *Fractal River Basins: Chance and Self-Organization*, Cambridge University Press, Cambridge, United Kingdom, 1997.
- Rubin, Y and Gomez-Hernandez, J.J., A stochastic approach to the problem of upscaling of conductivity in disordered media: Theory and unconditional numerical simulations, *Water Resources Research*, 26, 691, 1990.
- Sanchez-Villa, X., Girardi, J.P., and Carrera, J.A., synthesis of approaches to upscaling of hydraulic conductivities, *Water Resources Research*, 31, 867, 1995.
- Savenije, H.H.G., Equifinality, a blessing in disguise? *Hydrological Processes*, 15, 2835, 2001.
- Schmid, H.P., Footprint modeling for vegetation atmosphere exchange studies: A review and perspective, *Agricultural and Forest Meteorology*, 113, 159, 2002.
- Segol, G., *Classic Groundwater Simulations: Proving and Improving Numerical Models*, Prentice Hall, Englewood Cliffs, NJ, 1994.
- Shaman, J., Stieglitz, M., and Burns, D., Are big basins just the sum of small catchments? *Hydrological Processes*, 18, 3195, 2004.
- Shuttleworth, W.J., Yang, Z.-L., and Arain, M.A., Aggregation rules for surface parameters in global models, *Hydrology and Earth System Sciences*, 1, 217, 1997.
- Sivapalan, M., Process complexity at hillslope scale, process simplicity at the watershed scale: Is there a connection? *Hydrological Processes*, 17, 1037, 2003.
- Sivapalan, M. and Kalma, J.D., Scale problems in hydrology: Contributions of the Robertson workshop, *Hydrological Processes*, 9, 3, 1995.
- Sivapalan, M., Zhang, L., Vertessy, R., and Bloschl, G., Preface: Downward approach to hydrological prediction, *Hydrological Processes*, 17, 2099, 2003a.
- Sivapalan, M., Bloschl, G., Zhang, L., and Vertessy, R., Downward approach to hydrological prediction, *Hydrological Processes*, 17, 2101, 2003b.
- Sivapalan, M., Grayson, R., and Woods, R., Preface: Scale and scaling in hydrology, *Hydrological Processes*, 18, 1369, 2004.
- Skoien, J.O. and Bloschl, G., Sampling scale effects in random fields and implications for environmental monitoring, *Environmental Monitoring and Assessment*, 114, 521, 2006.

- Skoien, J.O., Bloschl, G., and Western, A.W., Characteristic space scales and timescales in hydrology, *Water Resources Research*, 39, doi:10.1029/2002WR001736, 2003.
- Soil Survey Staff, *Soil Taxonomy: A Basic System of Soil Classification for Making and Interpreting Soil Surveys*, 2nd ed., Agriculture Handbook Number 436, U.S. Department of Agriculture, Natural Resources Conservation Service, U.S. Government Printing Office, Washington, D.C., 1999.
- Souther, L.A. and Loague, K., Revisiting the Fresno DBCP case study simulations: The impact of upscaling, *Journal of Environmental Quality*, 29, 1794, 2000.
- Sposito, G., (ed.) *Scale Dependence and Scale Invariance in Hydrology*, Cambridge University Press, Cambridge, United Kingdom, 1998.
- Stewart, I.T. and Loague, K., Development of type transfer functions for regional-scale nonpoint source groundwater vulnerability assessments, *Water Resources Research*, 39, doi:10.1029/2003WR002269, 2003.
- Stewart, I.T. and Loague, K., Assessing ground water vulnerability with the type transfer function model in the San Joaquin valley, California, *Journal of Environmental Quality*, 33, 1487, 2004.
- Stewart, J.B., Engman, E.T., Feddes, R.A., and Kerr, Y., *Scaling Up in Hydrology Using Remote Sensing*, John Wiley and Sons, Chichester, England, 1996.
- Terzaghi, K., *Erdbaumechanik auf Bodenphysikalischer Grundlage*, Franz Deuticke, Vienna, 1925.
- Tsegaye, T.D., Crosson, W.L., Laymon, C.A., Schamschula, M.P., and Johnson, A.B., Application of a neural network-based spatial disaggregation scheme for addressing scaling of soil moisture, in *Scaling Methods in Soil Physics*, Pachepsky, Y., Radcliffe, D.E., and Selim, H.M., (eds.), CRC Press, Boca Raton, FL, 2003, chap. 15.
- Tyler, S.W. and Wheatcraft, S.W., The consequences of fractal scaling in heterogeneous soil and porous media, in *Scaling in Soil Physics: Principles and Applications*, Hillel, D. and Elrick, D.E., (eds.), Soil Science Society of America, Madison, WI, 1990, chap. 8.
- VanderKwaak, J.E., Numerical simulation of flow and chemical transport in integrated surface-subsurface hydrologic systems, PhD dissertation, University of Waterloo, Ontario, Canada, 1999.
- VanderKwaak, J.E. and Loague, K., Hydrologic-response simulations for the R-5 catchment with a comprehensive physics-based model, *Water Resources Research*, 37, 999, 2001.
- Wagenet, R.J. and Hutson, J.L., Scale-dependency of solute transport modeling/GIS applications, *Journal of Environmental Quality*, 25, 499, 1996.
- Webster, R. and Oliver, M.A., *Geostatistics for Environmental Sciences*, John Wiley & Sons, Chichester, England, 2001.
- Western, A.W., Grayson, R.B., and Bloschl, G., Scaling of soil moisture: A hydrologic perspective, *Annual Reviews of Earth and Planetary Sciences*, 30, 149, 2002.
- Wilby, R.L., Wigley, M.L., Conway, D., Jones, P.D., Hewitson, B.C., Main, J., and Wilks, D.S., Statistical downscaling of general circulation model output: A comparison of methods, *Water Resources Research*, 34, 2995, 1998.
- Wood, A.W., Leung, R.R., Sridhar, V., and Lettenmaier, D.P., Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs, *Climate Change*, 62, 189, 2004.
- Wood, E.F., Sivapalan, M., and Beven, K., Similarity and scale in catchment storm response, *Reviews in Geophysics*, 28, 1, 1990.
- Woolhiser, D.A. and Goodrich, D.C., Effects of storm rainfall intensity patterns on surface runoff, *Journal of Hydrology*, 102, 335, 1988.
- Zheng, C. and Gorelick, S.M., Analysis of solute transport in flow fields influenced by preferential flowpaths at the decimeter scale, *Ground Water*, 41, 142, 2003.
- Zheng, C. and Wang P.P., MT3DMS: A Modular Three-Dimensional Multispecies Transport Model for Simulation of Advection, Dispersion, and Chemical Reactions of Contaminants in Groundwater Systems, Release DoD_3.00A, Waterways Experiment Station, U.S. Army Corps of Engineers, Vicksburg, MS, 1998.

26

Accounting for Aquifer Heterogeneity in Solute Transport Modeling: A Case Study from the Macrodispersion Experiment (MADE) Site in Columbus, Mississippi

26.1	Introduction.....	26-2
26.2	Site Description	26-2
26.3	Natural-Gradient Tracer Tests	26-4
26.4	Modeling of the MADE-2 Tritium Tracer Test..... Objectives and Approaches • Model Setup and Boundary Conditions • Assignment of Hydraulic and Transport Properties	26-6
26.5	Analysis of Simulation Results	26-10
	Advection-Dispersion Model • Dual-Domain Mass Transfer Model • Sensitivity Analysis	
26.6	Summary and Concluding Remarks	26-13
	Acknowledgments.....	26-14
	Glossary.....	26-14
	References	26-15
	Further Information	26-17

Chunmiao Zheng
University of Alabama

26.1 Introduction

Field studies at well-instrumented sites have played a preeminent role in our efforts to better understand and characterize solute transport processes in geologic media. In particular, several field tracer tests conducted at well-known sites such as those in Borden, Canada (MacKay et al., 1986), Mobile, Alabama (Molz et al., 1986), Twin Lake, Minnesota (Killey and Moltyaner, 1988), Cape Cod, Massachusetts (LeBlanc et al., 1991) and Columbus, Mississippi (Boggs et al., 1992) have provided new insights and extensive data sets for development and testing of transport theories and mathematical models.

The case study presented in this chapter concerns the natural-gradient tracer tests conducted at the Columbus Air Force Base in Mississippi, southeastern United States, more commonly known as the Macrodispersion Experiment (MADE) site. Since the early 1990s, the data collected at the MADE site has been used extensively by numerous researchers around the world to gain insights into contaminant transport processes in highly heterogeneous aquifers. It has also served as a catalyst for the development of new and improved theories and computer models that enable more accurate description and prediction of contaminant transport and remediation. Recent studies related to the MADE site include Adams and Gelhar (1992), Boggs and Adams (1992), Boggs et al. (1992), Rehfeldt et al. (1992), MacIntyre et al. (1993), Eggleston and Rojstaczer (1998), Berkowitz and Scher (1998), Zheng and Jiao (1998), Harvey and Gorelick (2000), Feehley et al. (2000), Stapleton et al. (2000), Benson et al. (2001), Brauner and Widdowson (2001), Julian et al. (2001), Lu et al. (2002), Zheng and Gorelick (2003), Liu et al. (2004), Barlebo et al. (2004), Bowling et al. (2005), and Molz et al. (2006).

In this chapter, we will first describe the hydrogeologic setting of the MADE site and the details of the three natural-gradient tracer tests that have been conducted at the MADE site since 1986. We will then discuss the application of two modeling approaches, the classical advection-dispersion model and the alternative dual-domain (dual-porosity) mass transfer model, to interpret and analyze the tracer test data. The model application will elucidate the difficulties encountered by the classical advection-dispersion model in characterizing solute transport in an extremely heterogeneous aquifer, and reveal to what extent the dual-domain mass transfer model can be used effectively to lessen and overcome such difficulties. Finally, we will evaluate the sensitivity of the simulation results to key dual-domain model parameters and comment on the applicability of the dual-domain model under field conditions.

26.2 Site Description

Figure 26.1 shows the location of the MADE site located inside the Columbus Air Force Base in Mississippi, Southeastern U.S. The shallow unconfined aquifer that immediately underlies the site consists of alluvial

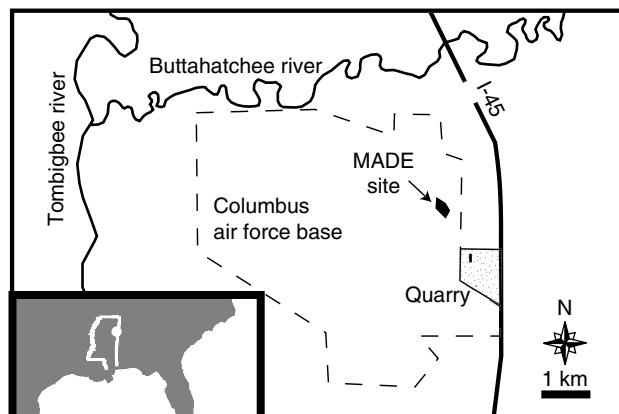


FIGURE 26.1 Location of the MADE site in Columbus, MS.

terrace deposits with an average thickness of 11 m. Pleistocene river deposits associated with nearby Tombigbee and Buttahachee Rivers may account for the heterogeneity observed in the field. The heterogeneous aquifer is composed of poorly sorted to well-sorted sandy gravel and gravelly sand with significant amounts of silt and clay. Sediments are generally unconsolidated. Soil facies occur as irregular lenses and layers having horizontal dimensions ranging up to 8 m and vertical dimensions of less than 1 m. Marine sediments, which belong to the Eutaw Formation and consist of clay, silts, and fine-grained sands, form an aquitard beneath the alluvial aquifer. Figure 26.2 shows a geologic facies model for the MADE site based on geophysical and sedimentological investigations (Bowling et al., 2005). A more detailed description of the site condition is provided in Boggs et al. (1990, 1992).

The spatial distribution of horizontal hydraulic conductivity (K) at the test site was first determined on the basis of 2187 measurements obtained from the borehole flowmeter tests conducted in 58 fully penetrating wells by Rehfeldt et al. (1992). The locations of these flowmeter test wells are shown in Figure 26.3. The vertical spacing of K measurements within each flowmeter well was approximately

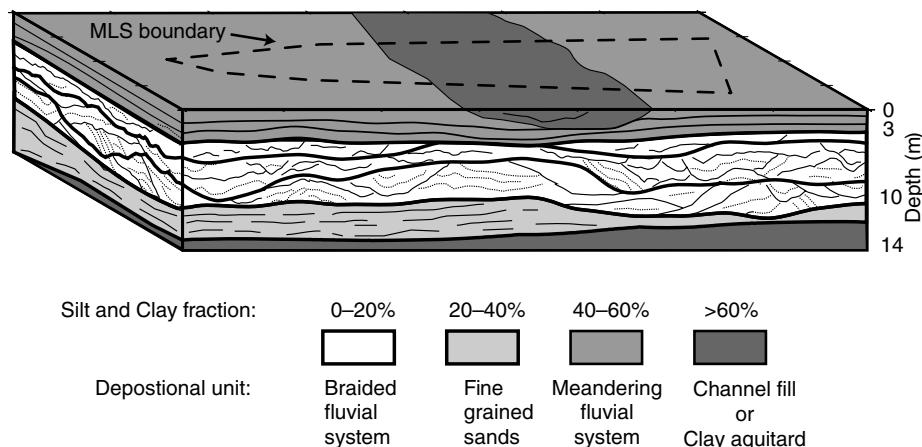


FIGURE 26.2 Geologic facies model of the MADE site based on resistivity and GPR data. (Bowling, J.C., A.B. Rodriguez, D.L. Harry, and C. Zheng, *Ground Water*, 43:890–903, 2005.)

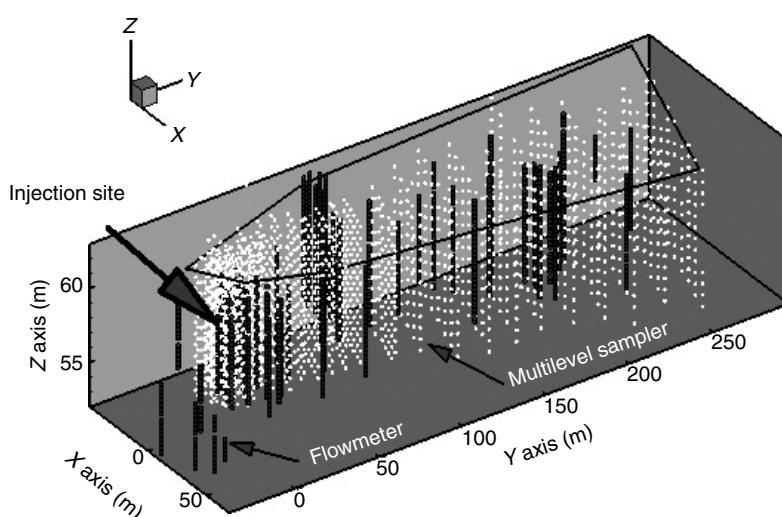


FIGURE 26.3 Three-dimensional view of the monitoring network at the MADE site. Vertical lines consisting of closely spaced dots indicate the locations of flowmeter test wells. Circles indicate multilevel sampler (MLS) locations.

15 cm. Details of the borehole flowmeter tests are described in Rehfeldt et al. (1992). More flowmeter test wells were added in subsequent studies, resulting in a cumulative total of 3925 measurements from 86 test wells (Boggs et al., 1995; Julian et al., 2001). The geometric mean of the 2187 measured horizontal K data is approximately 5×10^{-3} cm/s, but the variations in these data range from two to four orders of magnitude. The overall variance of the natural logarithm of horizontal hydraulic conductivity is 4.5 (Rehfeldt et al., 1992), which is significantly greater than that of any previously reported tracer test sites. For example, the reported variance of $\ln K$ is 0.29 for the Borden site (Sudicky, 1986), 0.26 for the Cape Cod site (LeBlanc et al., 1991), and 0.031 for the Twin Lake site (Killey and Moltyaner, 1988).

The hydraulic head field at the experiment site was monitored with piezometers installed at the flowmeter wells. The general direction of groundwater movement was northward along the y axis of Figure 26.3. The hydraulic head field was subject to complex temporal and spatial variability, resulting from the heterogeneity in the hydraulic conductivity distribution and large seasonal fluctuations of the water table. The averaged hydraulic gradient across the test site was approximately 0.03 with a 20–30% variation in the saturated thickness of the aquifer (Boggs et al., 1992).

26.3 Natural-Gradient Tracer Tests

Three natural-gradient tracer tests have been conducted at this site since 1986. In the first test (MADE-1) conducted between October 1986 to June 1988 (Boggs et al., 1990, 1992), a number of conservative tracers including bromide were injected into five source wells spaced 1 m apart in a linear array. Each well was screened between the elevations of 57.5 and 58.1 m above the mean sea level, or approximately halfway between the water table and the aquifer bottom. At the start of the test, approximately 10 m³ of solution containing a bromide concentration of 2500 mg/l was injected into the source wells at a uniform rate over 48.5 h. The tracer injection was so designed as to mimic a pulse release of tracers into the middle of the saturated aquifer zone with minimal disturbance to the natural flow field (Boggs et al., 1992).

The migration of the bromide plume in three dimensions was monitored through a network of 258 multilevel samplers (MLS) (increased to 328 for subsequent studies) (see Figure 26.3). Each MLS was equipped with 20 to 30 sampling points spaced 0.38 m apart in the vertical dimension. The network used for the MADE-1 test consisted of approximately 6000 sampling points. The groundwater samples collected from MLS were analyzed using high-pressure liquid chromatography (HPLC). Eight sampling events were conducted over a total monitoring period of 20 months. Each sampling event was designed to capture a “snapshot” of the evolving 3-D bromide plume.

Figure 26.4 shows a 3-D visualization of the 503-day snapshot of the bromide plume (Harvey and Gorelick, 2000). A 2-D contour map of vertically integrated concentrations is also shown beneath the 3-D rendering of the plume. The most noticeable aspect is the asymmetrical distribution of the plume along the general flow direction with the high-concentration region remaining close to the injection point (<20 m) and the low-concentration front spreading far down-gradient (>160 m) after 503 days.

In the second test (MADE-2) conducted between June 1990 and September 1991 (Boggs et al., 1993; MacIntyre et al., 1993), a tritium tracer was injected into the aquifer for 48.5 h at a uniform rate of 3.3 l/min, using the same experimental setup established for the MADE-1 test. The injected fluid had a tritium concentration of 55,610 pCi/ml, representing a total injected mass (activity) of 0.5387 Ci. Five snapshots of the tracer plume were taken over a period of 15 months from the 328 MLS at over 6,000 sampling points. The tritium concentration present in field samples was measured with a liquid scintillation analyzer.

Figure 26.5 shows the observed tritium plume snapshot of 328 days after the start of injection. Similar to the observed bromide plume for the MADE-1 test, the high-concentration region of the observed tritium plume stayed close to the injection point while the low-concentration front traveled far down-gradient. The location of the observed maximum concentration after 328 days (4721 pCi/ml) was within several meters of the injection point, while the leading plume edge as defined by the low concentration of 5 pCi/ml traveled more than 200 m from the injection point.

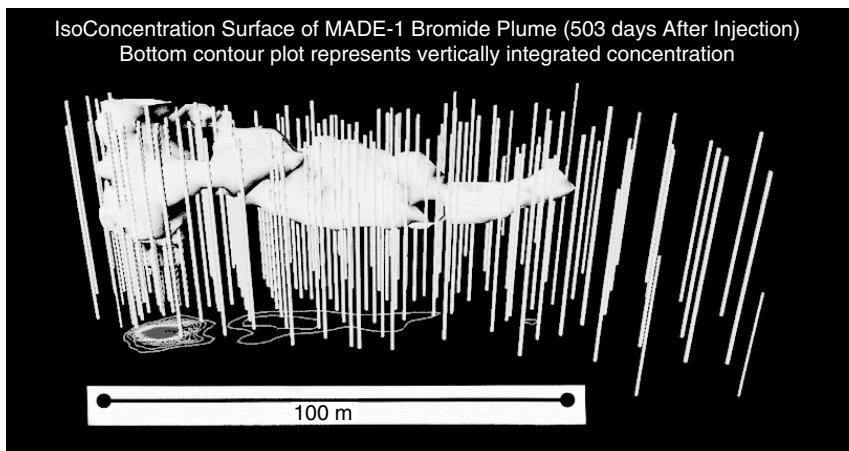


FIGURE 26.4 Observed MADE-1 bromide plume at 503 days after injection. (Harvey, C.F. and S.M. Gorelick, *Water Resour. Res.*, 36:637–650, 2000.)

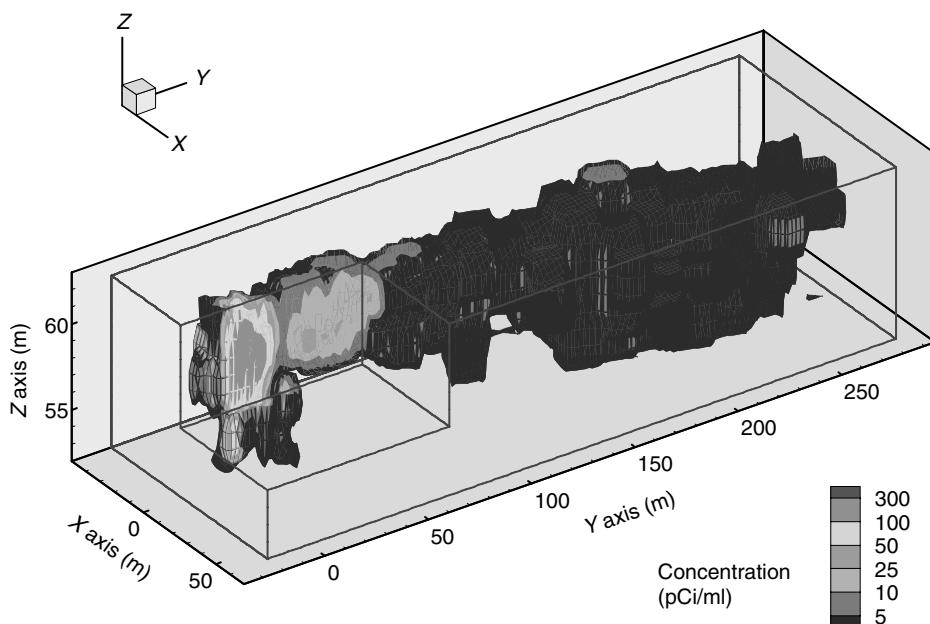


FIGURE 26.5 Observed MADE-2 tritium plume at 328 days after injection.

In the third test (MADE-3), also referred to as the Natural Attenuation Study or NATS, conducted between December 1995 and September 1997 (Boggs et al., 1995; Libelo et al., 1997; Stapleton et al., 2000; Julian et al., 2001), the injection well array was replaced by a source trench measuring approximately $14\text{ m} \times 0.8\text{ m} \times 1.7\text{ m}$. The source trench was emplaced in the aquifer between the elevations of 58.6 and 60.3 m with 30 m^3 of coarse sand coated with six hydrocarbon compounds and bromide. This setup was intended to approximate an instantaneous release for hydrocarbon and bromide tracers. A total of six snapshots of the plume were measured over 20 months.

Figure 26.6 shows a 2-D contour map of the observed “peak” bromide concentrations 152 days after the source release; this map was constructed using the highest concentrations measured among multiple sampling elevations at each MLS location. As in the case of the bromide plume for MADE-1

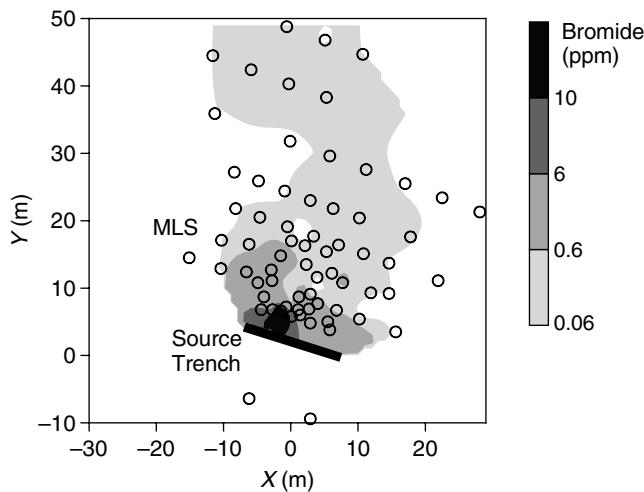


FIGURE 26.6 Observed MADE-3 bromide plume at 152 days after trench source release. The 2D contour map is constructed from the peak concentrations at each MLS location.

and the tritium plume for MADE-2, the bromide plume for MADE-3 shows the presence of a high concentration region close to the source trench and a low concentration front spreading to far-field, truncated by the extent of the sampling. This indicates that the characteristics of the conservative tracer plumes at the MADE site is independent of the method used for introducing the source into the aquifer.

The highly asymmetrical nature of the bromide and tritium plumes observed in the MADE-1, MADE-2, and MADE-3 experiments may be explained by the fact that the aquifer is extremely heterogeneous and the tracer source was located in a zone of relatively low hydraulic conductivity, thereby leading to the trapping of the tracer mass close to the source. The extensive spreading down-gradient may have been caused by the presence of small-scale preferential flow paths as discussed in subsequent sections.

26.4 Modeling of the MADE-2 Tritium Tracer Test

26.4.1 Objectives and Approaches

The data sets from the three natural-gradient tracer tests at the MADE site, described in the preceding section, provided unprecedented opportunities to test and evaluate various theories and mathematical models of solute transport in highly heterogeneous aquifers. The case study presented in this chapter focuses on the modeling of the MADE-2 test, adopted from the work of Zheng and Jiao (1998) and Feehley et al. (2000). The primary objectives of the transport modeling in this case study are to (1) test the applicability of the classical advection-dispersion model (ADM) to represent solute transport processes in the presence of extreme aquifer heterogeneity, and (2) evaluate the usefulness of the dual-domain (dual-porosity) mass transfer model (DDMT) as an alternative to the classical ADM.

The classical advection-dispersion transport equation for conservative tracers is (e.g., Bear, 1961; Zheng and Bennett, 2002),

$$\theta \frac{\partial C}{\partial t} = \frac{\partial}{\partial x_i} \left(\theta D_{ij} \frac{\partial C}{\partial x_j} \right) - \frac{\partial}{\partial x_i} (q_i C) + q_s C_s \quad (26.1)$$

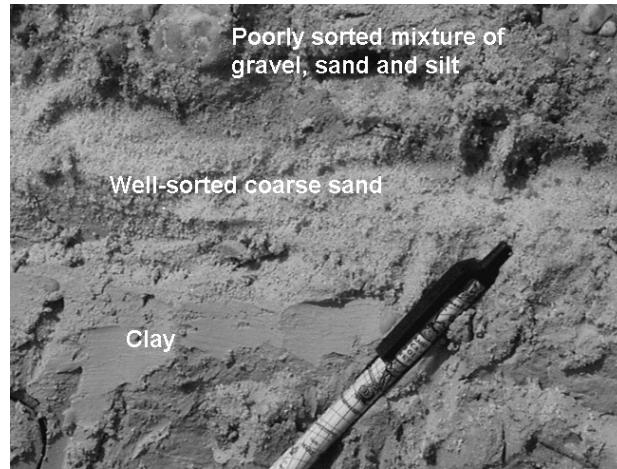


FIGURE 26.7 A newly exposed aquifer cross section near the MADE site indicating the existence of a decimeter-scale preferential flow path.

where C is the solute concentration, θ the porosity, t time, D_{ij} the hydrodynamic dispersion coefficient (including mechanic dispersion and molecular diffusion), q_i the Darcy flux, q_s the rate of fluid sink/source, and C_s the concentration of the fluid sink/source. The classical advection-dispersion model, first introduced by Taylor (1953) and Bear (1961) among others, has been the basis of practical field-scale transport modeling since the early 1970s (e.g., Bredehoeft and Pinder, 1973; Anderson, 1979).

A fundamental assumption embedded in Equation 26.1 is that the mechanical dispersion, caused by variations in seepage velocity due to aquifer heterogeneity, is analogous to molecular diffusion governed by Fick's law. Under this assumption, the mechanical dispersivities can be estimated from the geostatistical properties of the hydraulic conductivity field (e.g., Gelhar and Axness, 1983; Dagan, 1989). For solute transport in relatively homogeneous aquifers, as at the Borden site (MacKay et al., 1986) or the Cape Cod site (LeBlanc et al., 1991), the classic ADM has been shown to be a good one (e.g., Sudicky, 1986). However, for solute transport in extremely heterogeneous aquifers, such as at the MADE site, the classical ADM may be grossly inadequate (e.g., Carrera, 1993; Harvey and Gorelick, 2000; LaBolle and Fogg, 2001). This inadequacy could be attributed to the inability of the classical ADM to account for solute transport along connected preferential pathways that may persist at scales smaller than those that can be practically represented in numerical flow and transport models (Zheng and Gorelick, 2003; Liu et al., 2004). An example of such a small-scale preferential flow path from an aquifer outcrop adjacent to the MADE site is illustrated in Figure 26.7.

An alternative to the single-porosity advection-dispersion model is the dual-porosity or dual-domain mass transfer model, given below for conservative tracers (e.g., Coats and Smith, 1964; van Genutchen and Wierenga, 1976; Zheng and Bennett, 2002):

$$\begin{aligned} \theta_m \frac{\partial C_m}{\partial t} + \theta_{im} \frac{\partial C_{im}}{\partial t} &= \frac{\partial}{\partial x_i} \left(\theta_m D_{ij} \frac{\partial C_m}{\partial x_j} \right) - \frac{\partial}{\partial x_i} (q_i C_m) + q_s C_s \\ \theta_{im} \frac{\partial C_{im}}{\partial t} &= \beta (C_m - C_{im}) \end{aligned} \quad (26.2)$$

where θ_m is the porosity of the “mobile” domain, or pore spaces filled with mobile water, θ_{im} the porosity of the “immobile” domain, or pore spaces filled with immobile water, β the mass transfer rate coefficient between the mobile and immobile domains, C_m the concentration in the mobile domain, and C_{im} the concentration in the immobile domain.

The DDMT partitions transport into two domains: transport in the mobile domain is primarily by advection, while transport in the immobile domain is mostly by diffusion. The mobile and immobile domains are defined by two porosities: mobile (θ_m) and immobile (θ_{im}) porosities, with the total porosity $\theta = \theta_m + \theta_{im}$. The connection between the zones of mobile and immobile fluids is governed by a first-order mass transfer equation, with the transfer rate proportional to the concentration difference between the two zones multiplied by a coefficient of proportionality β . Other factors affecting mass transfer include sorption, measurement scale, fluid velocity, and characteristics of preferential flow paths (Griffioen et al., 1998; Haggerty et al., 2004; Gorelick et al., 2005). The larger the rate coefficient β , the faster mass transfer occurs between the mobile and immobile zones for a given concentration difference. As β approaches infinity, the connection is instantaneous, and the system acts like a single-porosity ADM with the porosity approaching total porosity. Conversely, as β approaches zero, the system behaves as a single-porosity ADM with the total porosity equal to θ_m . The mass transfer rate coefficient can be determined through laboratory and field procedures of Clothier et al. (1992), Jaynes et al. (1995), Hollenbeck et al. (1999), and Haggerty et al. (2001) or through trial-and-error model calibration (Feehley et al., 2000; Julian et al., 2001).

26.4.2 Model Setup and Boundary Conditions

Zheng and Jiao (1998) applied the classical ADM; Feehley et al. (2000) applied both the classical ADM and alternative DDMT to simulate and analyze the MADE-2 tritium tracer test. The 3-D finite-difference model grid used by Feehley et al. (2000) consists of 55 columns, 153 rows, and 22 layers, covering a model domain of 110 m × 306 m × 11 m. The grid spacing is 2 m × 2 m in the horizontal direction and 0.5 m in the vertical direction. Specified-head boundaries along the north and south, and no-flow boundaries to the east and west, bound the model domain laterally. The Eutaw Formation acts as a no-flow boundary at the bottom and a uniform recharge rate is specified at the water table. Transport simulation was conducted using the mesh developed for the flow model. The boundary conditions for the transport model are zero-mass flux at all boundaries except at the northern boundary where dispersive flux is assumed to be zero and advective flux is determined by outflow rate and concentration.

Only steady-state flow simulation was attempted using the MODFLOW code (McDonald and Harbaugh, 1988; Harbaugh et al., 2000) because the detailed data required for transient calibration (such as specific yield, storage coefficient, and time-dependent recharge distribution) were not available. Values along the specified-head boundaries were calibrated to achieve the best match between the calculated and observed heads at 48 piezometers. The calibration was done manually by minimizing the root mean of squared differences between the calculated and observed heads. More details of model development and calibration are provided in Zheng and Jiao (1998) and Feehley et al. (2000).

As noted above, the tritium tracer was injected into the aquifer over a period of two days. This case study only considered the 328-day plume snapshot because the plumes at earlier times were not as extensive, and the plume at 440 days was not sampled completely. To accommodate the injection in the flow model, the total transport simulation time of 328 days was approximated by two steady-state stress periods, one representing conditions during the two-day period of injection and the other representing conditions during the subsequent 326 days. In the first stress period, the tracer injection through five wells was represented by three cells with a total injection rate of 3.3 l/min and a concentration of 55,610 pCi/ml. The transport model was solved using the MT3DMS code (Zheng and Wang, 1999), which includes both the standard advection-dispersion approach and the alternative dual-domain mass transfer approach.

26.4.3 Assignment of Hydraulic and Transport Properties

The ordinary kriging technique was used to construct the hydraulic conductivity distributions needed by the numerical model. Ordinary kriging is a simple interpolation technique commonly used in field

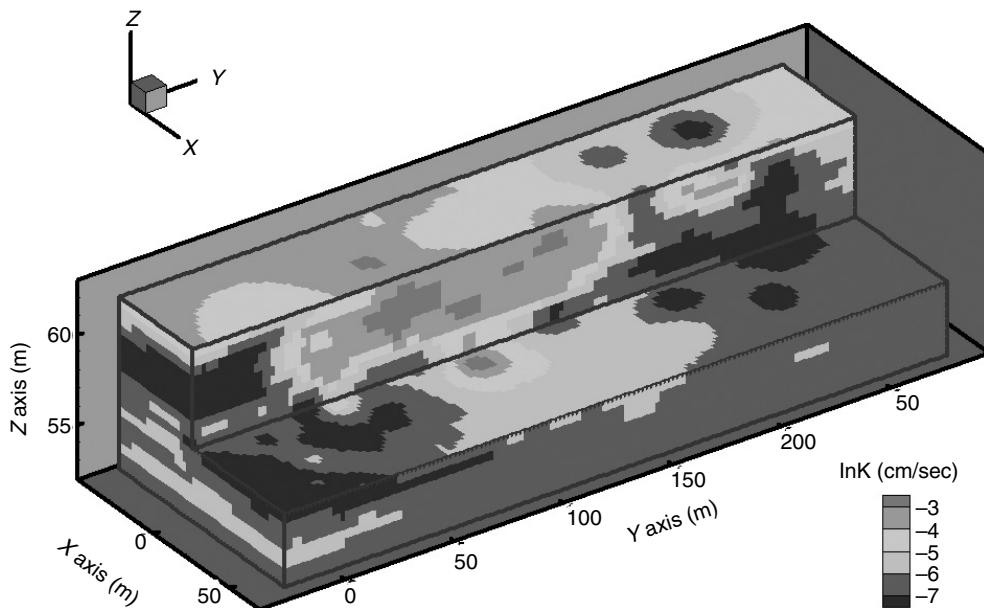


FIGURE 26.8 3D hydraulic conductivity field generated with ordinary kriging from over 3000 measurements. (Feehley C.E., C. Zheng, and F.J. Molz, *Water Resour. Res.*, 36:2501–2515, 2000.)

applications. In the kriging approach, the statistical parameters, such as the mean, variance and correlation scales, were determined from the flowmeter data and input into a three-dimensional ordinary kriging program (Deutsch and Journel, 1998). The result is a deterministic three-dimensional hydraulic conductivity field as shown in Figure 26.8. The corner of the cutaway block represents the approximate location of the observed tritium plume centerline. A stochastic fractal based method was also used by Feehley et al. (2000) to investigate the effect of different K generation schemes on the simulated plumes.

An average total porosity of 0.35 was used for the entire model area, which is close to the value of 0.32 determined from the soil cores collected from the test site (Boggs et al., 1990). The difference in the two values was intended to account for potential consolidation during handling of the core samples (Adams and Gelhar, 1992). A uniform recharge rate of 1.4×10^{-4} m/day was used for the steady-state flow model. Different values of longitudinal dispersivity were tested. The ratios of the transverse and vertical dispersivities to the longitudinal dispersivity were fixed at 0.01 and 0.001, respectively, consistent with the very limited transverse and vertical dispersion as observed at the MADE site and other field sites (Gelhar et al., 1992).

For the dual-domain mass transfer model, it is also necessary to specify mobile and immobile porosities and a mass transfer rate coefficient. These parameters were estimated through trial-and-error calibration based on a comparison of observed and calculated tritium plumes. In all simulations conducted, a mobile porosity of one-eighth of the total porosity was judged to produce the most consistent match between observed and calculated plumes. The mass transfer rate coefficient was found to be sensitive to the hydraulic conductivity distribution. For simulations with the kriged K field, a uniform value of 0.001 day^{-1} was used for the entire model domain. The effect of the mass transfer rate coefficient on the simulated plume is addressed subsequently.

The model domain was divided into six equally spaced zones along the general flow direction (the y axis) in which the mobile-domain mass (i.e., the zeroth moment) of the simulated plume was calculated and compared with that of the observed plume. The mass of each zone is determined by computing the mass per zone and dividing it by the total mass of all zones. This relative mass instead of absolute mass is used for the convenience of comparison.

26.5 Analysis of Simulation Results

26.5.1 Advection-Dispersion Model

For the advection-dispersion model, a single porosity equal to the total porosity of 0.35 was used. In a previous study by Zheng and Jiao (1998), a longitudinal dispersivity value of 5.0 m was found to result in a reasonable match of the overall plume configuration to the observed plume above a certain concentration limit. The longitudinal dispersivity was varied spatially during the calibration process, but the overall match was not significantly improved; thus, those results are not reported here.

Figure 26.9 shows the calculated tritium plume based on the single-porosity ADM and the kriged hydraulic conductivity field. A large amount of mass is retained near the injection wells, as was observed in the measured data (Figure 26.5). However, the simulated plume fails to represent rapid, anomalous spreading at low concentrations. Furthermore, the resulting plume does not reproduce the fringy appearance of the observed plume. The distribution of calculated relative mass in the different zones (shown in Figure 26.10 as columns with a darker shade) does not closely mimic the observed distribution (shown as columns with a light shade). In the injection area, zone 1, the relative mass of the calculated plume is much higher than that of the observed plume. There is a steady decline in calculated relative mass with increasing distance from the injection zone, while the relative mass for the observed plume shows a sharp decrease from zone 1 to zone 2, but increases and decreases again with distance. In the far-field (zone 5 and zone 6), the relative mass for the calculated plume is zero even though the observed plume shows significant mass in these zones.

Several possible reasons exist for the discrepancies between the observed and calculated plumes. First, the characterization of solute transport in highly heterogeneous porous media is inherently difficult; it is practically impossible to represent heterogeneity at a level where its effects on solute transport will be reproduced completely. Measured hydraulic conductivity distributions are incorporated into models by assigning a different hydraulic conductivity value to each model node, generally using interpolation. Interpolation methods, such as ordinary kriging, generate a unique hydraulic conductivity field, preserving local accuracy only near data points while smoothing data in areas of sparse sampling

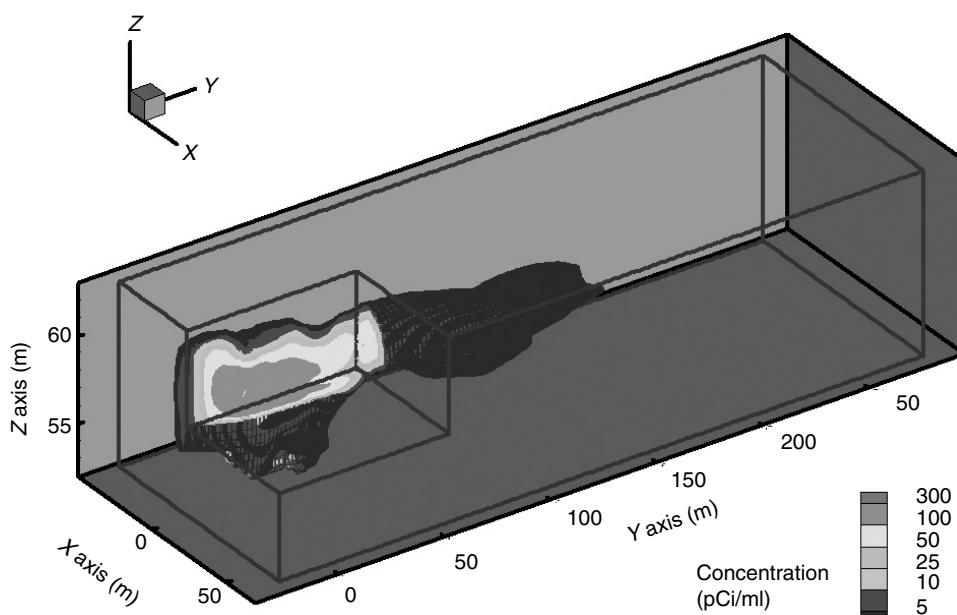


FIGURE 26.9 Simulated tritium plume at 328 days after injection with the advection-dispersion model (Feehley C.E., C. Zheng, and F.J. Molz, *Water Resour. Res.*, 36:2501–2515, 2000.). The longitudinal dispersivity used is 5 m.

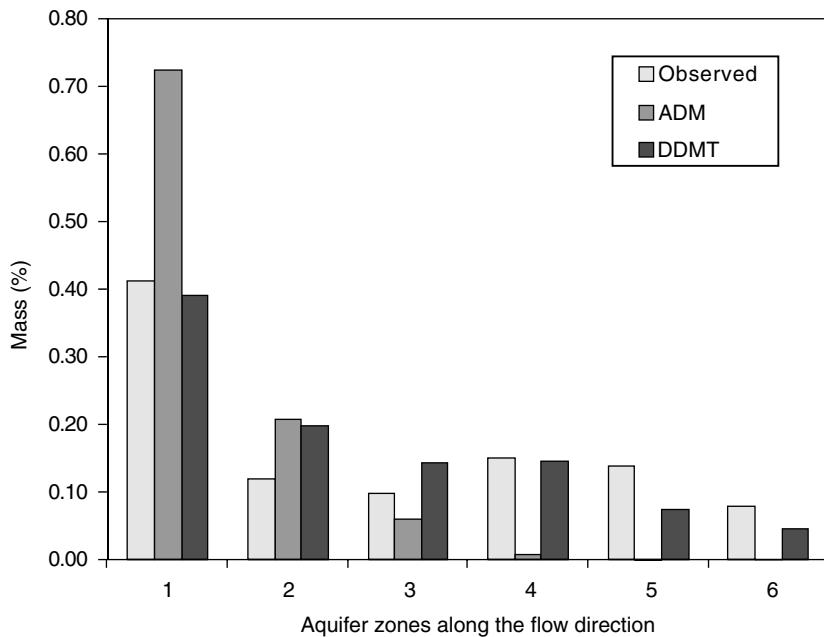


FIGURE 26.10 Bar graph showing results of the advection-dispersion model (ADM) and the dual-domain mass transfer model (DDMT). (Feehley C.E., C. Zheng, and F.J. Molz, *Water Resour. Res.*, 36:2501–2515, 2000.)

(Deutsch and Journel, 1998). Commonly, therefore, these methods fail to reproduce preferential pathways, which tend to be characterized by sharp local contrast in hydraulic conductivity. Furthermore, even if a more sophisticated hydraulic conductivity generation methodology were used, it would generally still be impossible to achieve a sufficiently accurate representation of the true hydraulic conductivity field and the resulting velocity field, particularly if the macroscopic or local scale heterogeneities in the aquifer are smaller than the numerical grid spacing, or if they occur in preferential directions. The failure to account for preferential pathways eliminates the mechanism for rapid spreading of a part of the tracer mass; this is probably the reason that the simulated plume in Figure 26.9 does not show mass transported a significant distance downstream.

26.5.2 Dual-Domain Mass Transfer Model

The second simulation was based on the DDMT with the same kriged hydraulic conductivity distribution as used in the first simulation. The mass transfer rate coefficient used for this simulation is 0.001 day^{-1} , and the mobile porosity was set to one-eighth of the total porosity. The longitudinal dispersivity used in the dual-domain mass transfer model is 1.0 m, compared to 5.0 m used in the single-porosity ADM. Note that the dispersivity is the driving force for all mechanical dispersion and mixing in the ADM. However, the dispersivity becomes secondary in the DDMT as much of the mechanical dispersion and mixing are accounted for by mass transfer between the mobile and immobile domains at each model cell. Thus, the dispersivity used in this simulation is necessarily smaller than that in the previous simulation, and is intended to account for the randomly occurring local dispersion not already accommodated by the mass transfer process. Additional analysis has shown that the simulated plume in the DDMT is not sensitive to the dispersivity value.

The overall configuration of the simulated plume, shown in Figure 26.11, resembles that of the observed plume, shown in Figure 26.5. There is a large amount of mass trapped near the source area, and also sufficient downstream spreading of low concentration fronts. The narrowing of the plume front as simulated by the ADM (Figure 26.9) has disappeared. The relative mass for the DDMT simulation is shown in

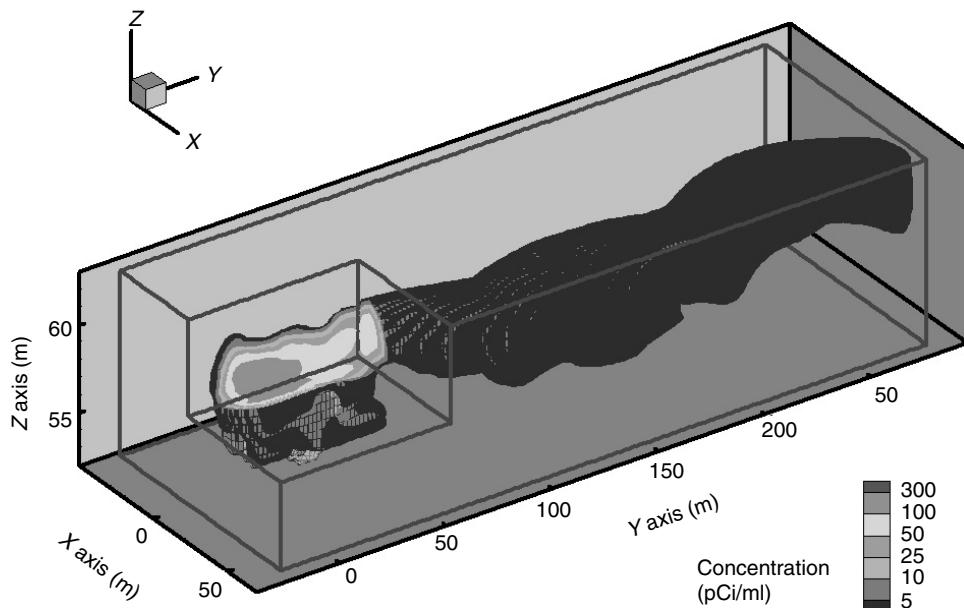


FIGURE 26.11 Simulated tritium plume at 328 days after injection with the dual-domain mass transfer model. (Feehley C.E., C. Zheng, and F.J. Molz, *Water Resour. Res.*, 36:2501–2515, 2000.). The longitudinal dispersivity used is 1 m; the ratio of mobile/total porosities used is 1/8; and the mass transfer rate coefficient is 0.001 day^{-1} .

Figure 26.10 as columns with the darkest shade. The calculated relative mass in each zone closely resembles that of the observed plume. While the match is not exact (the simulated plume is considerably smoother than the observed plume), the result is much more satisfying than that from the previous simulation based on the ADM.

By simulating a mobile and an immobile domain, the dual-domain mass transfer model allows for advective transport to occur in the mobile region at velocities increased in inverse proportion to the ratio of mobile/immobile porosities, thus providing a mechanism for rapid spreading of a certain amount of mass downstream, while retaining a bulk of mass near the source area. The smooth contour surface appears for the same reason as for the ADM simulation (i.e., due to the use of the kriging-generated hydraulic conductivity field), but the DDMT simulation shows a significant improvement in matching the overall size and shape of the observed plume compared to the ADM simulation.

26.5.3 Sensitivity Analysis

Many DDMT simulations, based on the kriging-generated hydraulic conductivity distribution, were run with different combinations of the dual-domain model parameters. The results of a sensitivity analysis showing how relative mass varies with changes in mass transfer rate coefficient are shown in Figure 26.12. In all the runs shown in Figure 26.12, mobile porosity, θ_m , was set to one-eighth of the total porosity. When the mass transfer rate coefficient is low, the transport regime approaches the behavior of a single-domain system having a total porosity equal to the mobile porosity of the dual-domain system. Therefore, the plume moves downstream at a faster rate, causing higher relative mass to appear in the last three zones. If the mass transfer rate coefficient is increased substantially, the connection between the domains becomes very rapid. The system acts like a single-domain regime having a porosity approaching the total porosity of the dual-domain system. Thus the plume travels through the domain at a slower rate; mass is trapped in the first zone due to the faster connection between mobile and immobile zones, and less mass moves downstream. These results suggest that a combination of mass transfer rate coefficients might best model the characteristics of the observed plume.

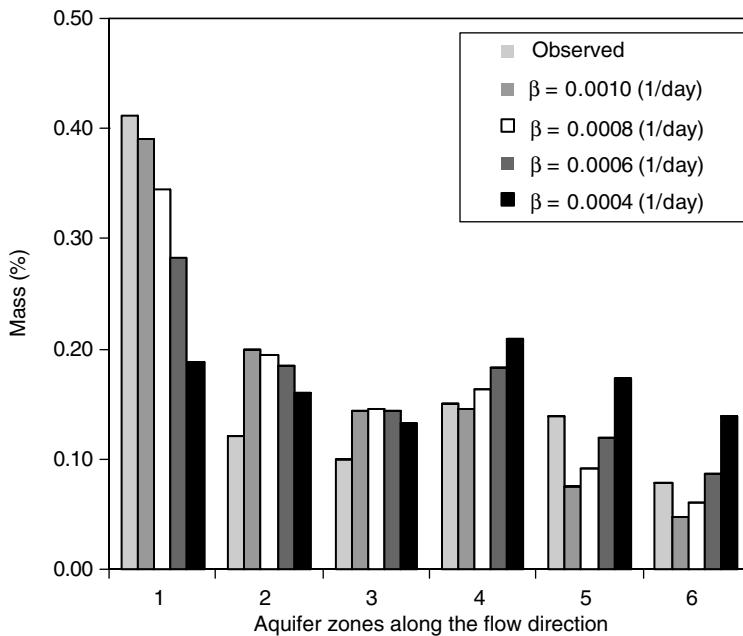


FIGURE 26.12 Bar graph showing the sensitivity of calculated relative mass to changes in mass transfer rate coefficient for the dual-domain mass transfer model. (Feehley C.E., C. Zheng, and F.J. Molz, *Water Resour. Res.*, 36:2501–2515, 2000.) Mobile porosity equal to one-eighth of total porosity is used.

The results of an analysis of how relative mass varies with changes in the ratio of mobile/total porosity are shown in Figure 26.13. The mass transfer rate coefficient of 0.0006 day^{-1} is used in the model and the ratio of mobile over total porosities is allowed to vary. It is clear that the calculated relative mass is not particularly sensitive to changes in the ratio of mobile over total porosities within the range examined (1 : 4–1 : 8).

26.6 Summary and Concluding Remarks

Three natural-gradient tracer tests have been conducted at the MADE site since 1986. The plumes for the conservative tracers used in all three tests exhibited a strongly asymmetrical pattern, with a more concentrated region remaining close to the source and a front of dilute concentration traveling far down-gradient. The non-Gaussian behavior was apparently caused by the extreme heterogeneity in the hydraulic conductivity distribution, which had a natural logarithm variance of 4.5. The same behavior was observed regardless of whether the tracers were introduced to the aquifer by injection wells or through an excavated trench.

A fundamental difficulty with applying the classical Fickian advection-dispersion model to the MADE site appears to be the existence of preferential flow pathways (high conductivity lenses) whose scale (cm to dm) is smaller than the grid spacing ($0.5 \sim 2 \text{ m}$) of the numerical model (e.g., Figure 26.7). Without hydraulic conductivity data sufficient to delineate these preferential flow pathways explicitly, and without the grid resolution required to incorporate them fully in the numerical model, the advection-dispersion approach cannot reproduce the observed plume behavior adequately, even when a large number of K measurements are available. More discussion on this subject can be found in Zheng and Gorelick (2003) and Liu et al. (2004).

The results of the dual-domain mass transfer model show a significant improvement over the advection-dispersion model in matching the observed plume. While the dual domain model does not represent preferential pathways explicitly, the possibility of reducing porosity for the mobile domain allows for

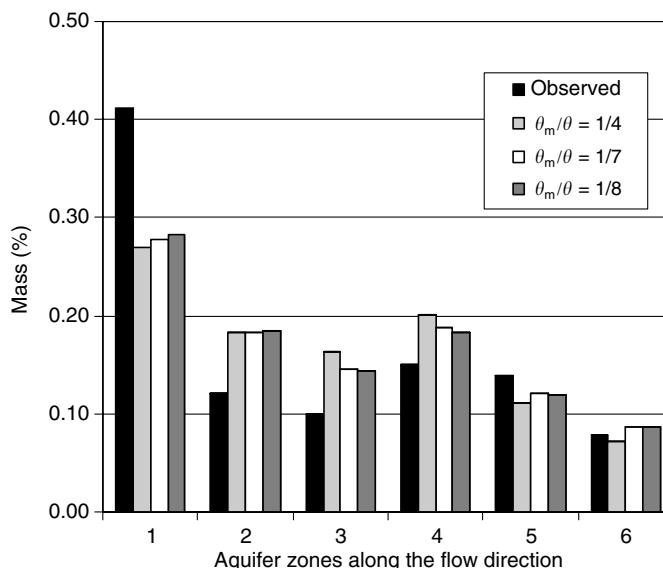


FIGURE 26.13 Bar graph showing the sensitivity of calculated relative mass to changes in mobile porosity ratio for the dual-domain mass transfer model. (Feehley C.E., C. Zheng, and F.J. Molz, *Water Resour. Res.*, 36:2501–2515, 2000.) Mass transfer rate coefficient of 0.0006 day^{-1} is used.

rapid solute movement, mimicking the effects of movement along preferential pathways. At the same time, the storage of mass in the immobile domain adequately represents the retention of mass in low-permeability regions of the subsurface. The DDMT simulation results in a satisfactory match between calculated and observed plumes and mass distributions, particularly in areas of low concentration; this represents a significant improvement over the ADM simulation. It is noteworthy that although this case study focuses only on the MADE-2 tritium plume, the DDMT approach has also been applied successfully to the MADE-1 and MADE-3 bromide plumes, respectively, by Harvey and Gorelick (2000) and Julian et al. (2001) with dual-domain model parameters similar to those used in this case study.

Acknowledgments

The funding for some of the materials presented in this chapter was provided by grants from the National Science Foundation and the U.S. Army Engineer Research and Development Center. Many individuals have contributed to the research discussed in this chapter, including Fred Molz, Steve Gorelick, Gaisheng Liu, Erin Feehley, Mark Boggs, Hank Julian, Tom Stauffer, Mark Dortch, Charles Harvey, Mary Hill, and Eileen Poeter. The author is grateful to Prabhakar Clement for reviewing this manuscript.

Glossary

Advection: Movement of solutes in groundwater with average seepage velocities.

Advection-dispersion model (ADM): A mathematical model for solute transport that is characterized by advection and hydrodynamic dispersion.

Dual porosity: A term used to describe the type of porous media where the total porosity can be separated into mobile and immobile porosities; equivalent to dual domain.

Dual-domain mass transfer model (DDMT): A mathematical model that conceptualizes solute transport as occurring in two overlapping domains: advection and hydrodynamic dispersion in the mobile domain, and rate-limited mass transfer between the immobile and mobile domains.

Fick's law: A mathematical equation that quantifies the diffusive mass flux as proportional to the concentration gradient.

Flowmeter test: An aquifer test technique in which the horizontal flow at multiple elevations in a single borehole is measured and differentiated to determine the vertical profile of hydraulic conductivity at the borehole location.

Heterogeneity: Variation of any aquifer property such as hydraulic conductivity with space or time.

Hydrodynamic dispersion: Sum of mechanical dispersion and molecular diffusion.

Immobile porosity: Ratio of the pore space filled by immobile or stagnant fluids to the total pore space.

Mass transfer: Exchange of solute mass between two regions such as between the mobile and immobile domains.

Mechanical dispersion: Spreading of solutes in groundwater due to variations in flow velocities.

Mobile porosity: Ratio of the pore space filled by mobile fluids to the total pore space.

Molecular diffusion: Movement of solutes in fluids induced by concentration gradients.

Natural gradient tracer test: A type of tracer tests in which conservative or reactive tracers are injected into the aquifer under natural conditions and their movement is monitored by a network of monitoring and sampling wells.

Ordinary kriging: A geostatistical method for estimating an unknown value from neighboring sampled data based on the minimization of interpolation error variance.

Preferential flow paths: Parts of porous media that are comprised of high conductivity materials to permit flow and/or solutes to move through preferentially.

Sensitivity analysis: A numerical procedure to quantify the effect of change in one variable on another variable.

References

- Adams, E.E. and L.W. Gelhar, Field study of dispersion in a heterogeneous aquifer, 2, Spatial moments analysis, *Water Resour. Res.*, 28:3293–3307, 1992.
- Anderson, M.P., Using models to simulate the movement of contaminants through ground water flow systems, *Crit. Rev. Environ. Control*, 9:97–156, 1979.
- Barlebo, H.C., M.C. Hill, and D. Rosbjerg, Investigating the macrodispersion experiment (MADE) site in Columbus, Mississippi, using a three-dimensional inverse flow and transport model, *Water Resour. Res.*, 40:W04211, doi:10.1029/2002WR001935, 2004.
- Bear, J., On the tensor form of dispersion in porous media, *J. Geophys. Res.*, 66:1185–1197, 1961.
- Benson, D.A., R. Schumer, M.M. Meerschaert, and S.W. Wheatcraft, Fractional dispersion, Levy motion, and the MADE tracer tests, *Transport in Porous Med.*, 42:211–240, 2001.
- Berkowitz, B. and H. Scher, Theory of anomalous chemical transport in random fracture network, *Phys. Rev. E*, 57:5858–5869, 1998.
- Boggs, J.M. and E.E. Adams, Field study of dispersion in a heterogeneous aquifer, 4, Investigation of adsorption and sampling bias, *Water Resour. Res.*, 28:3325–3336, 1992.
- Boggs, J.M., S.C. Young, D.J. Benton, and Y.C. Chung, Hydrogeologic characterization of the MADE site, interim report EN-6915, Electric Power Res. Inst., Palo Alto, CA, 1990.
- Boggs, J.M., S.C. Young, and L.M. Beard, Field study of dispersion in a heterogeneous aquifer, 1, Overview and site description, *Water Resour. Res.*, 28:3281–3291, 1992.
- Boggs, J.M., L.M. Beard, S.E. Long, M.P. McGee, W.G. MacIntyre, C.P. Antworth, and T.B. Stauffer, Database for the Second Macrodispersion Experiment (MADE-2), Technical Report TR-102072, Electric Power Res. Inst., Palo Alto, CA, 1993.
- Boggs, J.M., J.A. Schroeder, and S.C. Young, Data to support model development for natural attenuation study. Report No. WR28-2-520-197. TVA Engineering Laboratory, Tennessee Valley Authority, Norris, TN, 1995.
- Bowling, J.C., A.B. Rodriguez, D.L. Harry, and C. Zheng, Delineating alluvial aquifer heterogeneity using resistivity and GPR data, *Ground Water*, 43:890–903, 2005.

- Brauner, J.S. and M.A. Widdowson, Numerical simulation of a natural attenuation experiment with a petroleum hydrocarbon NAPL source, *Ground Water*, 39:939–952, 2001.
- Bredehoeft, J.D., and G.F. Pinder, Mass transport in flowing groundwater, *Water Resour. Res.*, 9:194–210, 1973.
- Carrera, J., An overview of uncertainties in modeling groundwater solute transport, *J. Contam. Hydrol.*, 13:23–48, 1993.
- Clothier, B.E., M.B. Kirkham, and J.E. Mclean, In situ measurements of the effective transport volume for solute moving through soil, *Soil Sci. Soc. Am. J.*, 56:733–736, 1992.
- Coats, K.H. and B.D. Smith, Dead-end pore volume and dispersion in porous media, *Soc. Petrol. Eng. J.*, 4:73–84, 1964.
- Dagan, G., *Flow and Transport in Porous Formation*, Springer-Verlag, New York, 465 p., 1989.
- Deutsch, C.V. and A.G. Journel, *GSLIB: Geostatistical Software Library and User's Guide*, 2nd edn., Oxford University Press, New York, 1998.
- Eggleston, J. and S. Rojstaczer, Identification of large-scale hydraulic conductivity trends and the influence of trends on contaminant transport, *Water Resour. Res.*, 34:2155–2168, 1998.
- Feehley C.E., C. Zheng, and F.J. Molz, A dual-domain mass transfer approach for modeling solute transport in heterogeneous porous media, application to the MADE site, *Water Resour. Res.*, 36:2501–2515, 2000.
- Gelhar, L.W. and C.L. Axness, Three-dimensional stochastic analysis of macrodispersion in aquifers, *Water Resour. Res.*, 19:161–189, 1983.
- Gelhar, L.W., C. Welty, and K. Rehfeldt, A critical review of data on field-scale dispersion in aquifers, *Water Resour. Res.*, 28:1955–1974, 1992.
- Gorelick, S.M., G. Liu, and C. Zheng, Quantifying mass transfer in permeable media containing conductive dendritic networks, *Geophys. Res. Lett.*, 32:L18402, doi:10.1029/2005GL023512, 2005.
- Griffioen, J.W., D.A. Barry, and J.Y. Parlange, Interpretation of two-region model parameters, *Water Resour. Res.*, 34:373–384, 1998.
- Haggerty, R., S.W. Fleming, L.C. Meigs, and S.A. McKenna, Tracer tests in a fractured dolomite. 2. Analysis of mass transfer in single-well injection-withdrawal tests, *Water Resour. Res.*, 37:1129–1142, 2001.
- Haggerty, R., C.F. Harvey, C.F. von Schwerin, and L.C. Meigs, What controls the apparent timescale of solute mass transfer in aquifers and soils? A comparison of diverse experimental results, *Water Resour. Res.*, 40: W01510, doi:10.1029/2002WR001716, 2004.
- Harbaugh, A.W., E.R. Banta, M.C. Hill, and M.G. McDonald, MODFLOW-2000: U.S. Geological Survey Modular Ground-water Model — User Guide to Modularization Concepts and the Ground-Water Flow Processes, U.S. Geological Survey Open-File Report 00-92, 121 p., 2000.
- Harvey, C.F. and S.M. Gorelick, Rate-limited mass transfer or macrodispersion: which dominates plume evolution at the Macrodispersion Experiment (MADE) site? *Water Resour. Res.*, 36:637–650, 2000.
- Hollenbeck, K.J., C.F. Harvey, R. Haggerty, and C.J. Werth, A method for estimating distribution of mass transfer rate coefficients with application to purging batch experiments, *J. Contam. Hydrol.*, 37:367–388, 1999.
- Jaynes, D.B., S.D. Logson, and R. Horton, A field method for measuring mobile/immobile water content and solute transport rate coefficient, *Soil Sci. Soc. Am. J.*, 59:352–365, 1995.
- Julian, H.E., J.M. Boggs, C. Zheng, and C.E. Feehley, Numerical simulation of a natural gradient tracer experiment for the Natural Attenuation Study: flow and physical transport, *Ground Water*, 39:534–545, 2001.
- Killey, R.W.D. and G.L. Moltyaner, Twin Lake tracer tests: methods and permeabilities, *Water Resour. Res.*, 24:1585–1613, 1988.
- LaBolle, E.M. and G.E. Fogg, Role of molecular diffusion in contaminant migration and recovery in an alluvial aquifer system, *Transport Porous Med.*, 42:155–179, 2001.
- LeBlanc, D.R., S.P. Garabedian, K.M. Hess, L.W. Gelhar, R.D. Quadri, K.G. Stollenwerk, and W.W. Wood, Large-scale natural gradient tracer test in sand and gravel, Cape Cod, Massachusetts. 1. Experimental design and observed tracer movement, *Water Resour. Res.*, 27:895–910, 1991.

- Libelo, E.L., T.B. Stauffer, M.A. Geer, W.G. MacIntyre, and J.M. Boggs, A field study to elucidate processes involved in natural attenuation, 4th International In Situ and On-Site Bioremediation Symposium, New Orleans, LA, 1997.
- Liu, G., C. Zheng, and S.M. Gorelick, Limits of applicability of the advection-dispersion model in aquifers containing high-conductivity channels, *Water Resour. Res.*, 40:W08308, doi:10.1029/2003WR002735, 2004.
- Lu, S., F.J. Molz, and G.J. Fix, Possible problems of scale dependency in applications of the three-dimensional fractional advection-dispersion equation to natural porous media, *Water Resour. Res.*, 38, doi:10.1029/2001WR000624, 2002.
- MacIntyre, W.G., J.M. Boggs, C.P. Antworth, and T.B. Stauffer, Degradation kinetics of aromatic organic solutes introduced into a heterogeneous aquifer, *Resour. Res.*, 29:4045–4051, 1993.
- MacKay, D.M., D.L. Freyberg, P.V. Roberts, and J.A. Cherry, A natural gradient experiment on solute transport in a sand aquifer. 1. Approach and overview of plume movement, *Water Resour. Res.*, 22:2017–2029, 1986.
- McDonald, M.G. and A.W. Harbaugh, A modular three-dimensional finite-difference ground-water flow model, *U.S. Geological Survey Techniques of Water Resources Investigations*, Book 6, U.S. Geological Survey, Reston, VA, 1988.
- Molz, F.J., O. Guven, J.G. Melville, R.D. Crocker, and K.T. Matteson, Performance, analysis, and simulation of a two-well tracer test at the Mobile site, an examination of scale-dependent dispersion coefficients, *Water Resour. Res.*, 22:1031–1037, 1986.
- Molz, F.J., C. Zheng, S.M. Gorelick, and C.F. Harvey, Discussion of “Investigating the Macrodispersion Experiment (MADE) site in Columbus, Mississippi, using a three-dimensional inverse flow and transport model” by H.C. Barlebo, M.C. Hill, and D. Rosbjerg, *Water Resour. Res.*, doi:10.1029/2005WR004265, 2006.
- Rehfeldt, K.R., J.M. Boggs, and L.W. Gelhar, Field study of dispersion in a heterogeneous aquifer. 3. Geostatistical analysis of hydraulic conductivity, *Water Resour. Res.*, 28:3309–3324, 1992.
- Stapleton, R.D., G.S. Sayler, and J.M. Boggs, Changes in subsurface catabolic gene frequencies during natural attenuation of petroleum hydrocarbons, *Environ. Sci. Technol.*, 34, 2000.
- Sudicky, E.A., A natural gradient experiment of solute transport in a sand aquifer: spatial variability of hydraulic conductivity and its role in the dispersion process, *Water Resour. Res.*, 22:2069–2082, 1986.
- Taylor, G., Dispersion of soluble matter in solvent flowing slowly through a tube, *Proc. R. Soc. London, A*, 219:186–203, 1953.
- van Genutchen, M.Th. and P.J. Wierenga, Mass transfer studies in porous media. 1. Analytical solutions, *Soil Sci. Soc. Am. J.*, 40:473–450, 1976.
- Zheng, C. and G.D. Bennett, *Applied Contaminant Transport Modeling*, Second Edition, Wiley, New York, 621 p., 2002.
- Zheng, C. and S.M. Gorelick, Analysis of the effect of decimeter-scale preferential flow paths on solute transport, *Ground Water*, 41:142–155, 2003.
- Zheng, C. and J.J. Jiao, Numerical simulation of tracer tests in a heterogeneous aquifer, *J. Environ. Eng.*, 124:510–516, 1998.
- Zheng, C. and P.P. Wang, MT3DMS: A modular three-dimensional multispecies model for simulation of advection, dispersion and chemical reactions of contaminants in groundwater systems; Documentation and User’s Guide, Contract Report SERDP-99-1, U.S. Army Engineer Research and Development Center, Vicksburg, MS, 1999.

Further Information

For a comprehensive treatment of the MADE-1 test, refer to a 4-part series of papers by Adams and Gelhar (1992), Boggs and Adams (1992), Boggs et al. (1992), and Rehfeldt et al. (1992). The information on

the MADE-2 test can be found in Boggs et al. (1993) and MacIntyre et al. (1993). The MADE-3 test is covered in Boggs et al. (1995), Libelo et al. (1997) and Stapleton et al. (2000). Zheng and Jiao (1998), Harvey and Gorelick (2000), Feehley et al. (2000), and Julian et al. (2000) provide an in-depth discussion of the difficulties in using the classical advection-dispersion model to simulate the MADE tracer tests and the applicability of the dual-domain mass transfer model as a practical alternative to the advection-dispersion model. Zheng and Gorelick (2003) and Liu et al. (2004) explore the concept of connected decimeter-scale high-conductivity networks as an explanation for the dual-domain mass transfer model applied to the MADE site.

27

Determining Sustainable Groundwater Development

27.1	Sustainability	27-1
27.2	Quantitatively Maintaining a Groundwater Development	27-2
27.3	Analytical Methods in Hydrogeology	27-2
27.4	The Water Budget	27-3
27.5	Aquifer Dynamics and Models	27-5
27.6	The Dynamics of a Basin and Range Aquifer	27-6
27.7	Paradise Valley	27-8
27.8	Conclusions	27-10
	References	27-10
	Further Reading	27-11

John D. Bredehoeft
The Hydrodynamics Group

27.1 Sustainability

The concept of sustainable development focuses on the concept of limiting resource use to levels that can be sustained over the long term. The Brundtland Commission (1987), known formerly as the World Commission on Environment and Development, defined sustainable development as that which meets the needs of the present without compromising the ability of future generations to meet their own needs. Implicit in this definition is a conservation ethic that states our generation should not use the resources of the planet, in the broadest context, in a manner that leaves our children and grandchildren (future generations) with problems of degraded, or impossibly depleted resources. The idea introduces a whole host of value judgments to be made by both individuals, and society collectively. History tells us that society's values with respect to the use of the planet's resources change with time. Society's judgment of what is sustainable today is different from that of older generations, and in turn, our children's and our grandchildren's societal values are likely to be different than ours.

The idea of this chapter is to introduce the concept of sustainability into groundwater development. Alley and Leake (2004) discuss how our ideas about groundwater developments that can be maintained indefinitely have evolved from the concept of safe yield to one of sustainability. Many things can degrade

the prolonged use of a groundwater system; the degradations generally fall in the categories of degraded groundwater quality, or a depletion of the quantity of water available from the resource, or a combination of both. Degradation is a serious issue that impacts groundwater sustainability. It is often important, especially in problems involving changes in groundwater quality to understand aquifer systems; this includes both the recharge and discharge from the systems. This is especially true where recharge brings with it undesirable water quality.

However, the focus in this chapter is on the quantitative issues associated with sustainability. Somehow the idea persists within the groundwater community that the sustained yield of an aquifer is some arbitrary fraction of the virgin recharge. Theis (1940) showed the fallacy of this notion, yet the idea persists. A second intent of this paper is to reinforce Theis' concept of how a groundwater system responds to development.

27.2 Quantitatively Maintaining a Groundwater Development

Much of our understanding of how groundwater systems respond to development goes back to the early work of Theis. Theis (1940) in a paper entitled: *The source of water derived from wells: essential factors controlling the response of an aquifer to development*, introduced the principle of capture. One can explain the principle in the following manner: before development a groundwater system is in a state of equilibrium in which the recharge is balanced by the discharge from the system. Pumping imposes an additional stress on the system. If the system is to reach a new equilibrium state, in which there is no further change in storage attributable to the pumping, then there must be an increase in the recharge, or a decrease in the discharge, or some combination of both.

The United States Geological Survey (Lohman, 1972) in *Definitions of selected groundwater terms* published the following definition of **capture**:

Water withdrawn artificially from an aquifer is derived from a decrease in storage in the aquifer, a reduction in the previous discharge from the aquifer, an increase in recharge, or a combination of these changes. The decrease in discharge plus the increase in recharge is termed capture.

Neither the (virgin) recharge nor discharge is of interest in determining the long-term productivity of an aquifer; rather it is the changes in these quantities, caused by the pumping, that one seeks. The purpose of this chapter is to show that the relevant quantity in determining the long-term productivity of a groundwater system is capture.

27.3 Analytical Methods in Hydrogeology

Before digital computer modeling codes, groundwater hydrologists used traditional analytical methods to assess the impacts of wells on groundwater systems. The traditional method of analysis used the principle of superposition (Reilly et al., 1987, see also Section 10.2.5). In this approach one assumes that the hydraulic head (or the water table) before development resulted from the inputs and outputs (recharge and discharge) from the system. One analyzes the impact of pumping independent of the initial (or virgin) hydraulic head. The cone of depression is calculated as a function of time. This cone of depression is then superposed upon the existing hydraulic head (or water table); the resulting head after superposition is the solution to the development.

What information does one need to make such a superposition calculation? One needs (1) the transmissivity and storativity distribution within the aquifer, (2) the boundary conditions that will be reached by the cone of depression, and (3) the rate of pumping. Those trained in classical hydraulic theory are well aware of reflection boundaries and image wells to account for the boundary conditions.

Missing from the classical analysis is any mention of recharge. The recharge is taken into account by the initial hydraulic head (or the water table). The initial head is a solution to initial boundary value problem that includes both the recharge and discharge.

Prior to the widespread use of digital computer models most analyses in groundwater flow were made using the principles of superposition. This was also the methodology used in the analog computer models of the 1950s, 1960s, and 1970s. With the advent of digital computer models it became feasible to specify the varying distributions of recharge and discharge with the idea of solving for the virgin water table. The calculated water table can then be compared to the observed water table (or hydraulic head). Such an analysis requires knowledge of the distribution of the virgin rate of both recharge and — in addition to the transmissivity distribution, and the boundary conditions for the aquifer of interest.

Usually one has an estimate of the rainfall; however, one has no idea of how large the recharge is except that it cannot exceed some unknown fraction of rainfall. One may know the transmissivity of the aquifer at a few places, and may also know that the aquifer discharge that makes up the baseflow of streams associated with the aquifer. Based upon this set of limited information a steady-state model analysis is made in an attempt to estimate the transmissivity of the aquifer. This is a common digital computer model analysis. In this context, knowledge of the virgin recharge and discharge are useful in estimating the transmissivity.

The recharge and the discharge are the inputs and outputs from a groundwater system. Both quantities are important in understanding how a particular groundwater system functions. However, it is not the purpose of this chapter to discuss recharge or discharge. The focus is in how recharge and discharge enter into the determination of the long-term quantitative yield of a groundwater system.

In the classical analytical method the important variables for determining the impacts of pumping are those that describe the dynamic response of the system — the distribution of aquifer diffusivity and the boundary conditions. This argument was the thrust of Brown's (1963) paper. The argument makes sense to one trained in classical analytical methods; it is obscure to others. Brown's paper made almost no impact on the profession. This chapter will attempt to further simplify the mathematical argument.

27.4 The Water Budget

To illustrate the basic principle, consider a simple aquifer system. A permeable alluvial aquifer underlies a circular island in a freshwater lake. The intent is to develop a well on the island. The island aquifer is shown schematically in Figure 27.1.

Before development, recharge from rainfall creates a water table. The recharge over the island is balanced by discharge from the permeable aquifer directly to the lake. One can write the following water balance for virgin conditions on our island:

$$R_0 = D_0 \quad \text{or} \quad R_0 - D_0 = 0$$

where R_0 is the virgin recharge (this is the recharge one generally refers to), and D_0 is the virgin discharge. A water table develops on the island in response to the distribution of recharge and discharge and the transmissivity of the alluvial aquifer.

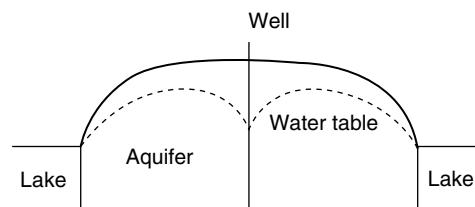


FIGURE 27.1 Cross-section of hypothetical island aquifer.

The discharge to the lake can be obtained at any point along the shore by applying Darcy's Law:

$$dD = T(dh/dl)$$

where D is the discharge through the aquifer at any point along the shore, T is the transmissivity at the same point, and dh/dl is the gradient in the water table at that point. If one integrates the point discharge along the entire shoreline of the island one obtains the total discharge from the island:

$$\int T(dh/dl) ds = D_0$$

Now at the middle of the island a well is installed and pumping is initiated (see Figure 27.1). At any time one can write a new water balance for the island:

$$(R_0 + \Delta R_0) - (D_0 + \Delta D_0) - P + dV/dt = 0$$

where ΔR_0 is the change in the virgin rate of recharge caused by the pumping, ΔD_0 is the change in the virgin rate of discharge caused by the pumping, P is the rate of pumping, and dV/dt is the rate at which pumping removes groundwater from storage on the island.

The virgin rate of recharge, R_0 , equals the virgin rate of discharge, D_0 , the water budget equation following the initiation of pumping reduces to

$$\Delta R_0 - \Delta D_0 - P + dV/dt = 0 \quad \text{or} \quad \Delta R_0 - \Delta D_0 - P = -dV/dt$$

For a long-term development one requires the rate of water taken from storage to be zero; in other words, we define prolonged development that can be maintained indefinitely as:

$$dV/dt = 0 \quad \text{as } t \text{ goes to } \infty$$

where dV/dt goes to zero at some future time. As long as the system can reach a future state where there is no continued depletion of storage we define the system as capable of being maintained indefinitely.

The water budget for long-term development (at $t = \infty$) is:

$$\Delta R_0 - \Delta D_0 = P$$

This states that in order to reach a development that can be maintained indefinitely the pumping must be balanced, at some future time, by a change in the virgin rate of recharge, ΔR_0 , or a change in the virgin rate of discharge, ΔD_0 , caused by the pumping. Traditionally, the sum of the change in recharge and the change in discharge caused by the pumping, the quantity $(\Delta R_0 - \Delta D_0)$, is defined as the *capture* attributable to the pumping. *For a development to be maintained indefinitely the rate of pumping must equal the rate of capture, at some future time.*

Notice that to determine a development that can be prolonged *we do not need to know the recharge*. The recharge may be of interest as are all the facets of the hydrologic budget, but it is not a determining factor in our analysis of quantitative sustainability.

Recharge is often a function of external conditions — rainfall, vegetation, soil permeability, and the like. In many, if not most, groundwater situations, the rate of recharge cannot be impacted by the pumping; in other words in terms of our water budget:

$$\Delta R_0 = 0$$

In most situations, prolonged development occurs when the pumping captures an equal amount of virgin discharge:

$$P = \Delta D_0$$

It is interesting to see how in the island aquifer, capture occurs conceptually. When pumping is initiated a cone of depression is created. Figure 27.1 (cross-section) shows the cone of depression at an early stage in the development of the island aquifer. The natural discharge from the island does not start to change until the cone of depression changes the slope in the water table at the shore of the island — remember Darcy's Law controls the discharge at the shoreline. Until the slope of the water table at the shoreline is changed by the pumping, the natural discharge continues at its virgin rate. Up until the point in time that the cone reaches the shore and changes the water table gradient significantly, all the water pumped from the well is supplied totally from storage in the aquifer. In other words, the cone of depression must reach the shoreline before the natural discharge is impacted. The rate at which the cone of depression develops, and reaches the shoreline, and changes the slope of the water table there, depends on the dynamics of the aquifer system — its transmissivity, storativity (or specific yield), and its boundary conditions. The rate of capture in a groundwater system is a problem in the dynamics of the system. Capture has nothing to do with the virgin rate of recharge.

At some point in time, if the pumping is large enough, the natural discharge will be almost eliminated — this occurs when the slope of the water table is flat everywhere at the shoreline. The island aquifer is an aquifer system in which one can induce water to flow from the lake into the aquifer (Figure 27.1). In this instance a prolonged development can exceed the virgin recharge (or the virgin discharge). This again suggests that the recharge is not a relevant input in determining the magnitude of a development that can be maintained indefinitely if another source of water is available — a lake, or stream.

Often the geometry of the aquifer restricts the capture. For example, were the aquifer on the island to be thin, one might run out of water at the pump long before the pumping could capture any fraction of the discharge. In this case, all the water pumped would come from storage — it would be “mined.” In the island example, with a thin aquifer, the well could run dry before it could impact the discharge at the shoreline. This again points out that the dynamic response of the aquifer system is all important to determining the impacts of development. It is for these reasons that hydrogeologists are concerned with the dynamics of aquifer system response. Hydrogeologists model aquifers in an attempt to understand their dynamics.

Clearly, the circular island aquifer example is a simple system. Even so, the principles explained in terms of this simple aquifer apply to all groundwater systems. It is the dynamics of how capture takes place in an aquifer that ultimately determines how large a prolonged groundwater development can be maintained.

Whether a groundwater system that can be maintained indefinitely is sustainable is an entirely different question having to do with the social, political, and ethical consequences of the development. Creating a groundwater system in which the pumping is balanced by the capture, a system that can be maintained indefinitely, may be a necessary condition for sustainability, but by itself it is not a sufficient condition. In an earlier paper (Bredehoeft, 2002) I argued that a system balanced at some future time by capture was both necessary and sufficient for sustainability; however, I no longer believe it to be sufficient — I stand corrected by my colleagues.

27.5 Aquifer Dynamics and Models

Since the development of the Theis equation in 1935 hydrogeologists have been concerned with the dynamics of aquifer response to stress — pumping or recharge. Once Theis (1935) and later Jacob (1940) showed the analogy of groundwater flow to heat flow, the groundwater community has been busy solving the appropriate boundary value problems that describe various schemes of development. This endeavor has gone through several stages.

The 1940s and 1950s were a time during which the groundwater profession was concerned with solving the problems of flow to single well. Numerous analytical solutions to the single well problem were produced. These solutions were used both to predict the response of the aquifer system, as well as to estimate aquifer properties — transmissivity (or permeability) and the storativity.

Hydrogeologists of that day saw the limitations in analyzing wells and sought a more robust methodology by which to analyze an entire aquifer, including complex boundary conditions and aquifer heterogeneity. The search led a group at the USGS to invent the analog model in the 1950s; the genius behind this development was Herb Skibitski — Herb was one of those individuals who rarely published. The new tool was the electric analog computer model of the aquifer. In the analogous system, voltage is equivalent to hydraulic head; electrical current is equivalent to the flow of water (see Section 3.4.1). The model consisted of a finite-difference network of resistors and capacitors. In the analog computer aquifer transmissivity is represented by the network of resistors; the storativity is represented by the network of capacitors. The resulting resistor–capacitor network is excited by electrical function generators that simulated pumping, or other stresses.

In reality, the analog models were elegant finite-difference computer models of aquifer systems. By 1960, the USGS had a facility in Phoenix, Arizona where analog models of aquifers were routinely built on a production basis. At about the same time, The Illinois Water Survey established an analog model facility in Urbana, Illinois. Some of these analog models had multiple aquifers; some had as many as 100,000 nodes. At the time, it was infeasible to solve the same problems with digital computers; the digital computers of the day were too small and too slow. However, by 1970, the power of the digital computers increased to the point that digital aquifer models could begin to compete with the analog models. By 1980, digital computer models had replaced the analog models, even at the USGS. The models of the 1980s have now grown to include solute transport, pre- and post-processors, and automatic parameter estimation. By far the vast majority of groundwater flow problems are simulated using the USGS code MODFLOW; there is a new version MODFLOW 2000 (see Chapter 23).

The groundwater model is a tool with which to investigate the dynamics of realistic aquifer systems. As suggested above, it is only through the study and understanding of aquifer dynamics that one can determine the impact of an imposed stress on an aquifer system.

27.6 The Dynamics of a Basin and Range Aquifer

One can illustrate the dynamic response of aquifers using as the prototype a closed basin aquifer like those in the Basin and Range of Nevada. The aquifer geometry is illustrated in plan view in Figure 27.2. The basin is approximately 50 miles in length by 25 miles in width. At the upper end of the valley are two streams that emerge from the nearby mountains and recharge the aquifer at an average combined rate of 100 cfs ($\sim 70,000$ acre-feet annually). At the lower end of the valley an area of phreatophyte vegetation discharges groundwater as evapotranspiration (ET) at an average rate of 100 cfs. The system before development is in balance; 100 cfs is being recharged, and 100 cfs is being discharged by ET.

Two scenarios of groundwater development are of interest in this aquifer. The size of both developments is the same, and is equal to the recharge (and the discharge) 100 cfs. Two locations for the well field are considered — shown as Case I and Case II in Figure 27.2. The Case II well field is closer to the area of phreatophyte vegetation. To simulate the system one needs aquifer properties; the aquifer properties are specified in Table 27.1.

In the hypothetical system, phreatophyte groundwater consumption is eliminated as the pumping lowers the water table in the area containing phreatophytes. This is a groundwater system in which capture of ET can occur. A linear function is used to cutoff the phreatophyte consumption. As the water table drops between 1 and 5 ft, the phreatophyte use of groundwater is linearly reduced. The reduction in phreatophyte use does not start until the groundwater declines 1 ft, by the time the water table drops 5 ft the phreatophyte use is eliminated in that cell. The phreatophyte-reduction function is applied cell by cell in the model.

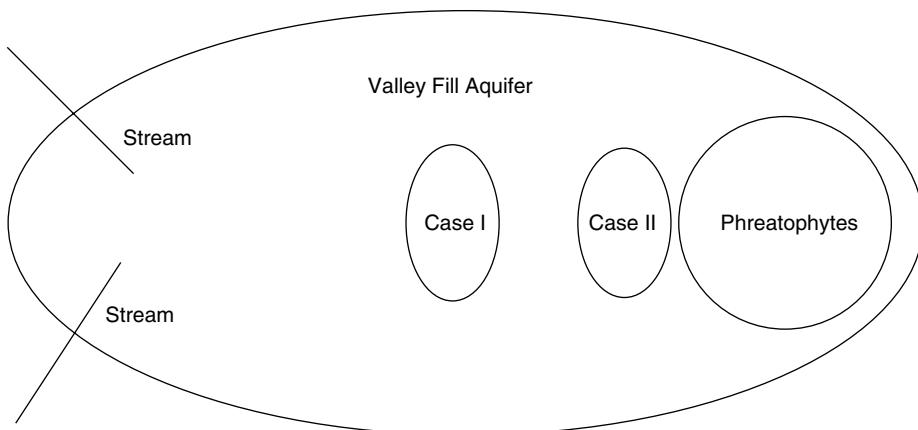


FIGURE 27.2 Plan of hypothetical valley fill aquifer. The two streams, to the left, provide on average 100 cfs of recharge to the aquifer. The area of phreatophytes, to the right, discharges on average 100 cfs of groundwater through ET before groundwater development. We consider two scenarios of groundwater development located as Case I and Case II; each development pumps at a rate equal to the recharge — 100 cfs.

TABLE 27.1 Aquifer Properties

Basin size	50 × 25 miles (see Figure 27.2)
Model cell dimensions	1 × 1 mile
Hydraulic conductivity	0.00025 ft/sec
Saturated thickness	2000 ft
Transmissivity	0.5 ft ² /sec (~43,000 ft ² /day)
Storage coefficient	0.1–10% specific yield
Phreatophyte consumption	100 cfs
Well field pumping	100 cfs
Recharge	100 cfs

For this system to reach a new equilibrium state, in which the yield can be maintained indefinitely, the phreatophyte consumption must be eliminated entirely. Using the model we can examine the phreatophyte use as a function of time. Figure 27.3 is a plot of the phreatophyte use in the system vs. time since pumping was initiated.

The ET is reduced faster where the pumping is closer to the phreatophytes — Case II. The point of considering Case I and Case II is to show that the location of the pumping makes a difference in the dynamic response of the system. In Case II, where the pumping is close to the phreatophytes, the ET is reduced by approximately 40 cfs in 10 years. In contrast in Case I, the ET is reduced by only approximately 5 cfs in 10 years. Both scenarios, Case I and Case II, take a long time to fully eliminate the ET; it takes almost 1000 years before the ET is totally eliminated. Even seasoned hydrologists are surprised at how long it takes a water table groundwater system to reach a new equilibrium state in which no more water is removed from storage.

One can examine the output from the model in another way by investigating the total amount of water removed from storage in the hypothetical valley fill aquifer (Figure 27.4). Both scenarios of development, Case I and Case II, take approximately the same time to reach a new equilibrium state. In Case II, where the pumping is close to the phreatophytes, the amount of water removed from storage is approximately 60% less than in Case I.

It is important to notice, that even though the two developments (Case I and Case II) are equal in size, the aquifer responds differently depending upon where the developments are sited. This again emphasizes the importance of studying the dynamics of the aquifer response — the response is different depending on where the development is located.

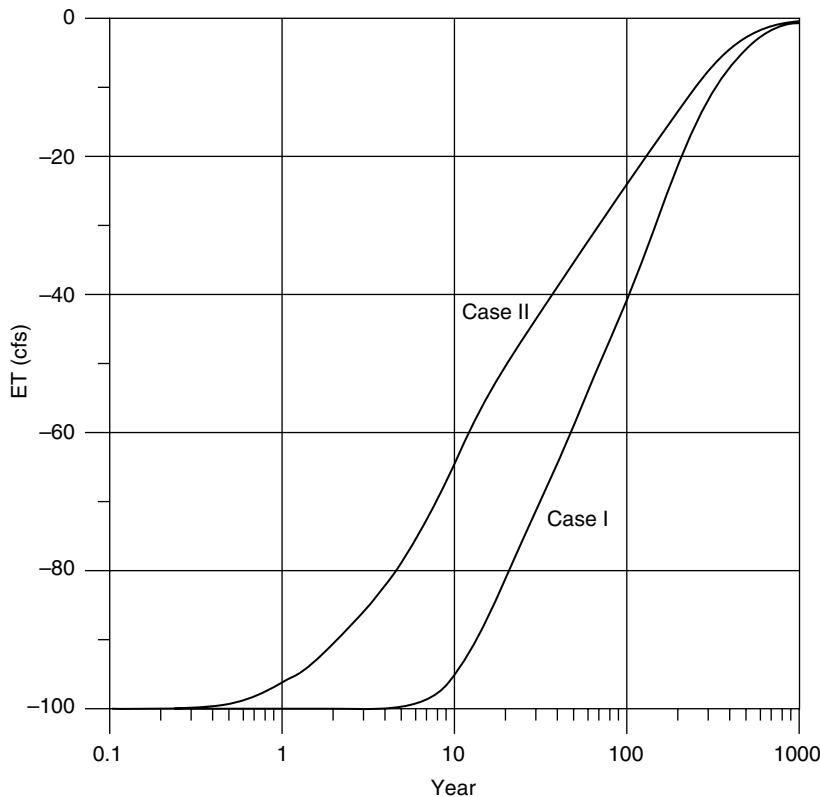


FIGURE 27.3 ET vs. time in our hypothetical valley-fill aquifer.

This example, of an overly simplified Basin and Range valley fill aquifer, illustrates the importance of understanding the dynamics of aquifer systems. Again, while this is a simple example, the principles illustrated apply to aquifers everywhere. It is the rate at which the phreatophyte consumptive use can be captured that determines how this system can be maintained indefinitely; this is a dynamic process. Capture always entails the dynamics of the aquifer system.

The question arises — is the basin and range aquifer described in this simple example sustainable? This depends upon one's values. To reach the new equilibrium state, in which the pumping can be maintained indefinitely, one has to eliminate all the phreatophyte consumptive use — kill the phreatophytes. This vegetation provides cover for wildlife in the area — is it reasonable to kill this vegetation? Nevada law restricts groundwater pumping to the natural recharge. Implicit in the law is a value judgment that discounts the value of phreatophyte vegetation that pumping to the full amount of the recharge will usually eliminate.

27.7 Paradise Valley

Paradise Valley is a real development in Nevada. Alley and Leake (2004) explored the concept of sustainability; they used as their example a development in Paradise Valley in northern Nevada. The Humboldt River flows across the southern end of the valley. They simulated groundwater pumping near the southern end of the valley, not far to the north of the Humboldt River. Alley and Leake (2004) examined the source of the groundwater pumped vs. time. There are four sources of water that support the pumping (1) water from storage, (2) capture of ET, (3) capture of water leaving the valley, and (4) induced recharge

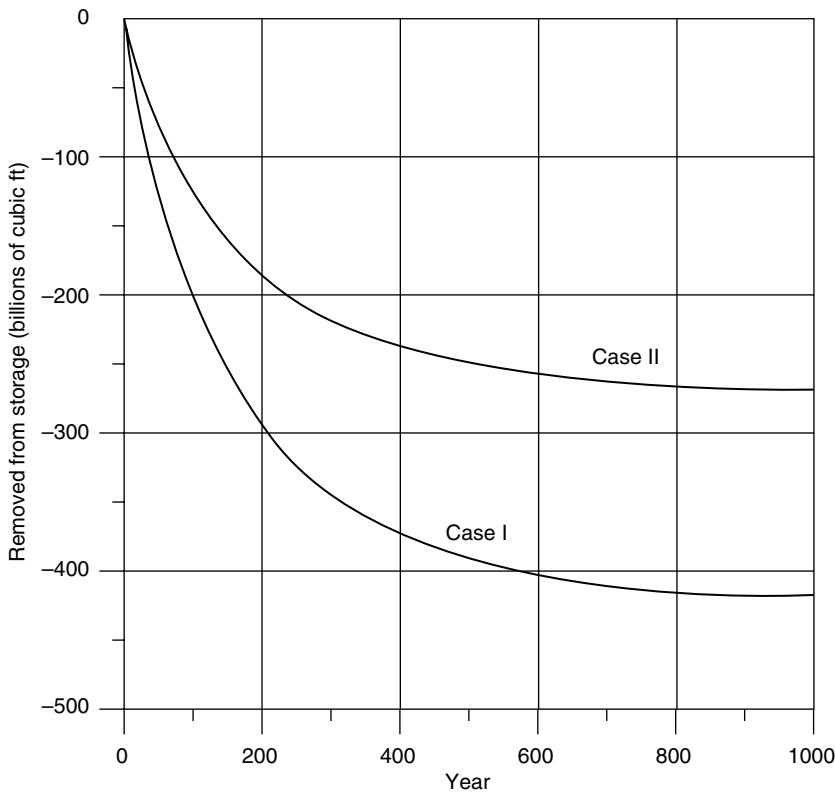


FIGURE 27.4 The volume of water removed from storage as a function of time in the hypothetical valley fill aquifer with two developments — Case I and Case II.

from the Humboldt River. Each of these sources varies with time. Figure 27.5 is a plot of the four sources of water that supply the pumping vs. time.

One sees the principal source of groundwater in Paradise Valley, during the early period, is from storage in the system. The storage declines to only 3% of the supply in 300 years. The capture of water from phreatophyte ET grows from 20% in 1 year, to approximately 74% of the total in 300 years. The induced recharge from the Humboldt River grows from zero in the early years to approximately 19% of the total in 300 years. The capture of outflow from the valley grows to 3% in 300 years.

The groundwater system in Paradise Valley takes a long time to reach a new equilibrium state. The time is similar to the valley fill aquifer explored above. Even after 300 years 3% of the water pumped is still coming from storage. The analysis suggests that the Paradise Valley pumping can be maintained indefinitely, that is, the pumping will at some point after 300 years be balanced by the capture.

Nowhere in the analysis by Alley and Leake (2004) does the recharge enter into the calculation of sustainability. The important quantities are the removal of groundwater from storage, the capture of ET, and capture (induced recharge) from the Humboldt River.

The groundwater system described by Alley and Leake (2004) can be maintained indefinitely. However, it reaches a new steady state by depleting streamflow in the Humboldt River; this is streamflow that is already claimed, and to which the downstream users have water rights. Any reasonable application of sustainability suggests that the Paradise Valley groundwater development is not sustainable; it infringes on the prior rights of others. Again, such a conclusion depends upon one's perspective. The Paradise Valley farmers are pumping groundwater to irrigate potatoes, while the downstream users are using the streamflow of the Humboldt River to irrigate alfalfa. Were the upstream users pumping to supply the nearby town of Winnemucca, the value judgment might not be so clear.

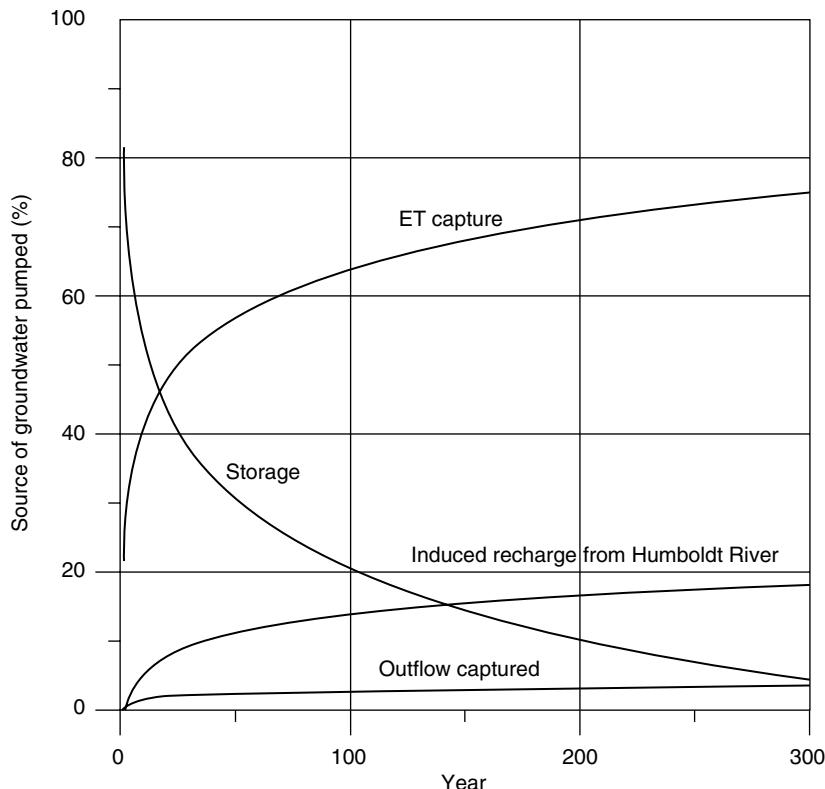


FIGURE 27.5 Sources of groundwater that supply the pumping in Paradise Valley, Nevada. (From Alley, W.M. and Leake, S.A., *Ground Water*, 42, 12–16, 2004.)

27.8 Conclusions

The important quantity in determining how a groundwater system reaches a new equilibrium is capture. Capture in an aquifer system is a dynamic process. For this reason hydrologists are occupied in studying aquifer dynamics. The principal tool for these investigations today is the groundwater model. These ideas are not new; Theis (1940) articulated this basic principle. Nowhere in the quantitative determination of long-term maintenance of pumping is the virgin groundwater recharge relevant.

Sustainability, on the other hand, is a much broader question that involves the application of societal values. In the final analysis, hydrogeologists seek to be good stewards of that small part of the earth for which they bear some management responsibility. Hopefully, one can apply the ethics of sustainability wisely.

References

- Alley, W.M. and Leake, S.A., The journey from safe yield to sustainability. *Ground Water*, 42, 12–16, 2004.
- Bredehoeft, J.D., The water budget myth revisited: why hydrogeologists model. *Ground Water*, 49, 340–345, 2002.
- Brown, R.H., The cone of depression and the area of diversion around a discharging well in an infinite strip aquifer subject to uniform recharge. U.S. Geological Survey Water-Supply Paper 1545C: pp. C69–C85, 1963.
- Brundtland (Report), *Our common future*, World Commission on Environment and Development. New York, Oxford University Press, 1987.

- Jacob, C.E., On the flow of water in an elastic artesian aquifer. *Transactions of American Geophysical Union, Part 2*: pp. 585–586, 1940.
- Lohman, S.W., Definitions of selected ground-water terms — revisions and conceptual refinements, *Water Supply Paper* 1988, U.S. Geological Survey, Reston, VA, 1972.
- Reilly, T.E., Franke, O.L., and Bennett, G.D., Principle of superposition and its application in ground-water hydraulics, *Techniques of Water-Resources Investigations, Book 3: Application of Hydraulics*, Chapter B6. U.S. Geological Survey, Denver, CO, 1987.
- Theis, C.V., The relation between lowering the piezometric surface and the rate and duration of discharge of a well using groundwater storage. *Transactions American Geophysical Union, 16th Annual Meeting, Part 2*: pp. 519–524, 1935.
- Theis, C.V., The source of water derived from wells: essential factors controlling the response of an aquifer to development. *Civil Engineer*, 10, 277–280, 1940.

Further Reading

- Alley, W.M. and Leake, S.A., The journey from safe yield to sustainability. *Ground Water*, 42, 12–16, 2004.
This paper explores sustainability using as an example a groundwater development in Paradise Valley, Nevada.
- Devlin, J.F. and Sophocleous, M., The persistence of the water budget myth and its relationship to sustainability. *Hydrogeology Journal*, 13, 549–554, 2005.
This paper further explores the idea of capture and its relationship to sustainability.

28

The Impact of Climate Change on Groundwater

28.1	Introduction.....	28-1
28.2	Background	28-2
	Climate Change • Impact of Climate Change on Water Resources • Groundwater Recharge • Factors Affecting Recharge • Recharge Estimation Methods • Recharge in Groundwater Modeling	
28.3	A Case Study: Recharge Analysis of the Grand River Watershed	28-20
	Introduction • Input Data • Methodology • Results • Conclusions	
28.4	Chapter Summary	28-34
	References	28-34

Mikko I. Jyrkama and
Jon F. Sykes
University of Waterloo

28.1 Introduction

The growing and competing demand for domestic, agricultural, industrial, and recreational water has made clean and safe drinking water a scarce natural resource in many regions of the world. Increasing populations, especially in urban areas, are not only stressing the capacity and sustainability of the existing water supplies, but they are also placing the supplies at a greater risk of contamination. Humans are also affecting the global water resources through climate change. Increases in the atmospheric concentrations of greenhouse gases and aerosols, as well as changes in land use due to deforestation and urbanization already have, and will continue to impact the Earth's climate, and hence the terrestrial hydrologic cycle in the future (IPCC, 2001).

The human impacts on Earth's water resources are not only being felt in developing countries, but they are increasingly, and often dramatically, being noticed in developed nations such as the United States and Canada. Understanding the interaction between the various components of the hydrologic cycle is, therefore, essential for the management of watershed and for ensuring the quality and sustainability of safe and clean drinking water resources.

The surface water and groundwater systems have often been considered as separate due to the pervasive complexity and increased computational effort required by their integration (El-Kadi, 1989). The systems, however, are intrinsically linked through the processes of recharge and discharge, and their interaction constitutes an essential part of the hydrologic cycle.

According to Robins (1998), “the management of both groundwater resources and of individual groundwater sources cannot sensibly be undertaken without some knowledge of recharge: its quantity, its seasonality and, above all, the different routes through the sub-soil and the unsaturated zone by which it can occur”. Recharge also plays an important role in the assessment of aquifer vulnerability to contamination. According to Foster (1998) “since the transport of most groundwater contaminants, with the exception of density driven contaminants such as DNAPLs, to saturated aquifers occurs in the aqueous phase as part of the recharge process, assessing aquifer pollution vulnerability is inextricably linked with understanding groundwater recharge mechanisms.” Aquifer vulnerability maps can provide valuable guidance to municipal planners and developers regarding land use allocation, well locations, and pumping regulations. This is especially important in areas where the underlying aquifers are exploited extensively for drinking water purposes.

Changes in future climate will alter regional hydrologic cycles and subsequently impact the quantity and quality of regional water resources (Gleick, 1989). While climate change affects surface water resources directly through changes in the major long-term climate variables such as air temperature, precipitation, and evapotranspiration, the relationship between the changing climate variables and groundwater is more complicated and poorly understood. Groundwater resources are related to climate change through the direct interaction with surface water resources, such as lakes and rivers, and indirectly through the recharge process. Therefore, quantifying the impact of climate change on groundwater resources requires not only reliable forecasting of changes in the major climatic variables, but also accurate estimation of groundwater recharge.

Accurate spatial and temporal characterization of groundwater recharge can be difficult, due to its dependence on a multitude of physical factors such as land use and hydrogeological heterogeneity (Lerner et al., 1990). Groundwater recharge can also be significantly impacted by snowmelt and frozen soil conditions in northern climates, and furthermore be complicated by the issue of scale.

It is evident that assessing the impact of climate change on groundwater resources requires a physically based approach for estimating groundwater recharge. The method must not only account for temporal variations in the climatic variables and their impact on the hydrologic cycle, but also consider the spatial variation of surface and subsurface properties across the study area.

Many climate change studies have generally focused on modeling the temporal change in the hydrologic processes and ignored or relied on average spatial characteristics due to model limitations or coarse discretization schemes. This chapter outlines a physically based methodology that can be used to characterize not only the temporal impact, but also the spatial effect of climate change on groundwater recharge. The method is based on the hydrologic software package HELP3 coupled with a geographic information system (GIS). The method is used in an example application to simulate the past conditions and possible future changes in the hydrologic cycle of the Grand River watershed in Ontario, Canada.

28.2 Background

28.2.1 Climate Change

The Earth’s energy cycle is driven by short wave radiation from the sun and balanced by the outgoing long wave terrestrial radiation. Any changes in this radiative balance will alter the global hydrologic cycle, and hence, the groundwater recharge as well as the atmospheric and oceanic circulation, thereby resulting in climate change. According to the Intergovernmental Panel on Climate Change (IPCC), the key climate variables impacted by changes in the Earth’s energy balance include temperature, precipitation and atmospheric moisture, snow cover, extent of land and sea ice, sea level, patterns in atmospheric and oceanic circulation, and extreme weather and climate events (IPCC, 2001).

The Earth’s radiative balance can be affected not only by human impacts, but also by a multitude of natural factors, resulting in both cooling and warming of the climate system. While increases in greenhouse gases, such as carbon dioxide, ozone, methane, and nitrous oxide, reduce the outgoing terrestrial radiation (i.e., heat) to space, thereby resulting in warming of the lower atmosphere and the Earth’s surface,

anthropogenic aerosols, such as those derived from fossil fuels and biomass burning, can reflect the incoming solar radiation, leading to cooling of the climate systems (IPCC, 2001). Changes in volcanic activity, the energy output of the sun, and slow variations in the Earth's orbit, not only have, but will also continue to impact the Earth's climate in the future. Climate variations can also occur in the absence of external factors, as a result of complex interactions between the components of the climate system, such as the coupling between the atmosphere and the oceans.

Variations in climate occur across a range of temporal and spatial scales, making both observation and modeling highly uncertain (Goddard et al., 2001). Nevertheless, advanced remote sensing from Earth-observation satellites is providing increasingly detailed images of our current climate, while the growing amount of paleoclimatic core data from trees, sediments, and ice are providing information about the Earth's climate in the distant past. At the same time, the continuing increase in computing power has led to the development of more detailed and sophisticated global climate and general circulation models (GCMs). The interaction between the Earth's orbit, ocean circulation, and atmosphere has contributed to nine glacial episodes in the northern hemisphere over approximately the past 900,000 years. Tarasov and Peltier (2004) integrated a GCM with a glacial model to simulate the evolution of ice coverage over the North American continent. McGuffie and Henderson-Sellers (2001) provide a review of climate modeling over the past 40 years. Significant improvements in hydrological parameterizations and grid resolution have also allowed coupling between the atmospheric circulation and hydrologic surface water models (Stewart et al., 1998; Pitman, 2003).

The greatest uncertainty in future predictions by climate models arises from clouds and their interactions with radiation. Clouds can both absorb and reflect solar radiation and absorb and emit long wave radiation, thereby either cooling or warming the Earth's surface. The variation in the two states is controlled by many factors and is difficult to model. Therefore, clouds represent a significant source of potential error in climate simulations (IPCC, 2001).

Through the co-ordination of researchers from around the world, and after a comprehensive and careful review of a vast number of observational records and results from a number of simple and complex GCM models, the IPCC has arrived at the following general predictions for the global climate during this century (IPCC, 2001):

- The global average surface air temperature is predicted to increase by 1.5 to 4.5°C. Furthermore, it is very likely that nearly all land areas will warm more rapidly than the global average, particularly those at northern high latitudes in the cold season.
- Globally averaged water vapor, evaporation, and precipitation are projected to increase. Both increases and decreases in precipitation are seen at the regional scale.
- Precipitation extremes are projected to increase more than the mean, and the intensity and frequency of extreme precipitation events are projected to increase.
- Ice caps and glaciers will continue their widespread retreat during the 21st century while sea ice and snow cover in the Northern Hemisphere are projected to decrease further.
- The global average sea level is predicted to increase by 0.09 to 0.88 m, with regional variations.

28.2.2 Impact of Climate Change on Water Resources

Various hydrologic models have been used to study the impact of climate change on water resources (e.g., Wilkinson and Cooper, 1993; Gureghian et al., 1994; Cooper et al., 1995; Bobba et al., 1997; Querner, 1997; Rosenberg et al., 1999; Kirshen, 2002). The hydrological effects of climate change are commonly evaluated by estimating the sensitivity of model outputs, such as streamflow hydrographs or soil moisture contents, to hypothetical changes in the magnitude and temporal distribution of model inputs such as precipitation and temperature. In addition to hypothetical perturbations, however, the results inferred from the GCMs have also been used to predict the effects of climate change on regional hydrology.

In one of the early studies, Vaccaro (1992) used a deep percolation model (DPM) to estimate the influence of climate change on recharge variability in a basin in north-western United States. In addition

to historical records, climate predictions from the synthetic weather generator WGEN (Richardson and Wright, 1984) and three GCMs were considered along with two different land use conditions. The results of the study indicated that the variability in annual recharge was less under the GCM conditions than using the historic data. The objective of the study was to investigate the degree of variability in climate and its impact on future recharge predictions, and not explicitly to predict the impact of climate change on recharge rates in the basin.

Bouraoui et al. (1999) developed a methodology to disaggregate the outputs of large-scale GCMs for use in hydrologic models. The method was used in conjunction with the runoff and sediment transport model ANSWERS to investigate the impact of doubling CO₂ on groundwater recharge in a watershed in France. The methodology was based on a statistical approach requiring a large number of simulations. The results of the study indicated that recharge would decrease in the basin due to increase in atmospheric CO₂.

Rosenberg et al. (1999) used the GIS-based program HUMUS and the hydrologic model SWAT (Arnold et al., 1998) to study the impact of climate change on the water yield and groundwater recharge of the Ogallala aquifer in the central United States. Three different GCMs were used to predict changes in the future climate due to anticipated changes in temperature and CO₂ concentrations. The study found that recharge was reduced under all scenarios, ranging up to 77%, depending on the simulation conditions.

Kirshen (2002) used MODFLOW to study the impact of global warming on a highly permeable aquifer in the northeastern United States. The lumped hydrologic model HSPF was used to compute streamflows from rainfall, temperature, and potential evapotranspiration, while groundwater recharge was estimated using a separate model based on precipitation and potential evapotranspiration. The sensitivity of the aquifer system to the possible effects of global climate change was simulated by introducing both hypothetical and GCM predicted changes to the input parameters of the HSPF and recharge models, and then running MODFLOW based on the outputs. The results of the study were mixed, showing higher, no different, and significantly lower recharge rates and groundwater elevations, depending on the climate scenario used.

Croley and Luukkonen (2003) studied the impact of climate change on groundwater using a regional scale MODFLOW model of Lansing, Michigan. The results from two separate GCMs were first used in an empirical streamflow model to estimate baseflow rates. The recharge in the regional groundwater model was then adjusted on the basis of the differences in baseflow from the two GCMs. The study used a steady-state approach and hence the transient changes in groundwater levels were not captured in the analysis. The results of the study indicated that the simulated groundwater levels were generally predicted to increase or decrease due to climate change, depending on the GCM used.

Eckhardt and Ulbrich (2003) used a revised version of the SWAT model to investigate the impact of climate change on groundwater recharge and streamflow. The model was applied to a small catchment in Germany. Climate change was simulated by adjusting stomatal conductance and leaf area in the SWAT model as a response to increased CO₂ levels. Four different climate scenarios were considered based on simulations from five different GCMs. The results of the study indicated that more precipitation will fall as rain in winter due to increased temperatures resulting in higher recharge and streamflow in January and February. They also found that the snowmelt increase in recharge in March disappears, while recharge and streamflow were shown to be potentially reduced in the summer months.

Loaiciga (2003) conducted a review of climate change predictions and associated hydrologic consequences and presented the results of a case study of an aquifer in south-central Texas. The study considered a confined karst aquifer that receives recharge only in sections of streams that are hydraulically connected to the underlying water table. The impact of climate change on the indirect recharge was estimated using runoff scaling factors based on the ratio of historical and future streamflows predicted from linked general and regional climate models. The study also considered the impact of changes in pumping rates (i.e., predicted changes in groundwater use) on groundwater resources of the aquifer. The study concluded that the rise in groundwater use associated with predicted growth would pose a higher threat to the aquifer than climate change.

Allen et al. (2004) used Visual MODFLOW to study the impact of climate change in an aquifer in southern British Columbia, Canada. The recharge boundary condition for the groundwater model was

estimated using Visual HELP, while rivers were represented using constant heads. Constant recharge was assumed across the domain in a steady-state analysis. The base case recharge analysis was based on synthetic weather generated with WGEN (Richardson and Wright, 1984) using Canadian weather normals (i.e., thirty-year averages) from a nearby weather station as input. Four climate change scenarios were considered based on general predictions for the study area. The average temperature and precipitation were perturbed based on the predictions and were then used in WGEN to generate synthetic records. Average annual recharge rates were then computed for each of the scenarios and applied across the whole domain. Results of the study indicated that there was little change in the overall water table configuration as a result of potential climate change. This is to be expected, however, as constant recharge (i.e., spatially) was assumed across the domain.

Brouyere et al. (2004) investigated the impact of climate change on a chalky groundwater basin in Belgium using the integrated hydrological model MOHISE. The integrated model was composed of three submodels, each for soil, surface water, and groundwater, which were linked through flux boundaries. Three GCMs were used to compute monthly change rates in precipitation and temperature that were then used to adjust the daily historical record to obtain climate change scenarios. The results showed decreases in groundwater levels for two scenarios, while one of the models showed no significant change. Instead of seasonal changes in groundwater levels, a monotonic increase over time was observed. The impact of seasonal variation was likely smoothed out by a thick saturated zone found at their basin.

Understanding the impact of climate change is most crucial for studies concerned with the storage and containment of high-level nuclear waste. Due to the slow decay of nuclear materials, modeling efforts must be concerned with changes in boundary conditions over very long periods of time. However, predictions of climate change are highly uncertain even for shorter periods such as the next century, making the predictions for the next 10,000 to 100,000 years extremely uncertain.

Yucca Mountain in Nevada is being evaluated as a potential site for high-level nuclear waste disposal in the United States. Because the proposed repository is to be located within the deep unsaturated zone, estimating rates of groundwater recharge at the site is critically important (Flint et al., 2002). Gureghian et al. (1994) studied the impact of climate change on the groundwater recharge rate at Yucca Mountain using a quasi-linear form of the Richards' equation. They used two different climatic variation models for temperature and precipitation over the next 10,000 years based on the recommendations by a panel of experts. The results of the study indicate minimal differences between the two climate models on the overall movement of the wetting front.

In summary, climate change will have an impact on future recharge rates and hence on the underlying groundwater resources. The impact may not necessarily be a negative one, as evidenced by some of the investigations. Quantifying the impact is difficult, however, and is subject to uncertainties present in the future climate predictions. Simulations based on general circulation models have yielded mixed and conflicting results, raising questions about their reliability in predicting future hydrologic conditions.

Groundwater recharge is influenced not only by hydrologic processes, but also by the physical characteristics of the land surface and soil profile. Many climate change studies have focused on modeling the temporal changes in the hydrologic processes and ignored the spatial variability of physical properties across the study area. While knowing the average change in recharge and groundwater levels over time is important, these changes will not occur equally over a regional catchment or watershed. Long-term water resource planning requires both spatial and temporal information on groundwater recharge to properly manage not only water use and exploitation, but also land use allocation and development. Studies concerned with climate change should therefore also consider the spatial change in groundwater recharge as a result of future changes in hydrologic processes.

28.2.3 Groundwater Recharge

The word "recharge" has several meanings in the groundwater literature. Generally, it means any water that is added as an input to the system, while discharge is normally considered to be any water that exits the groundwater system. Within the context of this chapter, groundwater recharge is defined as the soil

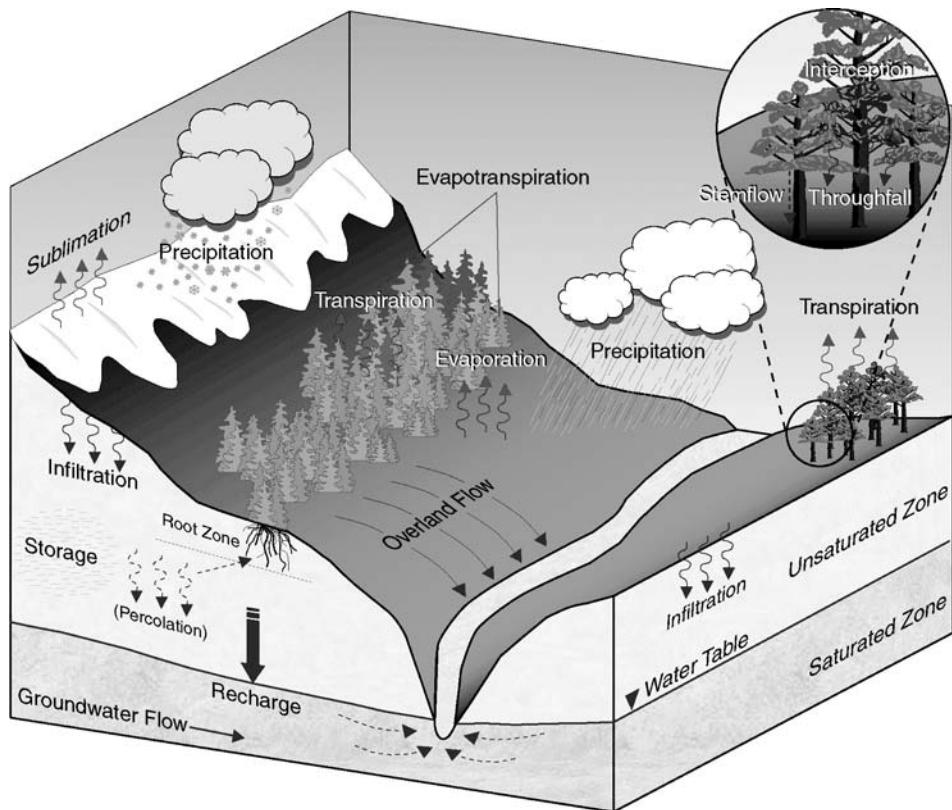


FIGURE 28.1 The hydrologic cycle.

water (from precipitation) in the vadose zone that crosses the water table into the saturated zone. Recharge of deeper confined aquifers can also occur both laterally and vertically through aquitard leakage in the saturated zone (Gerber and Howard, 2000).

Groundwater recharge is part of the vadose zone soil water budget. Figure 28.1 illustrates the main processes. For clarification, infiltration is defined as the volume rate of water flowing into a unit area of soil surface, whereas percolation is the process by which water migrates down through the soil profile in the unsaturated zone. Deep soil water percolation, often referred to as groundwater recharge by soil scientists, is the water that has moved past the evaporative and root zones in the vadose zone and is no longer available to plants.

The driving force for natural recharge is precipitation; however, groundwater recharge may also result from other processes. According to Lerner et al. (1990) groundwater recharge can be divided into three distinct processes:

- *direct recharge*, which is water added to the groundwater system in excess of soil moisture deficits and evapotranspiration by direct vertical percolation of precipitation through the unsaturated zone.
- *localized recharge* is an intermediate form of recharge resulting from the horizontal surface concentration of water in local joints and depressions.
- *indirect recharge* is the flow of water to the water table through the beds of surface water courses, such as rivers and lakes.

As direct recharge due to precipitation forms the highest contribution to the groundwater system in humid climates (Knutsson, 1988), it will receive the main focus in this section. Surface water bodies commonly act

as discharge areas rather than recharge areas in humid climates and therefore provide a minor contribution to the overall production of groundwater recharge. However, for areas of an aquifer in close proximity to a surface water body, seasonal variations in surface water levels may result in alternating periods of groundwater recharge and discharge. Climate change can impact the levels of the surface water and hence the relative importance of the recharge and discharge.

As shown in the hydrologic cycle in Figure 28.1, precipitation is formed either as rain or snow. Before reaching the ground, however, the vegetation canopy intercepts some of the precipitation, which then either evaporates, or is channeled to the ground via stemflow, or drips directly to the ground as part of throughfall. Depending on the rainfall intensity and ground surface cover, the precipitation reaching the ground surface may then flow overland directly into streams and ditches, or infiltrate into the soil. The infiltrated water then percolates downwards through the vegetative root zone or evaporative zone where a portion of it may be taken up by the plant roots and subsequently transpired through the vegetation canopy. The remaining water will continue percolating deeper into the soil column, eventually becoming groundwater recharge when crossing the water table into the saturated zone. The percolating soil water can also partition into the air phase in the vadose zone through evaporation at any time, depending on the moisture content and distribution in the unsaturated zone.

28.2.4 Factors Affecting Recharge

28.2.4.1 Precipitation

Precipitation is the most important parameter in the groundwater recharge process. It is the driving force in the hydrologic cycle and provides the water that will eventually recharge the groundwater system. Precipitation is effected by climatic factors such as wind and temperature, resulting in a very complex and dynamic distribution. Therefore, determining the amount and rate of precipitation over an area is exceedingly difficult due to the high spatial and temporal variation. Storm events can vary widely in duration, velocity, and intensity, while lower temperatures can result in snowfall or mixed precipitation. Singh (1997) provides a literature review on the effects of rainfall variability across a watershed on streamflow hydrographs. He concludes that the velocity and direction of a storm has a significant impact on the hydrographs and that temporally varying rainfall leads to higher peak flows than constant rainfall. In addition, low intensity rainfall may cause no recharge due to a high rate of evapotranspiration, whereas the same amount in a shorter time period may be sufficient to saturate the soil and cause recharge (Lerner et al., 1990). In arid and semiarid areas, where evaporation dominates the water balance, recharge is determined by the distribution of extreme events in excess of threshold levels (Lloyd, 1986).

In addition to estimation, accurate measurement of precipitation is also very difficult. Point measurements using standard rain gauges, such as weighing, capacitance, tipping-bucket, and optical types, are often used to measure the accumulation of rainfall at a particular location. However, these instruments can be subject to large measurement errors (Habib et al., 2001), and provide poor estimates of the spatial distribution of rainfall over a catchment. Ball and Luk (1998) analyzed the spatial distribution of rainfall over a catchment using several mathematical models in the GIS environment. They showed that all the methods based on spatial extrapolation of point data, such as Theissen polygons, inverse distance weights, kriging, and surface fitting, can result in significant errors in predicted rainfall distributions.

An accurate representation of rainfall is especially difficult in remote areas where the available information is limited. In many cases, synthetic weather generators based on historical weather records and stochastic analysis, such as WGEN (Richardson and Wright, 1984) and MORECS (Hough and Jones, 1998) are often used to generate synthetic weather data and rainfall distributions. However, these methods commonly use the extrapolation methods described by Ball and Luk (1998) and hence are subject to potentially large errors. The use of weather radar to predict precipitation distributions seems promising; however, it is also subject to errors (e.g., due to signal noise) and problems with scale or resolution (Kouwen and Garland, 1989).

In most groundwater modeling applications, rainfall rates are often considered constant spatially, or averaged over time in steady state analyses. Individual storm events are rarely accounted for and monthly

or annual averages are commonly used. The assumption of temporally averaged rainfall seems reasonable, at least for shorter time periods, as the fluxes in the vadose zone contributing to recharge are usually much slower than an individual storm. However, rainfall varies considerably even over short distances, and therefore, the spatial variation in precipitation should not be ignored, especially in larger scale modeling studies.

28.2.4.2 Land Use and Cover

The type of land cover can have a significant impact on infiltration, and hence on groundwater recharge. For example, the clearing of native vegetation and replacing it with shallow-rooted annual crops and pastures have resulted in substantial waterlogging and salinity problems in Australia due to increased recharge rates (e.g., Peck, 1978; Williamson, 1990; Walker et al., 2002). Urban areas can also have a profound impact on groundwater recharge by increasing the amount of impervious areas and generally altering the land surface from its natural state (Lerner et al., 1990; Lerner, 2002). However, the most significant impact of urbanization is often its adverse impact on the quality of recharge, rather than on its quantity (Foster, 2001).

The impact of land use changes on groundwater recharge has been the focus of modeling studies. Walker et al. (2002) give a comprehensive review of modeling approaches used in Australia to combat dryland salinity problems. Carmon et al. (1997) and Collin and Melloul (2001) investigated the impact of urbanization and changes in land use types in the context of sustainable groundwater development in Israel. They concluded that measures should be taken to reduce the negative effects of urban development on groundwater quality. Finch (2001) estimated the response of mean annual groundwater recharge to land cover changes in a rural catchment in southern England. He observed that land use changes lead to changes in the recharge pattern, but had no impact on the overall catchment total. Bellot et al. (2001) studied the impact of land use changes on runoff and recharge due to wildfires, afforestation, and land abandonment in Spain. They found that the increase in vegetated land cover has a negative effect on human water availability in semiarid areas, due to decreased aquifer recharge rates. Batelaan et al. (2003) used MODFLOW to study land use changes in Belgium by perturbing recharge rates at selected areas. Their study yielded mixed results.

28.2.4.3 Vegetation

Vegetation reduces recharge by directly interfering with the passage of precipitation from the atmosphere to the water table. The vegetation canopy intercepts a portion of the rainfall, which then either evaporates, or is channeled to the ground through stemflow, or drips directly to the ground as part of throughfall (Le Maitre et al., 1999). Arguably, these processes are very difficult to quantify as they are dependent on a multitude of climatic parameters, such as intensity and duration of rainfall, temperature, wind speed, as well as the physical characteristics of the individual plants (Larcher, 1983). However, they can still have a significant effect on the recharge process as shown by Taniguchi et al. (1996) and Finch (1998).

Perhaps the greatest influence on recharge by vegetation is through evaporation and transpiration. These terms are often lumped together with direct evaporation from the soil and referred to as evapotranspiration. Evapotranspiration is an important component of the water budget with up to 40 to 60% of the annual rainfall being lost through it in humid climates (Knutsson, 1988).

Methods to calculate actual evapotranspiration include the Bowen ratio method, derived from a simplified energy balance, and the eddy correlation method based on the correlation between wind speed and vapor density (e.g., Dingman, 1994). The downside of these models is that they require variables, such as aerodynamic, thermal, and plant specific parameters, which are very difficult to measure in practice. Other methods, such as Penman (1948), Penman-Monteith (Monteith, 1965), Thornthwaite and Mather (1957), and Priestley and Taylor (1972) have been developed to estimate evapotranspiration. However, these models also require many parameters that are derived from ideal conditions. Furthermore, most of these methods estimate potential evapotranspiration, which is not always equal to actual evaporation rates. The Penman–Grindley model (Grindley, 1969) is a widely used and simple soil moisture budgeting

and evaporation model, which also allows for the determination of recharge indirectly. Like the other models, however, it can be very sensitive to errors in the input parameters (Finch, 1998).

Plant roots also play an important role in the recharge process. Not only do they enable plants to draw water from deep in the vadose zone (and even from the saturated zone), thereby reducing the amount of percolating water that reaches the water table, but they also create preferential flow paths and channels that aid water flow through the soil profile (Le Maitre et al., 1999). Studies have revealed that in addition to absorption, plant roots can also exude water into the soil in response to gradients in water potential between themselves and the soil (Burgess et al., 1998; Schulze et al., 1998). When root systems span soil layers of different moisture content, water is redistributed by the roots in the direction of the difference in water potential. This hydraulic redistribution of water not only removes excess water from the topsoil and out of the reach of evaporation, shallow-rooted competitors, runoff, and lateral subsurface flow, but it may also increase the capture of mobile nutrients such as nitrogen from the topsoil (Burgess et al., 2001). Although the water stored at depth is unlikely to be significant for drought avoidance by plants, the downward transfer of water to dry soil layers following rain may be important to plant establishment and the reduction of water logging in certain soil types (Burgess et al., 2001).

Both Finch (1998), and Zhang et al. (1999b, 1999c) found that rooting depth, which depends on soil type, available water, and the type of plant, had a significant impact on groundwater recharge rates. The depth of plant root penetration is difficult to measure and can vary significantly even within the same plant community (Canadell et al., 1996; Jackson et al., 1996). Furthermore, only a few roots penetrating deep into the ground can sustain large plant communities (Le Maitre et al., 1999).

Finch (1998) and Zhang et al. (1999b, 1999c) also found that the leaf area index, or the ratio of leaf area to ground cover, has a significant influence on groundwater recharge. Larger leaf areas result in greater interception of precipitation as well as potentially higher transpiration rates leading to decreased groundwater recharge. Although standard methods exist for measuring the leaf area index (Pearcy et al., 1989), these methods are impractical for large areas. The values for leaf area index reported in the literature are also very difficult to apply to many sites due to the high variability of types of vegetation across catchments and watersheds. In addition, the leaf area index not only has high variability between different plants, but it also varies seasonally due to climatic changes.

In summary, the effect of vegetation on recharge can be both positive and negative. Vegetation intercepts rainfall and transpires water obtained from the rooted soil profile; however, it can also facilitate infiltration by reducing overland flow and creating surface storing opportunities. Plant root systems may also increase recharge rates by creating macropores and channels in the soil profile (Le Maitre et al., 1999), and by downward transfer of water due to differences in moisture potentials (Burgess et al., 2001).

28.2.4.4 Urbanization

The water balance of an area can be drastically altered by urbanization. Not only is the recharge rate affected, but the entire climate can also be altered (Lerner et al., 1990). A micro-climate may develop in large urban centers, causing changes in temperature, humidity, wind speed, and air clarity (Hall, 1984). While surface runoff is considerably increased due to increased imperviousness, the amount of direct groundwater recharge due to precipitation is decreased (Burges et al., 1998; Rose and Peters, 2001). However, most urban centers import their water for consumption from outside sources thereby increasing the amount of water in the area. A portion of this imported water then becomes recharged through septic tanks and leaking sewer and water distribution systems, and over-irrigation of lawns and parks (Lerner et al., 1990; Simmers, 1998; Yang et al., 1999; Foster, 2001; Lerner, 2002). This kind of indirect or man-induced recharge can result in higher recharge rates, or balance the loss of precipitation recharge due to impermeable areas (Lerner, 2002). However, determining the exact spatial and temporal distribution of this indirect recharge is exceedingly complicated and difficult to account for in any modeling efforts (Yang et al., 1999).

In summary, any changes in land cover, whether seasonal changes in vegetation or permanent changes due to urbanization, can impact groundwater recharge by altering the interception, infiltration, surface runoff, evaporation, and transpiration processes.

28.2.4.5 Overland Flow

Overland flow is a rare event in humid climates due to less intensive rainfall, well-developed vegetation, and sufficient infiltration capacity of most soils (Knutsson, 1988). Surface runoff occurs only over short distances or during special events such as intense snowmelt over frozen ground (Johnsson and Lundin, 1991). Furthermore, overland flow only takes place in the discharge areas, which usually occupy only a small part of a catchment (Knutsson, 1988). In areas of high topographic relief, with very low permeability soils such as clays and tills near the ground surface, surface ponding can occur, resulting in overland flow.

Overland flow can be estimated using the diffusion wave approximation of the depth-integrated shallow water equations with Manning's equation used for velocity calculations (Woolhiser et al., 1996; VanderKwaak, 1999). Due to small-scale variations in surface slope, the estimation of Manning's roughness coefficients across a catchment is difficult and subject to spatial averaging.

A popular method among surface hydrologists to estimate overland flow is using the Natural Resources Conservation Service (NRCS, formerly the Soil Conservation Service) curve number method (NRCS, 1986). This rainfall-runoff analysis method is based on a simple relationship between precipitation, infiltration, an initial abstraction, which is a function of interception and surface ponding, and an empirical curve number obtained from numerous field experiments. This method is widely applied due to its simplicity (e.g., Colosimo and Mendicino, 1996; Srinivas et al., 1999); however, it only provides estimates of runoff volumes, and ignores any subsurface flow mechanisms.

28.2.4.6 Infiltration and Flow in the Unsaturated Zone

Infiltration is the volume rate of water flowing into a unit area of soil surface. It is generally considered to be one-dimensional in the vertical direction; however, as noted before, local changes in ground cover and near surface permeability can lead to lateral flows. Percolation is the process by which water migrates down through the soil profile, whereas deep soil water percolation is the water that has moved past the evaporative and root zones in the unsaturated zone and is no longer available to plants.

Both infiltration and percolation are very complicated processes as they are governed by such factors as the rate of precipitation, antecedent moisture conditions in the soil, soil hydraulic properties, topography, and others. While infiltration equations (e.g., Green and Ampt, 1911; Horton, 1940; Philip, 1957) simply describe the rate of water movement into the soil at the ground surface, other relationships (e.g., Richards, 1931) are used to describe the percolation of water through the unsaturated zone.

The processes of infiltration and flow in the unsaturated zone have been much studied with many methods derived for their estimation. These processes are examined in detail in other chapters of this book.

28.2.4.7 Soil Properties

The hydraulic properties of the soils in the unsaturated zone are very sensitive to the moisture content and pressure head distributions. A small change in the volumetric water content can often result in a change in the hydraulic conductivity by two or more orders of magnitude (Rushton, 1988). In addition, the soils in the unsaturated zone rarely exhibit homogeneous properties, often consisting of layered sands, silts, and clays, resulting in non-uniform moisture distributions. Instability in the wetting front and subtle changes in the permeability structure can also lead to flow fingering (Kung, 1990; Selker et al., 1992). The fingering process is enhanced by initially dry, layered, coarse-grained systems (Sililo and Tellam, 2000); however, the antecedent moisture contents prevalent in the field in humid climates will generally inhibit its occurrence.

The unpredictable occurrence of preferred pathways, due to plant roots, cracks, and fissures, in even relatively homogeneous materials furthermore complicates the hydraulic characterization of soils in the unsaturated zone (Simmers, 1990). Large variations in recharge can also occur even across uniform soils due to topography, resulting in depression focused recharge (e.g., Freeze and Banner, 1970; Winter, 1983; Schuh et al., 1993).

28.2.4.8 Temperature

In northern climates, subzero temperatures result in snow accumulations in winter time, as well as frost formation in the near surface soil layers due to pore water freezing. Furthermore, in northern Canada, the layer of frost remains in the ground throughout the entire year and is referred to as permafrost. The presence of both a snowpack and a frozen soil layer will clearly have a significant impact on the recharge process.

Similar to rainfall, the spatial and temporal distribution of snow accumulation is very complex, and even further complicated by its high sensitivity to temperature and wind velocities (i.e., drifting) (Deng et al., 1994). Snowpack melting is a fairly well-understood process and is primarily based on the energy balance at the air–snow interface, and on the physical characteristics of the snowpack (Harms and Chanasyk, 1998). However, the presence and extent of the frost layer influences the amount and distribution of the infiltrating snowmelt. Snowmelt or rain located in one portion of a watershed may be redistributed by either surface runoff or interflow on top of the frozen soil layer depending on the topography (Johnsson and Lundin, 1991).

As the soil frost layer develops, the water in the soil pores freezes. The growing ice constricts or blocks the infiltrating water, thereby increasing the tortuosity and producing a net reduction in hydraulic conductivity and infiltration rates (Kane and Stein, 1983; Granger et al., 1984; Black and Miller, 1990). Infiltration rates can also vary depending on the initial soil water content before freezing (Engelmark, 1988).

The depth of the frozen soil layer is dependent on several factors such as temperature, the duration of freezing temperatures, snow depth at the ground surface, and initial soil water content (Daniel and Staricka, 2000). The porosity of the frozen soil may also change due to expansion caused by ice formation. The freeze–thaw process can also create large, temporary fractures resulting in higher hydraulic conductivities as the ice melts (Johnsson and Lundin, 1991). As shown by Jyrkama (1999), during and immediately after thawing of the frozen soil layer in the spring, rapid infiltration of the accumulated snowmelt can create a transient groundwater mound, resulting in significant recharge to the water table.

Therefore, due to the complex interaction of all the above processes, modeling unsaturated zone flow in a seasonally (or permanently) frozen environment is very challenging.

28.2.5 Recharge Estimation Methods

The link between climate change and groundwater is the precipitation recharge at the land surface and the recharge and discharge that occurs at surface water bodies. Reviews on the unsaturated zone flow processes and methods to quantify groundwater recharge rates from precipitation can be found in texts by Lerner et al. (1990) and Stephens (1996). In addition, Simmers (1988) and Sharma (1989) provide several examples on how these methods have been applied in many field studies. An overview of estimation problems and developments in groundwater recharge are also given by Simmers (1998) and the special theme issue of the *Hydrogeology Journal* (vol. 10, 2002). It should be noted that much of the literature and the associated methods have been developed for estimating recharge in arid and semiarid areas, where recharge plays both an economically and environmentally critical role.

Lerner et al. (1990) highlight five main methods of estimating direct recharge from precipitation:

- direct measurement
- empirical methods
- water budget methods
- Darcian approaches
- tracers.

Other methods include, for example, plane of zero flux, temperature and electromagnetic methods, baseflow separation, remote sensing, and inverse groundwater modeling (Stephens, 1996). A brief discussion of these and other additional methods is presented in the following sections.

28.2.5.1 Field (or Direct) Techniques

The most common and practical instrument to measure recharge directly is with a lysimeter. A lysimeter is a block of soil instrumented such that all the parameters in the water budget can be quantified, with the recharge flux being either measured directly at the bottom (e.g., Lerner et al., 1990), or estimated based on the changes in the block weight (Kirkham et al., 1984; Yang et al., 2000). The inherent problems with lysimeters are that they are difficult and expensive to construct and only provide a point or an average value of recharge for the materials within the block. Soil disturbance is also a problem, as well as leaks and other failures resulting in potentially erroneous measurements. The major difficulty with weighing lysimeters is in measuring small weight changes in relation to the large and heavy soil mass.

As an example, Fayer et al. (1996) used three lysimeters, in conjunction with other methods, to estimate recharge rates at the Hanford nuclear waste site in the United States. They measured drainage from the bottom of each instrument over several years to estimate an average annual recharge rate. The lysimeters were of different depths and filled with soils typical of their surroundings. The main problem with their measurements, however, was that the instruments were too shallow (i.e., within or above the root or evaporative zone). Therefore, in spite of the long measurement period used, the collected drainage may have overestimated the actual recharge rates in the area. To avoid problems with the larger size instruments, Holder et al. (1991) developed a special wick pan lysimeter, also called the passive capillary sampler (PCAPS). The PCAPS is generally constructed from a small sealed bucket and instrumented with a wetted fiberglass wick that acts as a hanging water column and develops suction in the soil water depending on the flux. Louie et al. (2000) used a number of PCAPS and found them to perform satisfactorily under various field conditions.

The matric potential and water content measurements using tensiometers and time-domain reflectometry (TDR) can also be used in the field to estimate hydraulic gradients and storage changes in the vadose zone. As a pulse of infiltrated water moves through the unsaturated zone, a plane of zero flux, that is, where the hydraulic gradient is zero, may develop as the water above the plane moves upward due to evapotranspiration while below the plane, water moves downward due to gravity (e.g., Scanlon et al., 2002). Water below the zero flux plane will therefore become recharged, and the amount or rate can be calculated by monitoring the change in water content below the plane. However, this requires extensive monitoring in the field and is therefore unfeasible for large areas. Furthermore, the location of the zero flux plane changes with time and only applies to a pulse of infiltrated water. Therefore, the method breaks down when there is a downward hydraulic gradient due to a continuous precipitation event (Stephens, 1996).

Temperature and electromagnetic methods have been developed based on the fact that surface waters and groundwaters have different temperatures and chemical compositions. The rate of recharge can be estimated based on measured borehole temperature profiles (Taniguchi, 2002) or through soil temperature measurements (Taniguchi and Sharma, 1993). Electromagnetic methods are mostly used for reconnaissance investigations to screen sites where more quantitative methods for recharge could be applied (Cook and Kitty, 1992). Similarly, remote sensing, such as Landsat imagery (Salama et al., 1994) or InSAR analysis of ERS-1 and ERS-2 images (Lu and Danskin, 2001) have been used to distinguish regions of recharge and discharge over larger areas.

28.2.5.2 Empirical Methods

Recharge rates have been estimated by correlating precipitation with recharge using an empirical relationship often based on regression (e.g., Sinha and Sharma, 1988; Shade and Nichols, 1996; Yang et al., 1999; Nichols and Verry, 2001; Xu and van Tonder, 2001; Chen et al., 2002) or stochastic analysis (Wu et al., 1997; Gau and Liu, 2000). Although these methods may be adequate for the particular study area and purpose, they are not applicable to other areas. Expressing recharge as a function of precipitation has no physical basis, as all the unsaturated zone processes described in the previous section are completely ignored.

Methods have also been developed to quantify groundwater recharge based on the analysis of the recession curve in the streamflow hydrographs (e.g., Rorabaugh, 1964; Chapman, 1999; Wittenberg and

Sivapalan, 1999; Arnold et al., 2000). It is assumed that the recession part of the hydrograph represents primarily groundwater discharge to the river as a direct result of precipitation recharge to the aquifer. Erskine and Papaioannou (1997) further developed a relationship called the aquifer response rate based on maximum permissible abstraction and minimum flow requirements for a river.

Halford and Mayer (2000) studied various hydrograph separation techniques and concluded that baseflow is frequently not equivalent to groundwater discharge because other hydrologic phenomena can significantly affect stream discharge during the recession periods. Drainage from bank storage, wetlands, surface water bodies, soils, and snowpacks also decreases exponentially during recession periods, and along with evapotranspiration, were shown to affect the stream discharge more than groundwater discharge during the recession periods. These methods have obvious limitations, therefore, and at best only provide an approximation of the overall volume of groundwater recharge over a selected time period.

28.2.5.3 Water Budget Methods

Soil water budget methods are based on a soil moisture balance whereby all the components of the water balance equation are estimated with groundwater recharge being the residual (also see Figure 28.1)

$$R = P - ET \pm O \pm \Delta S \quad (28.1)$$

where R is the recharge, P is precipitation, ET is actual evapotranspiration, O is the lateral surface runoff in and out of the area, and ΔS represents the change in water storage in the unsaturated zone. Assuming a unit cross-sectional area, each process can be described as a rate or a flux. As indicated, water is added or depleted from the soil storage with any excess becoming groundwater recharge.

Due to the difficulty in accurately estimating each of the processes, propagation of errors in the water budget method can result in large uncertainty in the calculated recharge. Furthermore, this method ignores the movement of moisture, that is, percolation, through the unsaturated zone, and therefore fails to account for the time it takes for the infiltrated water to reach the water table. Recharge is assumed to occur instantaneously as water is released from soil storage. Therefore, the temporal recharge pattern or the timing of recharge is difficult to capture with this method. The water budget method has been used to calculate average recharge rates in various situations (e.g., Bekesi and McConchie, 1999; Srinivas et al., 1999; Finch, 2001; Chapman and Malone, 2002).

28.2.5.4 Darcian Approaches

Water movement in the unsaturated zone can be modeled based on Darcy's equation and conservation of mass (Richards, 1931). The strength of this method is that it is physically based, however, in addition to the uncertainty in the antecedent moisture conditions and heterogeneities of the soil column, the high nonlinearity between the hydraulic conductivity, pressure head, and water content in the unsaturated zone pose several difficulties in its application. This method is therefore computationally intensive and requires detailed information about the hydraulic characteristics of the unsaturated zone. These inherent difficulties are compounded by the potential presence of root channels and other fissures, which may dominate the recharge process. However, accounting for the moisture movement through the unsaturated zone explicitly provides a robust method of estimating groundwater recharge (assuming errors in the input data can be minimized).

28.2.5.5 Tracers

A host of environmental tracers have been used to estimate groundwater recharge. These include tritium and chlorine-36, chloride, carbon-14, chlorofluorocarbons, and stable isotopes such as deuterium, and ^{18}O . These methods require the measurement of a concentration profile in the soil column, which provides information about water movement in the unsaturated zone. Recharge is then estimated based on a tracer mass balance, total water content analysis, or numerical modeling. These methods usually assume piston displacement and result in an estimate of either a long-term average recharge rate or a total volume of recharge (e.g., Allison, 1988; Scanlon et al., 2002).

28.2.5.6 Inverse Groundwater Modeling

Recharge is commonly used purely as a calibration parameter in many groundwater modeling studies (e.g., Martin and Frind, 1998; Bonomi and Cavallin, 1999; Javed and Bonnell, 1999; Varni and Usunoff, 1999; Ella et al., 2002). Therefore, the resultant recharge pattern can potentially be used to quantify recharge rates in the area. Zoning of recharge into equivalent areas is often used to reduce the total number of calibration parameters.

The concepts of inverse calibration and zoning are based on earlier conceptual models (e.g., Toth, 1963; Freeze and Witherspoon, 1968; Winter, 1978, 1983) where the specific recharge and discharge zones emerge naturally as a response to geologic or topographic conditions. The distribution of recharge can therefore be delineated with the knowledge of the depth of the water table and the underlying hydraulic conductivity distribution, while the rate of recharge can furthermore be estimated using the Darcy equation. Stoertz and Bradbury (1989) and Levine and Salvucci (1999) used this approach with numerical groundwater flow modeling to delineate recharge areas and rates.

Using the same concept, Lin and Anderson (2003) developed a recharge and discharge mapping procedure based on pattern recognition and image analysis. Similar to Stoertz and Bradbury (1989) their method is based on a mass balance calculation performed using MODFLOW, and therefore is critically dependent on the accuracy of the water table interpolation. Zoning is used to smooth out the results as “it is not practical to work with numerous grid specific rates in a parameter estimation code, nor are the individual values expected to be very accurate” (Lin and Anderson, 2003). Scale is also a major problem with this method as the resolution of the hydraulic gradient between adjacent model cells decreases with finer grid spacing. As the method ignores all the physically based processes in the hydrologic cycle contributing to recharge, and only considers parameters used in the groundwater flow model, Lin and Anderson (2003) admit that the method may result in unreasonable estimates of groundwater recharge, for example, recharge greater than precipitation.

The main problem with inverse groundwater model calibration is that it relies entirely on the accuracy of the groundwater flow model which often has great uncertainty associated with it due to limitations in the input data. Small changes in the calculated hydraulic conductivity field, for example, can result in potentially large variations in the recharge estimates. Furthermore, because hydraulic conductivity ranges over several orders of magnitude, estimated recharge rates are highly uncertain. The results are also non-unique because the same distribution of hydraulic heads can be produced with a range of recharge rates, as long as the ratio of recharge to hydraulic conductivity remains the same (Scanlon et al., 2002). This approach, therefore, ignores all the physical processes in the hydrologic cycle and disregards the dynamic interaction between the surface water and groundwater systems.

28.2.5.7 Combined Modeling

Groundwater recharge has also been estimated by using a mixed approach, based on a combination of the water budget methods and the Darcy equation. The problem with the Darcian approach alone is that it lacks the description of many of the important hydrologic processes. It simply calculates the movement of water through the soil column based on limited boundary conditions. Using the water budget method first to estimate important processes such as surface runoff, evapotranspiration, and snowmelt, for example, and then calculating percolation through the unsaturated zone using the Darcy equation, will preserve the hydrologic moisture balance and allow recharge to be quantified in both space and time.

This approach is also well suited to the investigation of the impact of land use and climate changes on groundwater recharge. The various weather input parameters, such as precipitation and temperature, can be perturbed according to predicted changes in the future climate, while the ground surface properties, such as runoff potential and rooting depths, can be modified due to expected changes in land use.

Examples of combined models include Bauer and Vaccaro (1987), UNSAT-H (Fayer and Jones, 1990), TOPOG_IRM (Dawes and Hatton, 1993), PRZM-2 (Mullins et al., 1993), HELP3 (Schroeder et al., 1994), Sandstrom (1995), ANSWERS (Bouraoui and Dillaha, 1996), PERFECT (Abbs and Littleboy, 1998), SWAT (Arnold et al., 1998), and WAVES (Zhang and Dawes, 1998). The rate of recharge in each model is simulated using a combination of various vegetation, climate, and hydrologic models.

28.2.6 Recharge in Groundwater Modeling

28.2.6.1 Introduction

Groundwater models are employed for numerous hydrologic investigation purposes such as vulnerability assessments, remediation designs, and water quantity estimations. Most often, however, they are used in groundwater quality studies dealing with the increasing number of emerging groundwater contamination problems. Groundwater models are applied across various scales, from very small scale flow analyses (e.g., Srinivas et al., 1999; Ella et al., 2002), to the regional, or even national scale groundwater flow studies (e.g., Brodie, 1999; Varni and Usunoff, 1999; Vermulst and De Lange, 1999). Both steady-state and transient modeling strategies are employed, depending on the extent and nature of the particular problem.

Whereas fully saturated models, such as the United States Geological Survey (USGS) modular finite-difference groundwater model MODFLOW (McDonald and Harbaugh, 1996), only simulate groundwater flow below the water table, variably saturated models that depict flow in the unsaturated zone also exist, and are used in some instances (e.g., Freeze, 1971; Huyakorn et al., 1986; Deng et al., 1994; Therrien and Sudicky, 1996). Infiltration at the ground surface is used as the top boundary condition in variably saturated models, while groundwater recharge is calculated explicitly as part of the flow solution. Although subsurface fluid flow (both air and water) is therefore modeled as a continuum, this approach is still limited by the scaling problem. The highly nonlinear relationships between water content, pressure head, and unsaturated hydraulic conductivity requires very fine discretization, limiting the application to smaller areas due to increased computational costs.

In addition to added input data, variably saturated models also need to account for the near surface hydrologic processes such as infiltration, evapotranspiration, frost penetration, and snowmelt that influence not only the top boundary condition, but also the upper few meters of the unsaturated zone. Fully saturated groundwater models, on the other hand, only require the specification of the recharge boundary condition at the water table and are therefore more commonly used in most groundwater modeling investigations.

Variably saturated flow is described by the Richards' equation, while transient three-dimensional saturated groundwater flow in heterogeneous anisotropic porous media is expressed as (Bear, 1972)

$$\frac{\partial}{\partial x_i} \left(K_{ij} \frac{\partial h}{\partial x_j} \right) + Q = S_s \frac{\partial h}{\partial t} \quad (28.2)$$

where K_{ij} is the hydraulic conductivity tensor, h is the hydraulic head, Q represents sources or sinks, and S_s is the specific storage. As fully saturated models only describe groundwater flow below the water table, the specification of either the location of the water table (Dirichlet — Type I) or the recharge flux (Neumann — Type II) is required at the top or upper boundary. Depending on the model, the flow equations are commonly solved either by the finite difference method, for example, MODFLOW (McDonald and Harbaugh, 1996), or the finite element method, for example, WATFLOW (Molson et al., 1992).

As the exact location of the phreatic surface is often difficult to determine, and small errors in its location may potentially lead to large errors in groundwater fluxes, the recharge flux is often specified as the upper boundary condition. However, due to the many difficulties in recharge estimation, as alluded to earlier, the recharge boundary condition is commonly used purely as a calibration parameter to improve the model fit (e.g., Martin and Frind, 1998; Bonomi and Cavallin, 1999; Varni and Usunoff, 1999). Where information about precipitation is available, a fraction of it is also often assigned as the recharge boundary condition (Kennett-Smith et al., 1996; Brodie, 1999). The concept of the recharge spreading layer (RSL) has also been proposed to redistribute recharge from areas of low permeability to areas of relatively high permeability (Therrien and Sudicky, 1996; Martin and Frind, 1998; Beckers and Frind, 2000). Although these assumptions may be reasonable for the long-term simulation of a regional groundwater flow system or flow in a deep confined aquifer, a physically based and more accurate description of the recharge

boundary condition is required for studies involving climate change, where the detailed knowledge of the spatial and temporal variation of the recharge is imperative.

28.2.6.2 Coupled Models

To obtain a more physically based recharge boundary condition for saturated groundwater models, several researchers have adopted a coupled approach (Chiew et al., 1992; Prathapar et al., 1994; Fayer et al., 1996; Beverly et al., 1999; Srinivas et al., 1999; Vermulst and De Lange, 1999; Zhang et al., 1999a; Sophocleous and Perkins, 2000; Batelaan and De Smedt, 2001; Ruud et al., 2001; Coppola et al., 2002). Groundwater recharge is estimated separately, using methods described Section 28.2.5, and the results are then used as a boundary condition in a saturated groundwater flow model. While most, if not all, of the methods presented in Section 28.2.5 have been used in groundwater modeling studies, the following discussion will focus on combined vadose zone–water budget modeling, as it provides the most versatile and physically based means of estimating the recharge boundary condition.

Chiew et al. (1992) developed a coupled approach based on the rainfall-runoff model HYDROLOG and the finite element groundwater flow model AQUIFEM-N to estimate regional groundwater recharge rates. The surface hydrologic model was used to calculate daily recharge rates, which were then summed over a month and input into the groundwater flow model as the recharge flux boundary condition. While their methodology may provide improved results as calibration is conducted against both streamflow and potentiometric head data, there are still many shortcomings and limitations in their approach. First, they divided their catchment into ten sub-areas based on drainage divides, rather than into areas with similar hydrologic characteristics. This resulted in spatially averaged values being used for each sub-area in the recharge analysis. Second, the rainfall-runoff model HYDROLOG completely ignores flow in the unsaturated zone and is simply calibrated against streamflow hydrographs with recharge estimated as a result. Even though the analysis is conducted based on daily precipitation records, the combination of ignoring the unsaturated zone and using average parameters to represent each sub-area results in an averaged estimate of recharge for a given sub-area. This is clearly shown in their simulated groundwater levels, which fail to match the transient variations observed in the groundwater flow field resulting from significant recharge events.

Prathapar et al. (1994) developed a soil water and groundwater simulation model SWAGSIM to estimate water table fluctuations in extensively irrigated regions. They calculated daily recharge at the water table based on an analytical solution to the Richards' equation and added it to a two-dimensional finite difference groundwater flow model through the source/sink term. The main problems with their model include the simplification of important hydrologic processes, the requirement of homogeneous and isotropic conditions for the analytical solution, and the two-dimensional nature of the groundwater flow model (essentially assumes a one-layer system with assigned transmissivities and specific yields). Therefore, their model is only applicable to very simple systems.

Using a geographic information system (GIS), Fayer et al. (1996) estimated the spatial distribution of recharge with four different methods for their groundwater study at the Hanford nuclear waste site in the United States. They used a combination of field techniques such as lysimeters, water content measurements, and tracer studies in conjunction with numerical modeling (UNSAT-H) to generate a temporally averaged recharge boundary condition map for their groundwater flow model of the study area. The use of the GIS aided significantly in the analysis by allowing the identification of all possible combinations of soil type and vegetation in the area that could be assigned an appropriate estimate of recharge. The inclusion of field studies provided additional confidence to the unsaturated zone modeling results. However, the resulting recharge map only depicts average annual recharge rates and neglects to provide information regarding the change of recharge over time. The results from the groundwater model were not discussed.

Due to data limitations required by more physically based models, Beverly et al. (1999) developed a simple unsaturated module SMILE to provide recharge estimates to MODFLOW. Their model accounts for overland flow and evaporation from both the unsaturated and saturated zones and recharge is calculated based on either a matrix or crack flow approximation, or an empirical Darcy approach. The two models

are again linked through the recharge boundary condition; however, their approach also allows for groundwater discharge, due to heads in excess of the ground surface, to be distributed in overland flow and reinfiltrated in the subsequent time steps. This dynamic linking seems very attractive with a small computational burden due to the simplicity of the model; however, the process descriptions are overly simplified and may lead to significant errors in the results.

Srinivas et al. (1999) used the water budget method to calculate monthly recharge rates for their nested squares groundwater flow model NEWSAM of a small fractured granitic aquifer in India. They used the SCS curve number method to calculate an average curve number for the watershed and applied the water balance method on a monthly basis to estimate recharge. The estimated temporal recharge pattern was then coupled to the groundwater flow model through the boundary condition. Their results showed the importance of including temporally varying recharge in groundwater flow modeling to capture the temporal fluctuations observed at the water table. However, their model consistently over predicted the observed water levels by a wide margin, although it seemed to capture the timing of changes in the flow field. Using the water budget method across the entire watershed leads to spatial averaging, which may account for the problems in their simulated data. As mentioned previously, the water budget method ignores percolation through the unsaturated zone resulting in average results.

Vermulst and De Lange (1999) used a GIS interface to link the unsaturated flow model MOZART and the analytic element groundwater flow model NAGROM to aid in national groundwater policy analysis in the Netherlands. As their models operated at different spatial and temporal scales, they coupled them through the Cauchy boundary condition using up- and down-scaling procedures. The models were run sequentially until a steady state was obtained. Although recharge was averaged both spatially and temporally, their model provides a valuable tool for policy decisions at the national scale. However, the model has limited potential for groundwater quality studies of point source contamination.

To simulate catchment scale responses under different land management options in the context of salinity control, Zhang et al. (1999a) used a distributed parameter ecohydrological model TOPOG-IRM to estimate soil moisture and groundwater recharge rates. The model accounts explicitly for the spatial distribution of topography, soil characteristics, and vegetation properties and predicts the dynamic interactions within the soil–vegetation–atmosphere system. It also uses the Richards' equation to simulate water movement in the unsaturated zone and links soil and canopy processes with catchment topography. Evapotranspiration is based on a Penman–Monteith type approach, with transpiration based on a canopy resistance model. The model uses daily climate variables such as precipitation, temperature, vapor pressure deficit, and solar radiation to estimate the recharge rate at the bottom of the soil column. Their model, therefore, seems to offer a very sophisticated approach for estimating groundwater recharge.

Using the results from the recharge analysis, Zhang et al. (1999a) also used MODFLOW to simulate groundwater flow in their catchment. The simulated water levels, however, failed to match the temporal variations in the observed data. This was likely not caused by problems in the recharge analysis, but by the simplicity of the groundwater flow model. Due to the lack of soil information in the study area, they assumed homogeneous conditions across the entire site. Furthermore, the calculated recharge was only approximately 5% of the annual rainfall in the area, resulting in a very small recharge boundary flux.

Sophocleous and Perkins (2000) combined the semi-distributed agricultural watershed model SWAT (Arnold et al., 1998) with MODFLOW to study conjunctive water use and management scenarios in three watersheds in Kansas. They developed a linking routine where the results from SWAT were distributed or averaged over the groundwater model grid and passed as inputs into MODFLOW at each MODFLOW time step. They also utilized GIS to aid in parameterization of hydrologic variables using overlaying and averaging operations. They calibrated the model against actual streamflows and groundwater levels in three watersheds and found reasonable agreement between the simulated and observed results. In addition to allowing calibration to multiple targets, the main advantage of their method is in the flexible modular linking approach that allows other watershed models to be used in place

of SWAT. The main shortcoming, however, is that the linked model is based on the HRU concept resulting in averaging of important parameters. Mean sub-basin responses are first estimated by averaging the mean results from each of the HRUs within the sub-basin, and then distributed over the underlying groundwater model grid. The inputs to the groundwater model are therefore not only averaged temporally, but spatially as well.

Batelaan and De Smedt (2001) developed a GIS-based methodology to provide recharge estimates for regional groundwater modeling. Their model, WetSpass, estimates recharge on a raster cell basis using a simplified water balance formulation. They used the results of the recharge analysis as boundary conditions in MODFLOW to study the effect of land use changes on groundwater discharge in a catchment in Belgium (Batelaan et al., 2003). While their method may provide reasonable estimates of average annual recharge rates at the regional scale, it is unsuitable for small temporal and spatial scales. One of the main drawbacks of the method is the coarseness of the raster approach, which requires average parameter values to be used for each raster cell. However, the central weakness of their method is the crude water balance, which employs simple empirical relationships to describe important hydrologic processes, with recharge estimated as the residual.

Ruud et al. (2001) used a similar GIS-based approach in their coupled groundwater and surface water flow model at the basin scale. They linked a surface water model, a land-atmosphere interface and unsaturated zone model, and a groundwater flow model (MODFLOW) in GIS to evaluate conjunctive use alternatives and their impacts on groundwater resources in response to land use demands and potential changes to surface water deliveries in their study area. Monthly rates of recharge were calculated from the surface water models and input as the recharge flux boundary condition in MODFLOW. The main weakness in their approach is that the modeling is conducted at a very large scale leading to spatial averaging of important processes and parameters. Furthermore, they used the water budget method (and a quasi-hydraulic model) to compute the recharge rates at the water table which ignores or simplifies the impact of the unsaturated zone soil column.

As part of a groundwater flow study in New Jersey, Coppola et al. (2002) developed a fuzzy-rule based approach to estimate monthly recharge rates for the groundwater flow model. Instead of estimating recharge using a physically based description of hydrologic processes, they employed a fuzzy-rule based approach relating monthly groundwater recharge to mean monthly air temperature, monthly precipitation, and the previous month's total streamflow. They used the method to estimate monthly basin wide recharge values during transient calibration of the groundwater flow model. Similar to other empirical approaches, however, their method is only a statistical tool and lacks physical basis. Furthermore, at best, it can only estimate average basin wide values as the analysis is based on monthly streamflow totals from the basin.

In summary, various types of coupled models have been developed for different purposes and applications. The approaches portray various degrees of sophistication, from the rigorous numerical treatment of many hydrologic processes, to simple water budget or empirical models. Regardless of the many obvious weaknesses present in many of the models, these methods provide a more versatile and practical way to estimate groundwater recharge than the more expensive and intensive field methods. Physically based method are especially useful, as they not only provide current estimates of recharge, but they can also be used to study the impact of changes in the physical processes, for example, in temperature and precipitation due to climate change, on recharge distribution and rates.

28.2.6.3 Integrated Models

To overcome the problems inherent in the coupled approaches, several researchers have attempted to model the entire hydrologic system as a whole. The basic design of the integrated surface-subsurface modeling approach was first outlined by Freeze and Harlan (1969). Yan and Smith (1994) proposed a framework for an integrated approach based on integrating a surface water management model SFWMM with MODFLOW. Although the idea of integrating existing and widely used models sounds appealing, their approach, however, suffers from many simplifying assumptions common with the coupled methods.

Furthermore, besides the proposed framework, details of model development and applications have not been found in literature to date.

Using the GIS, Xiao et al. (1996) developed a raster-based spatially and temporally continuous surface-subsurface hydrologic flow model. The model is based on a dynamic linkage between precipitation, surface flow, infiltration, and subsurface flow submodels and uses simple physical relationships instead of the more rigorous higher order transient differential equations. The processes of interception and evapotranspiration are ignored due to sparse vegetation and the assumption that evapotranspiration rates would be low during a rainfall event due to high humidity. The equations are solved sequentially, and the percolation of water through the unsaturated zone is ignored. The model was applied to a large watershed in northern Alaska to simulate the system response to a single rainfall event. In addition to the inherent problems with a simplified model, the lack of evapotranspiration, snow pack, and snowmelt processes limit the application of the model to simple systems.

Querner (1997) developed a physically based groundwater and surface water model MOGROW and used it to investigate the effects of human interventions and natural factors on groundwater recharge at the regional scale in the Netherlands (Querner, 2000). The model combines a variably saturated finite element groundwater flow model SIMGRO with a surface water model SIMWAT. Although advertised as an integrated model, the two models operate at different time steps and are linked through the surface boundary condition. Furthermore, the flow in the unsaturated zone is modeled with a pseudo-steady-state approach with empirical parameters introduced for surface runoff, perched water tables, and preferential flow. Results from their study indicate that the model may be suitable for long-term assessments of changes in the groundwater flow field due to changes in land use, groundwater abstractions, and meteorological conditions.

VanderKwaak (1999) developed a fully integrated surface water and groundwater model that rigorously considers the flow and transport processes on the land surface and in the variably saturated, dual-continua subsurface. The two-dimensional diffusion wave equation is used to describe the shallow surface water flow, while groundwater flow is described using the three-dimensional form of the Richards' equation. The linkage between the surface and subsurface systems is through first-order, physically based flux relationships or through continuity assumptions. The entire system of equations is also solved simultaneously resulting in a truly integrated model. The weakness of the model, however, is that it ignores important hydrologic processes such as interception and evapotranspiration, which can account for significant losses in long-term flow simulations. Snowmelt processes are also ignored, limiting the application of the model to warmer climates or short-term simulations. Parameterization and the scaling problem between the surface and subsurface processes furthermore restrict the application of the model to smaller areas with small variations in properties or characteristics (VanderKwaak and Sudicky, 2000).

Montgomery Watson (1993) has also created an Integrated Groundwater Surface Water Model (IGSM) based on a quasi three-dimensional finite element approach. The major components of the hydrologic cycle simulated by the model include rainfall, runoff, groundwater recharge, pumping, consumptive use, evaporation, subsurface flows, and seawater intrusion. As stated by Fuller (1995) "integration of groundwater and surface water flows is carried out using a soil mixture accounting and unsaturated flow model. This, combined with runoff and soil percolation, allows for full interaction between surface and groundwater. Water quality simulation is also included which can track a contaminant through the processes of advection, dispersion, and dissolution." The model has been applied in basin management studies in California, Florida, and Colorado. Due to its increasing popularity, LaBolle et al. (2003) conducted a comprehensive review of the underlying assumptions and mathematical formulations present in the IGSM model. In addition to using empirical land use and vadose zone models to compute net percolation or recharge to the water table, they found that "the model fails to properly couple and simultaneously solve groundwater and surface water models with appropriate mass balance and head convergence under reasonable conditions" (LaBolle et al., 2003). They further concluded that due to instabilities and errors caused by some of the algorithms in the model, the model results may produce misleading predictions and interpretations that may influence planning and management decisions.

HydroGeoLogic (2002) developed an integrated hydrologic modeling system MODHMS based on the USGS MODFLOW model. They added modules for three-dimensional flow in the unsaturated zone using the Richards' equation, two-dimensional overland flow using the diffusion wave approximation, and flow through a network of one-dimensional channels or pipes using the diffusion wave approximation with Priesmann Slot conceptualization for pressurized flow in pipes. The model is fully compatible with GIS and provides an interface with ArcView for accessing and analyzing spatially distributed input parameters. The model has been applied to various problems, such as groundwater interaction with rivers and saltwater intrusion analysis. Due to the lack of published data, however, a more detailed description of the model cannot be given here.

Graham and Refsgaard (2001) give a descriptive overview of the fully distributed hydrologic model MIKE-SHE, which is an extension of the original Système Hydrologique Européen (SHE) code (Abbott et al., 1986). They assert that "MIKE-SHE is one of the only distributed, physically based, fully integrated surface water/groundwater models available today." The model includes descriptions for many of the important hydrologic processes, including snowmelt, with well-known non-empirical equations used to represent the physical processes in the different parts of the hydrologic cycle (Graham and Refsgaard, 2001). The model and its earlier versions have been around for quite some time and therefore have been applied to numerous hydrologic problems. One of the weaknesses of the model includes the strictly one-dimensional flow assumption in the unsaturated zone and sheet flow approximation of overland flow. Furthermore, as with any of the integrated hydrologic models, the central problem is with parameterization and increased computational requirements due to rigorous mathematical equations as well as coupling, or integration, of processes that take place on vastly different spatial and temporal scales.

Singh and Woolhiser (2002) give a comprehensive review of hydrologic models reported in literature over the years. While most of the models listed in their paper deal only with surface water hydrology, references to more comprehensive models are also included.

In the end, the problem comes down to the purpose and scope of the model, and its associated conceptual model. As Reddi and Danda (1994) pointed out, "while simple lumped parameter models may be adequate in regional groundwater quantity studies on homogeneous soil domains, models addressing multidimensionality and spatial heterogeneity are necessary in groundwater quality studies, such as remediation designs at contaminated sites." Therefore, depending on the objectives and purpose of the model, the use of steady-state or average parameters may very well be appropriate. However, this does not imply that using average values, especially across many spatial or temporal scales, results in an accurate model. Parameter sampling and modeling must be done at the appropriate scale to accurately describe the processes that the model intends to simulate.

28.3 A Case Study: Recharge Analysis of the Grand River Watershed

28.3.1 Introduction

The Grand River watershed is located in south-western Ontario, draining an area of nearly 7000 km² into Lake Erie. The location of the watershed is shown in Figure 28.2. The main tributary is approximately 290 km in length with an elevation differential of about 362 m from its source to the mouth. The landscape of the watershed has mainly been shaped by the last period of glaciation, resulting in highly variable soils and topography. The southern part of the watershed consists of low permeability lacustrine clay deposits and low topographic relief. The central part is formed mostly of higher permeability sand and gravel kame moraines with moderately high relief, while the northern portion of the watershed is comprised of lower permeability till plains with varying surface relief (Holysh et al., 2000).

Although 90% of the watershed is classified as rural, the watershed contains some of the fastest growing urban areas in Ontario, such as the cities of Kitchener, Waterloo, Cambridge, and Guelph. Not only is



FIGURE 28.2 Location of the Grand River watershed.

increasing urbanization stressing the existing water supply, but it is also placing the supply at a greater risk of contamination. There is growing concern about the environmental impact of such rapid urbanization and the ability of the river and groundwater systems to meet the rising demand for water. The recent and continuing drought conditions in southern Ontario have also placed an additional stress on the hydrology and water resources of the watershed.

The objective of this study is to assess the spatial and temporal distribution of groundwater recharge in the Grand River watershed and to estimate the impact that climate change may have on groundwater recharge, evapotranspiration, and runoff. Approximately 80% of the population in the watershed derive their drinking water from groundwater. Therefore, quantifying the input to the groundwater system is critical for developing an effective groundwater management strategy that will ensure an adequate and clean supply of drinking water, now and in the future.

28.3.2 Input Data

Numerical modeling at the regional watershed scale, such as the Grand River, involves the handling of large amounts of input and output data. A geographic information system (GIS) is ideally suited for this purpose as it provides an integrated platform to manage, analyze, and display model parameters and results. Due to the large volume of temporally varying climate data (e.g., daily precipitation, temperature, and solar radiation), a separate relational database management system (RDBMS) was also utilized in this work to facilitate data organization and storage.

The physically based hydrologic model HELP3 was used to calculate the spatially varying daily groundwater recharge rates for the watershed. The model was linked to ArcView GIS and the database management system (MS Access) using simple Visual Basic and ArcView Avenue scripts. Details of the methodology can be found in Jyrkama et al. (2002). HELP3 was chosen mainly because it was readily available and easy to use. Furthermore, HELP3 simulates all of the important hydrologic processes in the water budget, including the effects of snowmelt and freezing temperatures. Following is a summary of the input data used in the model.

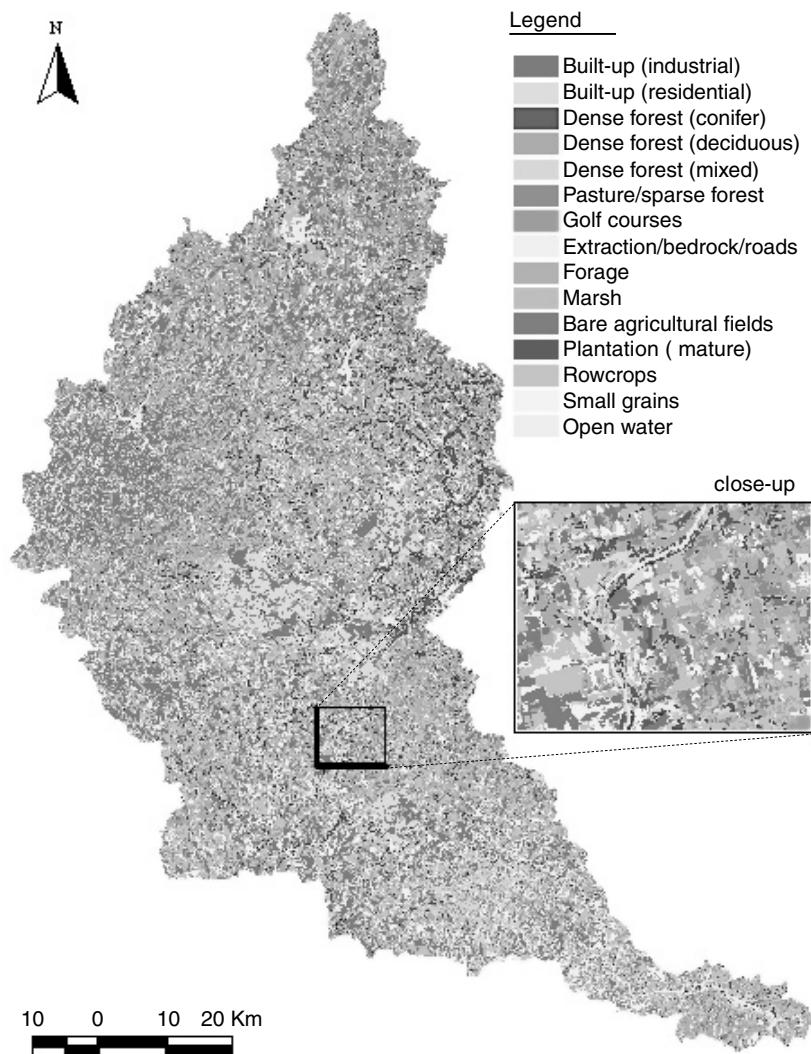


FIGURE 28.3 Land use/land cover (LULC) map.

28.3.2.1 Land Use/Land Cover Data

Digital land use and land cover (LULC) data for the Grand River watershed were obtained using 1999 satellite imagery from the Landsat 7 thematic mapper. As shown in Figure 28.3, the raster LULC coverage is based on a 25-m grid with 15 unique field descriptors.

28.3.2.2 Soil Data

The surface soil information in Figure 28.4, was assembled from various regional soil surveys conducted in the watershed. All the information was obtained from the Canadian Soil Information System (CANSIS) website in digital form. However, several difficulties were encountered during the construction of the soil map and the associated database for the watershed.

The Grand River watershed spans a number of different counties and municipalities, each with its own soil survey and database, as shown in Table 28.1 and Figure 28.2. Due to differences in scales and mapping methods, relatively minor to severe discontinuities and overlaps existed between adjacent soil map sheets. Therefore, to generate a continuous soil map for the entire watershed, these gaps and overlaps needed to

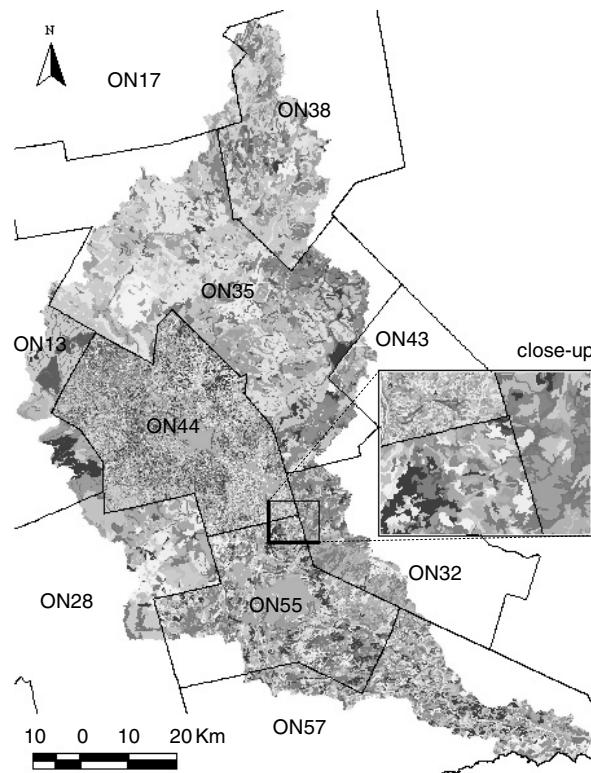


FIGURE 28.4 Soil surveys and soil map.

TABLE 28.1 Soil Surveys in the Grand River Watershed

Survey	County	Scale	Year
ON13	Perth	1:63360	1975
ON17	Grey	1:63360	1981
ON28	Oxford	1:63360	1961
ON32	Wentworth	1:63360	1965
ON35	Wellington	1:63360	1963
ON38	Dufferin	1:63360	1963
ON43	Halton	1:63360	1971
ON44	Waterloo	1:20000	1971
ON55	Brant	1:25000	1989
ON57	Haldimand-Norfolk	1:25000	1984

be corrected. It was assumed that newer maps and maps with finer scales were more accurate and therefore were used to fill gaps (by extending edge polygons), and to trim the surrounding map sheets.

There were a total of 723 unique soil types identified in the watershed. In addition to physical and chemical details, the associated soil database also contained information on soil type, number of layers, layer depths, and soil texture classifications.

28.3.2.3 Vegetation Data

Plant roots can have a significant impact on recharge as they remove infiltrated water from the soil. However, the determination of plant root penetration and subsequently the evaporative zone depth are very difficult tasks. In this study, the average evaporative zone depths (based on the combination of soils

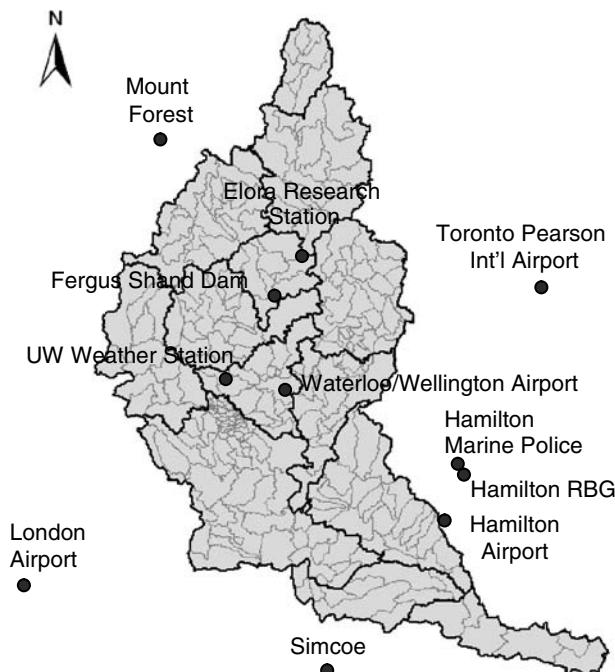


FIGURE 28.5 Zones of uniform meteorology (ZUM), sub-basins, and location of relative humidity and average wind speed observations.

and land cover types) were estimated using the guidelines given in the recharge methodology developed by the New Jersey Geological Survey (NJGS) (Charles et al., 1993).

HELP3 also requires a value for the maximum leaf area index (LAI) to calculate transpiration rates for the vegetative cover. Following the guidelines in HELP3, where the LAI ranges from 0 for bare ground to 5.0 for maximum vegetal leaf coverage, the values for LAI were assigned based on the LULC data. For example, the maximum LAI for bare agricultural fields was assumed to be 0, and for golf courses 2.0.

28.3.2.4 Weather Data

Due to its large size, the weather varies significantly across the Grand River watershed. Actual daily precipitation and temperature records from January 1960 to December 1999 were obtained from the Grand River Conservation Authority (GRCA) for the watershed. The data were based on point observations at various locations within (as well as outside) the watershed that were then used to represent the weather patterns within 13 subregions, or zones of uniform meteorology (ZUM) shown in Figure 28.5. The built-in weather generator in HELP3 was then used to generate daily synthetic solar radiation values for each ZUM as a function of latitude, precipitation, and temperature. A preliminary study revealed, however, that these subregions were still quite large resulting in discontinuities in recharge along their boundaries. Therefore, a new method based on an interpolation algorithm was developed.

The Grand River watershed was divided into 293 smaller sub-basins, also shown in Figure 28.5, each with its distinct values of daily precipitation, temperature, and solar radiation. The values were interpolated using the inverse distance squared (IDS) algorithm as

$$P_{\text{SUB}} = \frac{\sum_{i=1}^{13} P_i^{\text{ZUM}} / d_i^2}{\sum_{i=1}^{13} 1/d_i^2} \quad (28.3)$$

where P is the daily precipitation, temperature, or solar radiation, and d is the distance from the centroid of each ZUM to the centroid of each sub-basin. This provided a much smoother transition of weather data across the entire watershed, while still honouring the original observations.

The average quarterly relative humidities and average annual wind speeds were obtained from Environment Canada as well as the University of Waterloo weather station for various locations in and around the watershed, as shown in Figure 28.5. As before, the inverse distance squared weighting scheme was then used to estimate values for each of the 293 sub-basins.

The dates for the start and end of growing season were estimated using guidelines provided in HELP3 (Schroeder et al., 1994). The values were estimated for each ZUM, therefore, constant growing season dates were assumed for each sub-basin within each of the ZUM areas. Growing season starting dates varied from May 2nd to May 6th, while the ending dates ranged from October 7th to October 12th.

28.3.3 Methodology

As mentioned previously, the methodology was based on the work by Jyrkama et al. (2002), therefore, only a brief description is given in the following. Figure 28.6 illustrates a schematic diagram of the method. First, the LULC and soil maps were overlaid in the GIS to produce a combination map. These combinations were then used to determine vegetation data such as evaporative zone depths, as well as curve numbers used by HELP3 in the recharge analysis. HELP3 was then run daily from January 1960 to December 1999 for each of the unique combinations within each of the 293 sub-basin using the interpolated weather data for each sub-basin. This resulted in a total of approximately 47,000 unique combinations of LULC, soil,

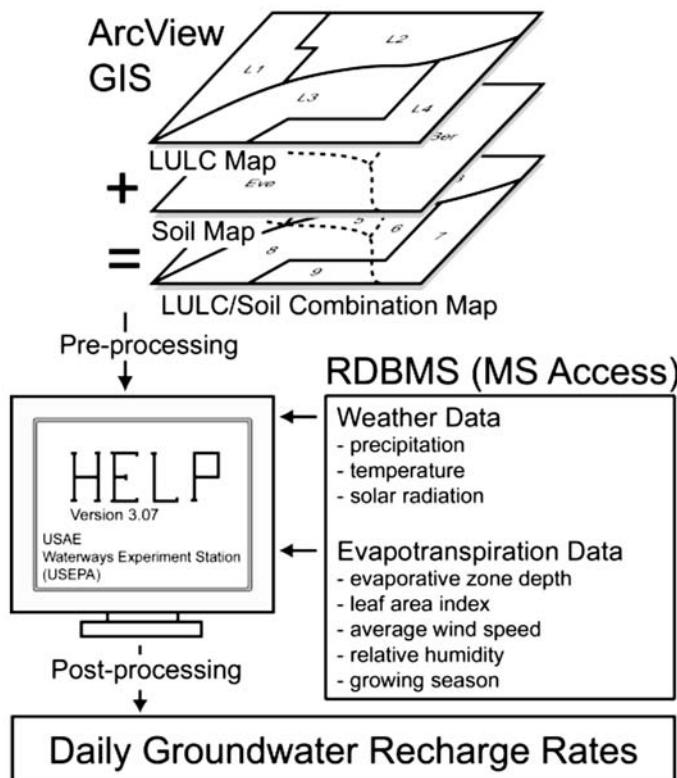


FIGURE 28.6 Methodology for estimating groundwater recharge. (After Jyrkama, M.I., Sykes, J.F., and Normani, S.D. 2002. *Ground Water*, 40, 638–648.)

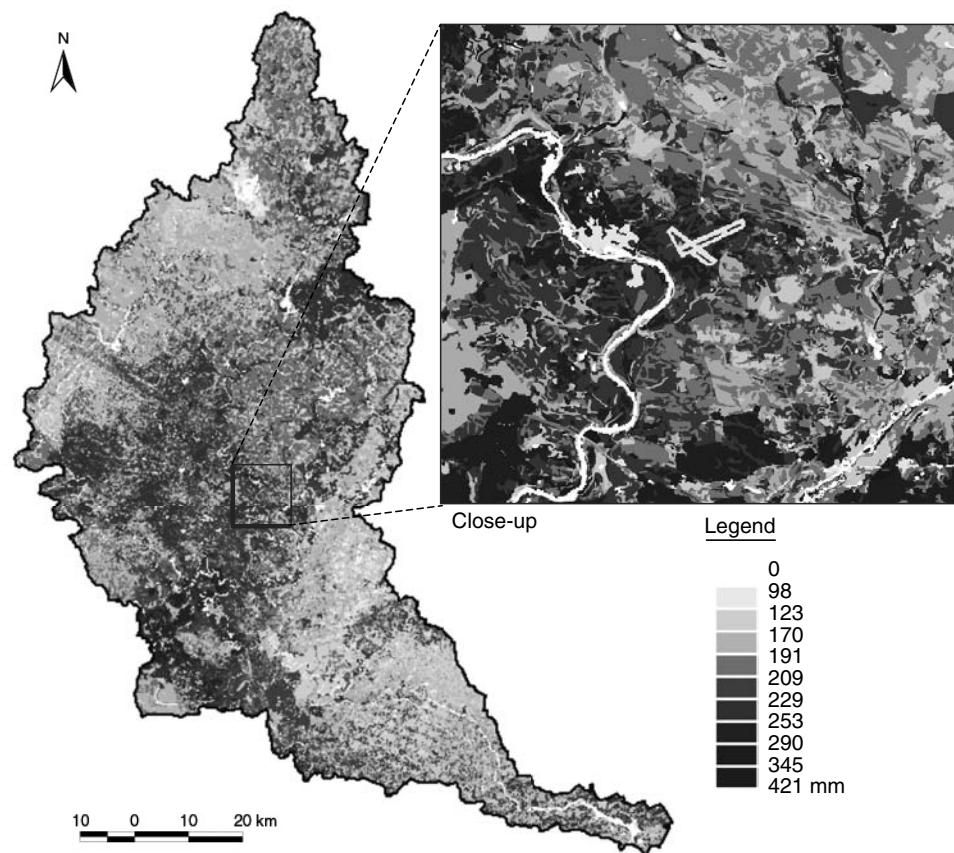


FIGURE 28.7 Average annual recharge for the Grand River watershed (mm/yr).

and weather data (i.e., ~47,000 HELP3 simulation runs). Areas classified as open water were ignored in the recharge analysis.

28.3.4 Results

Figure 28.7 shows the average annual recharge rates obtained from the HELP3 analysis for the Grand River watershed. Figure 28.8 illustrates the histogram of the same distribution. The average annual groundwater recharge in the watershed is approximately 200 mm/yr, which is approximately one-fifth of the average annual precipitation (950 mm/yr). As shown in Figure 28.7, recharge varies considerably across the watershed, responding directly to variations in land use and the hydraulic characteristics of the underlying soils.

As discussed previously, areas of high recharge may indicate areas where the underlying aquifers are subjected to increased vulnerability from contamination. This may have significant implications on land use planning near the urban areas, where existing lands are rapidly being converted into residential subdivisions and industrial areas.

The recharge rates also vary considerably over time, as shown by the monthly rates in Figure 28.9. Due to warmer temperatures in December 1992, most of the precipitation is able to infiltrate and become groundwater recharge, while during the much colder period in December 1998, most of the precipitation occurs as snow over frozen ground resulting in negligible recharge rates across the watershed.

As shown in Figure 28.9a, there is also considerable variation in recharge rates across the watershed. This variation is not only due to the variations in the land cover and soils, as discussed previously, but

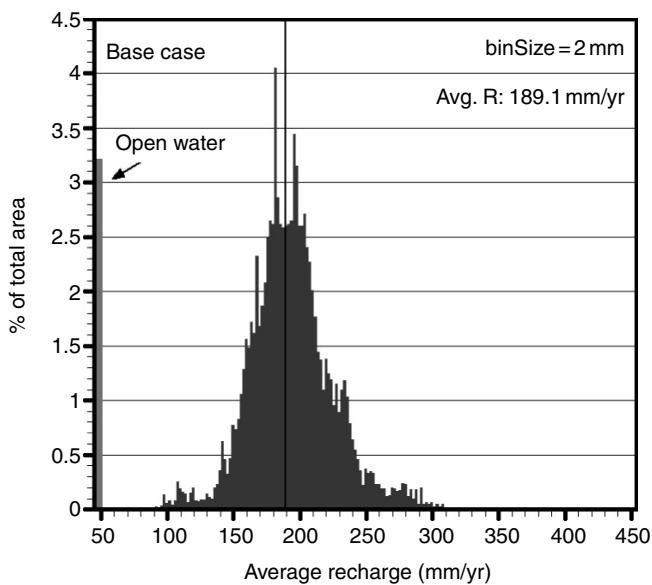


FIGURE 28.8 Recharge histogram.

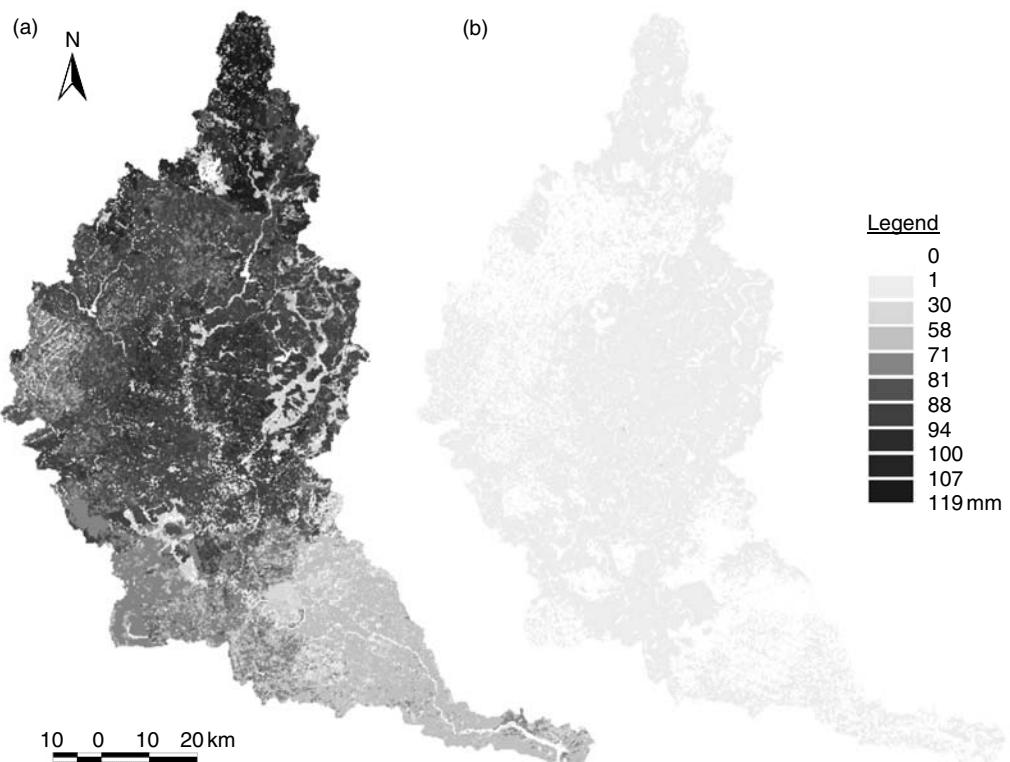


FIGURE 28.9 Monthly recharge rates for (a) December 1992 and (b) December 1998 (mm/month).

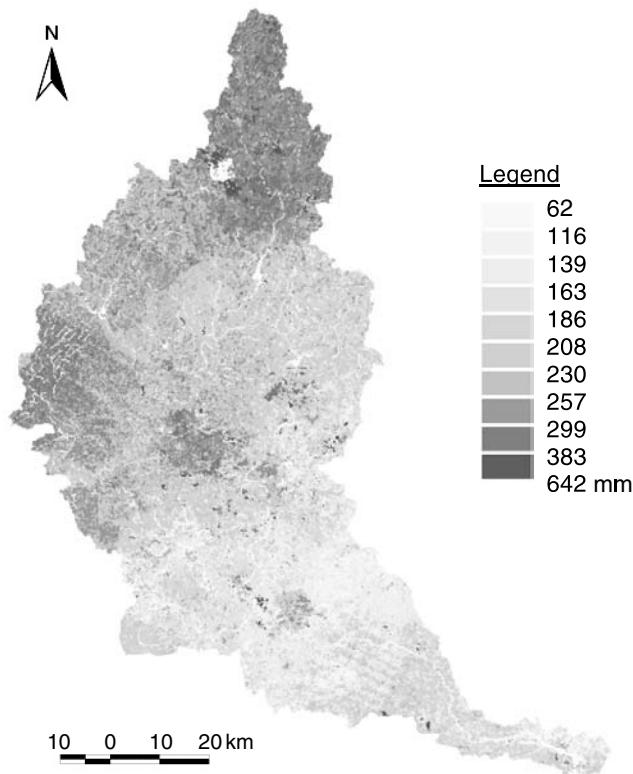


FIGURE 28.10 Average annual surface runoff (mm/yr).

also as a response to variations in climatic conditions across the watershed. Contrary to what might be expected, the northern portion of the watershed has considerably higher recharge rates in December 1992 than the southern end. This is explained by lower precipitation rates and unusually cool temperatures prevalent in the southern part of the watershed at the time.

The physically based recharge methodology also provides estimates of evapotranspiration and surface runoff, shown in Figure 28.10 and Figure 28.11, respectively, which can be used in an assessment of the watershed's overall water budget. As expected, both evapotranspiration and surface runoff vary considerably across the watershed, responding directly to variations in the land cover and soils. The northern portion of the watershed has significantly higher surface runoff rates, as expected, due to the presence of low permeability tills soils.

Approximately 210 mm of annual precipitation is subjected to overland flow while approximately 510 mm is evapotranspired by the vegetation in the watershed on an annual basis. Evapotranspiration, therefore, constitutes the largest portion of the watersheds' overall water budget.

As demonstrated by the above results, the methodology can potentially be used to assess the impact of climate change and changes in land use due to urbanization. The results further prove the applicability of the methodology across various scales, that is, from the small sub-basin level analysis to the regional watershed scale.

28.3.4.1 Impact of Climate Change

As discussed previously, climate change has the potential to significantly impact and alter the key processes in the Earth's water and energy cycles. Specific evidence from a wide variety of General Circulation (GC) and other climate models predict increases in the globally averaged surface air temperature, water vapor, evaporation, and precipitation. The intensity and frequency of extreme events is also projected to grow

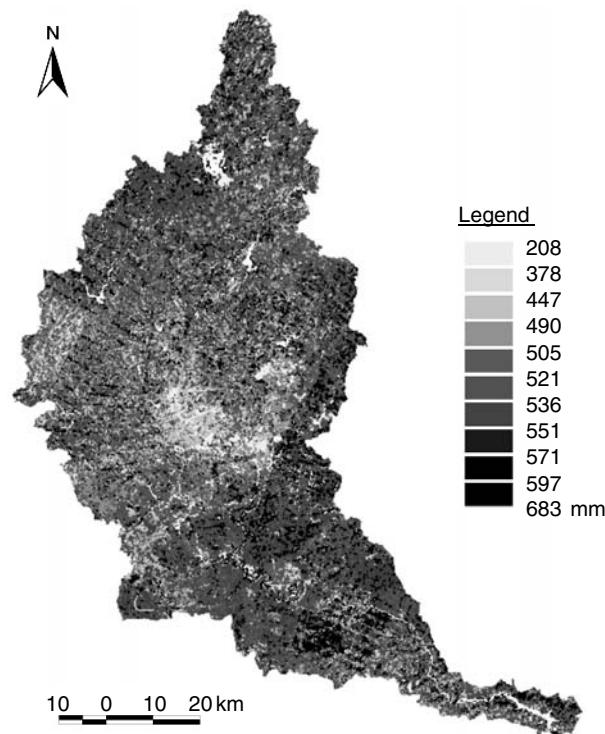


FIGURE 28.11 Average annual evapotranspiration (mm/yr).

(IPCC, 2001). Understanding the impact of these potential changes in the hydrologic cycle is essential for ensuring the quality and sustainability of water resources in the Grand River watershed.

The impact of climate change was modeled by perturbing the HELP3 model input parameters using predicted changes in the climate of the Grand River watershed. In addition to global changes, the IPCC reported the following general predictions for the regional climate in north-eastern North America over the next 100 years (IPCC, 2001):

- precipitation is projected to increase with an average change between 5 and 20% in the winter
- precipitation extremes are projected to increase more than the mean with higher intensities and higher frequency of extreme events
- greater than average warming in both summer and winter temperatures
- a possible reduction in incoming solar radiation due to increases in greenhouse gases.

Using the actual historical weather data as a reference for the Grand River study area, several scenarios were constructed to simulate the impact of climate change on the hydrologic cycle in the future. These scenarios are shown in Table 28.2.

All of the simulation parameters were scaled over the length of the study period. That is, they were assumed to increase linearly over time. For example, the temperature change of $+0.016^{\circ}\text{C}/\text{yr}$ corresponds to a predicted increase of 1.6°C in 100 years, or to a daily increase of approximately $4.4 \times 10^{-5}^{\circ}\text{C}$. The changes in precipitation intensity were simulated using

$$\Delta x = \frac{\sum_{i=1}^N P \cdot \delta}{\sum_{i=1}^N iP} \quad (28.4)$$

TABLE 28.2 Climate Change Simulation Scenarios

Case	Description
Base	Actual daily temperature, precipitation, and simulated solar radiation
1	Precipitation +5% for Dec, Jan, and Feb
2	Precipitation +20% for Dec, Jan, and Feb
3	Precipitation +20% for all months
4	Temperature +0.016°C/yr
5	Temperature +0.070°C/yr
6	Solar radiation 2% for all months
7	Combination of Cases 1, 5, and 6
8	Combination of Cases 3, 5, and 6

and

$$P_i^{cc} = P_i(1 + i\Delta x) \quad (28.5)$$

where P is the actual daily precipitation on day i , P^{cc} is the new daily precipitation due to climate change, Δx is the calculated change in daily precipitation, δ is the percent change in the average precipitation, and N is the total number of days in the study.

Figure 28.12 presents the cumulative differences in surface runoff, evapotranspiration, and recharge between all the cases and the Base Case scenario, for the Grand River watershed, averaged spatially over the study area. As shown, changing the precipitation (Case 1 to Case 3) has the highest influence on the hydrologic cycle, while solar radiation (Case 6) has a minimal impact. Groundwater recharge is predicted to increase under all scenarios, while evapotranspiration increases in all cases, except when incoming solar radiation is reduced (Case 6).

Figure 28.12a illustrates that, as expected, surface runoff increases with increasing precipitation. Furthermore, increasing the precipitation rate will generally increase all three hydrologic parameters as there is more water available in the system. Increasing temperature, however, has both a negative and positive impact on the hydrologic processes.

As demonstrated by Case 4 and Case 5 in Figure 28.12a for the Grand River watershed, temperature has a significant influence on the runoff process. The cumulative surface runoff decreases with increasing temperature mainly due to a reduced period of ground frost. Due to warmer winter temperatures, less water is stored in the snowpack and more water is able to infiltrate into the ground thereby reducing runoff and increasing groundwater recharge. This effect is less pronounced for more southerly study areas where the average winter temperatures are generally higher. At the watershed scale, climate change has the highest relative impact on groundwater recharge resulting in a 53% increase over 40 years, while surface runoff and evapotranspiration increase by 10% and 12%, respectively, over the same time period.

The temporal variabilities in the hydrologic processes are further demonstrated using the results from Case 8. Figure 28.13 shows the spatially averaged monthly differences for Case 8 and Base Case for the Grand River watershed. The increase in temperature and precipitation intensity over time is directly reflected in the hydrologic processes.

Figure 28.14 shows the same results for each month. As shown in Figure 28.14 for the Grand River watershed, it is evident that there is a significant reduction in the average runoff in the spring (e.g., April) as the springmelt is shifted into the winter months due to warmer temperatures. The amount of runoff is consequently increased during January, February, and March as moisture is released from the snowpack (as opposed to being stored or accumulated).

Groundwater recharge also increases significantly during the winter months as more water is able to infiltrate into the ground. Evapotranspiration rates are increased for the study area during the summer months due to higher temperatures and increased amount of available water.

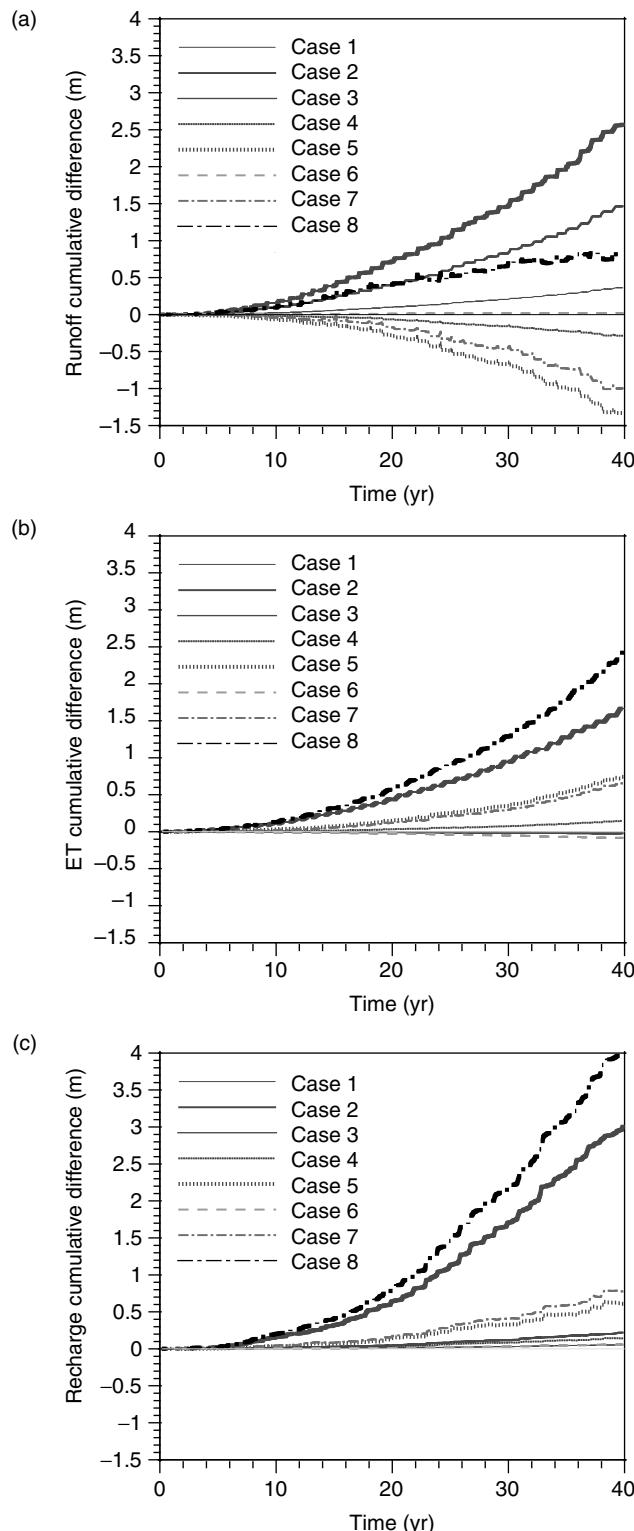


FIGURE 28.12 Cumulative differences between the climate change scenarios and the Base Case for (a) surface runoff, (b) evapotranspiration, and (c) groundwater recharge.

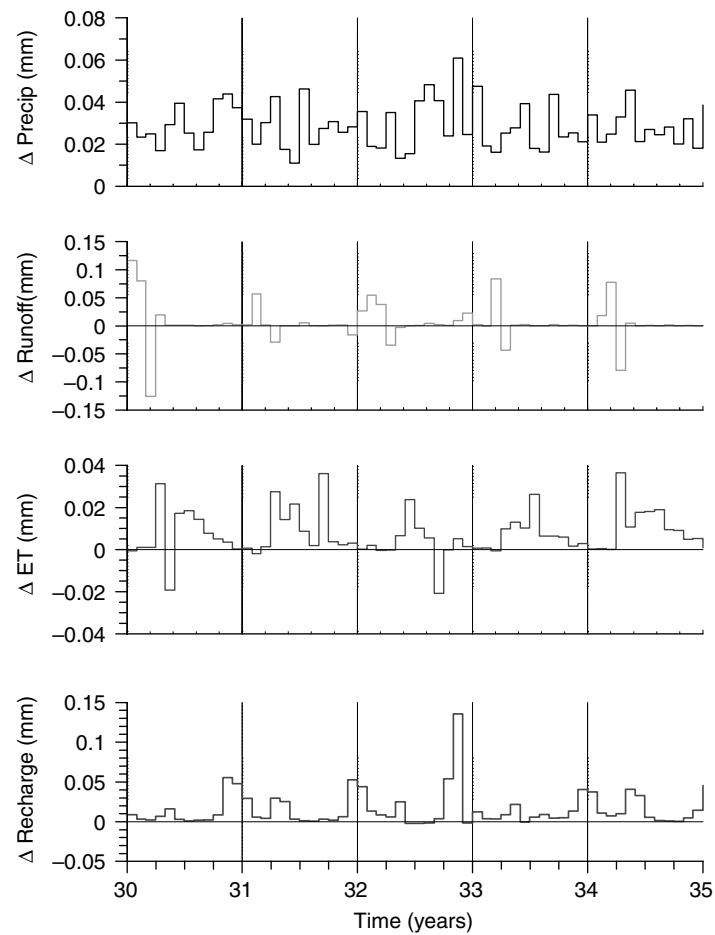


FIGURE 28.13 Monthly differences in precipitation, surface runoff, evapotranspiration, and groundwater recharge between Case 8 and the Base Case for a selected time period.

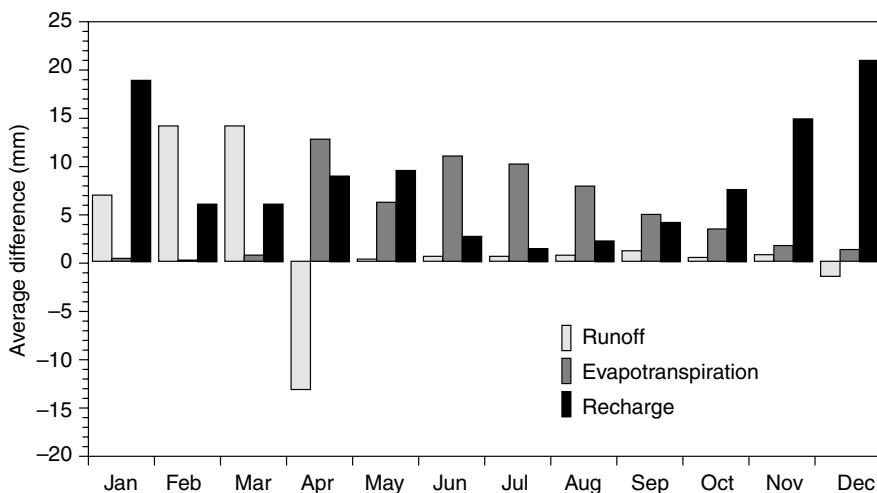


FIGURE 28.14 Average differences for each month between Case 8 and the Base Case.

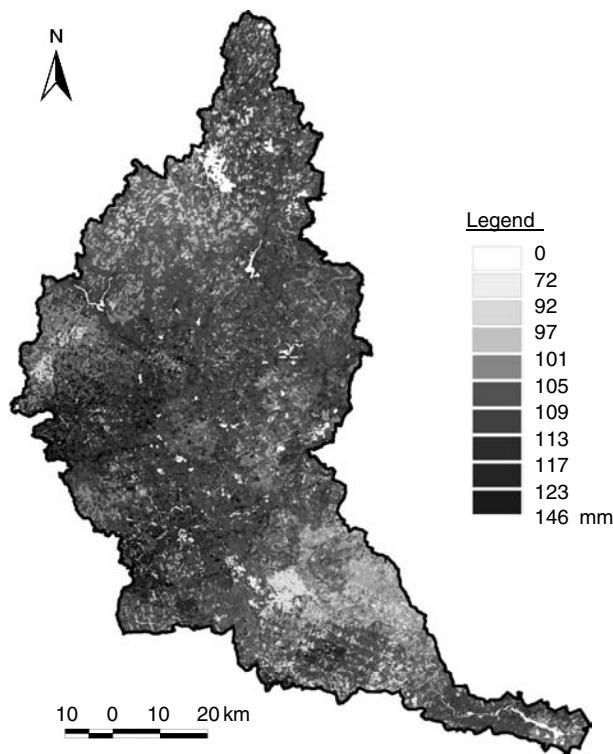


FIGURE 28.15 Average annual change in recharge between Case 8 and the Base Case.

Figure 28.15 shows the temporally averaged groundwater recharge rates for the Grand River watershed for the Base Case and Case 8. As illustrated, there is an overall increase in the recharge rate across the areas due to climate change. The average rate is predicted to increase by approximately 100 mm/yr from 189 mm/yr to 289 mm/yr over the 40-year study period for the Grand River watershed.

Unlike the significant reduction in recharge rates as predicted by Rosenberg et al. (1999) for the Ogallala aquifer in the United States, the results of this study predicted increased recharge rates for the study area due to climate change. The reason for the contrasting results is not only due to the temperature impacts in cooler climates, but also due to differences in the input data. Rosenberg et al. (1999) only included predicted changes in temperature and CO₂ concentrations in their simulations. As shown by the results for the Grand River watershed study area, increased precipitation has a significant influence on the estimated recharge rates.

Land use changes due to potential increases in urbanization, especially in the Grand River watershed, will also impact the hydrologic processes beyond the possible effects of climate change. The predicted increases in overall groundwater recharge due to climate change will, therefore, not necessarily guarantee sustainability in the future because human impacts, such as increased pumping, also exert a major influence in the aquifer system and can lead to depletion of the drinking water resources.

In conclusion, understanding the impact of potential changes in the hydrologic cycle due to climate change is essential for ensuring the quality and sustainability of our water resources in the future. The results of the study demonstrate the benefit of using the developed recharge methodology to address these issues.

28.3.5 Conclusions

Estimating groundwater recharge is critical for developing an effective watershed management strategy that will ensure the protection of the fresh water resources of the watershed. The physically based hydrologic

model HELP3 was integrated with ArcView GIS to estimate spatially distributed daily groundwater recharge rates for the Grand River watershed over a period of 40 years, from January 1960 to December 1999. The use of the GIS and a relational database management system (RDBMS) was essential due to the large volume of data required for numerical modeling of watershed processes at the regional scale.

The results of the study indicated that on average, only one-fifth of the overall precipitation in the watershed contributes to direct groundwater recharge, with an average recharge rate of approximately 200 mm/yr. In addition to quantifying the temporal distribution and amount of recharge, the results also identify areas of varying recharge potential. As the transport of most groundwater contaminants to saturated aquifers occurs in the aqueous phase as part of the recharge process, areas of high recharge may indicate areas where the underlying aquifers are subjected to increased vulnerability from surface contamination. This may have significant implications on land use planning, especially near urban areas. The developed recharge methodology also provides estimates of evapotranspiration and surface runoff, which may be used in estimating the watershed's overall water budget.

The results also demonstrate the applicability of the methodology across various scales. The method can easily be applied to hydrologic analyses, for example, due to climate change, at the regional watershed scale, or a detailed evaluation of the impact of land use changes at the much smaller sub-basin scale.

28.4 Chapter Summary

Climate change can directly affect the water balance of a surface water system; the impact on the groundwater system is indirect and is influenced through the spatially and temporally varying recharge and discharge boundary conditions that link the groundwater system to the surface water system. Recharge can occur as a result of precipitation events while both recharge and discharge can occur at surface water bodies. Assessing the impact of climate change on groundwater systems will generally require physically based models of these boundary conditions. Alternatively, fully integrated or coupled surface water and groundwater models can be used to assess climate change impact. Regardless of the modeling methodology used, it is essential that an assessment include the evaluation of historical conditions. For some groundwater systems, the spatial and temporal variation in historical groundwater recharge and discharge rates may be greater than the perturbation caused by climate change. In this chapter, the model HELP3 has been used as a simple tool to assess both the historical recharge and the impact of climate change for the Grand River Ontario watershed. For engineers and scientists, the value of the tool lies in the model's ready availability and its easy-to-use database. However, there are many other physically based models that are suitable for the assessment of the impact on groundwater of climate change. This chapter has reviewed many of these models and analyses.

References

- Abbott, M.B., Bathurst, J.C., Cunge, J.A., O'Connell, P.E., and Rasmussen, J. 1986. An introduction to the European Hydrological System — systeme hydrologique Europeen, “SHE”, 1: History and philosophy of a physically-based, distributed modelling system. *Journal of Hydrology*, 87, 45–59.
- Abbs, K. and Littleboy, M. 1998. Recharge estimation for the Liverpool Plains. *Australian Journal of Soil Research*, 36, 335–357.
- Allen, D.M., Mackie, D.C., and Wei, M. 2004. Groundwater and climate change: a sensitivity analysis for the Grand Forks aquifer, southern British Columbia, Canada. *Hydrogeology Journal*, 12, 270–290.
- Allison, G.B. 1988. A review of some of the physical, chemical and isotopic techniques available for estimating groundwater recharge. In: Simmers, I. (ed.) *Estimation of natural groundwater recharge*. NATO ASI Series C, vol. 222 (Proceedings of the NATO advanced research workshop, Antalya, Turkey, March 1987), D. Reidel Publication Company, Dordrecht, pp. 49–72.

- Arnold, J.G., Srinivasan, R., Muttiah, R.S., and Williams, J.R. 1998. Large area hydrologic modeling and assessment, Part I. Model development. *Journal of the American Water Resources Association*, 34, 73–89.
- Arnold, J.G., Muttiah, R.S., Srinivasan, R., and Allen, P.M. 2000. Regional estimation of base flow and groundwater recharge in the Upper Mississippi river basin. *Journal of Hydrology*, 227, 21–40.
- Ball, J.E. and Luk, K.C. 1998. Modeling spatial variability of rainfall over a catchment. *Journal of Hydrologic Engineering*, 3, 122–130.
- Batelaan, O. and De Smedt, F. 2001. WetSpass: a flexible, GIS based, distributed recharge methodology for regional groundwater modelling. In: Gehrels, H., N.E. Peters, E. Hoehn, K. Jensen, C. Leibundgut, J. Griffioen, B. Webb, and W.J. Zaadnoordijk. (eds.) *Proceedings of the Conference on Impact of Human Activity on Groundwater Dynamics*. IAHS Publication No. 269, pp. 11–17.
- Batelaan, O., De Smedt, F., and Triest, L. 2003. Regional groundwater discharge: phreatophyte mapping, groundwater modelling and impact analysis of land-use change. *Journal of Hydrology*, 275, 86–108.
- Bauer, H.H. and Vaccaro, J.J. 1987. *Documentation of a deep percolation model for estimating groundwater recharge*. U.S. Geological Survey, Open-File Report 86-536, Tacoma, WA.
- Bear, J. 1972. *Dynamics of Fluids in Porous Media*. Elsevier, New York.
- Beckers, J. and Frind, E.O. 2000. Simulating groundwater flow and runoff for the Oro Moraine aquifer system. Part 1. Model formulation and conceptual analysis. *Journal of Hydrology*, 229, 265–280.
- Bekesi, G. and McConchie, J. 1999. Groundwater recharge modelling using the Monte Carlo technique, Manawatu region, New Zealand. *Journal of Hydrology*, 224, 137–148.
- Bellot, J., Bonet, A., Sanchez, J.R., and Chirino, E. 2001. Likely effects of land use changes on the runoff and aquifer recharge in a semiarid landscape using a hydrological model. *Landscape and Urban Planning*, 55, 41–53.
- Beverly, C.R., Nathan, R.J., Malafant, K.W.J., and Fordham, D.P. 1999. Development of a simplified unsaturated module for providing recharge estimates to saturated groundwater models. *Hydrological Processes*, 13, 653–675.
- Black, P.B. and Miller, R.D. 1990. Hydraulic conductivity and unfrozen water content of air-free frozen silt. *Water Resources Research*, 26, 323–329.
- Bobba, A.G., Singh, V.P., Jeffries, D.S., and Bengtsson, L. 1997. Application of a watershed runoff model to north-east Pond River, Newfoundland: To study water balance and hydrological characteristics owing to atmospheric change. *Hydrological Processes*, 11, 1573–1593.
- Bonomi, T. and Cavallin, A. 1999. Three-dimensional hydrogeological modelling application to the Alvera mudslide (Cortina d'Ampezzo, Italy). *Geomorphology*, 30, 189–199.
- Bouraoui, F. and Dillaha, T.A. 1996. ANSWERS-2000: Runoff and sediment transport model. *Journal of Environmental Engineering*, ASCE, 122, 493–502.
- Bouraoui, F., Vachaud, G., Li, H., Le Treut, and Chen, T. 1999. Evaluation of the impact of climate changes on water storage and groundwater recharge at the watershed scale. *Climate Dynamics*, 15, 153–161.
- Brodie, R.S. 1999. Integrating GIS and RDBMS technologies during construction of a regional groundwater model. *Environmental Modelling & Software*, 14, 119–128.
- Brouyere, S., Carabin, G., and Dassargues, A. 2004. Climate change impacts on groundwater resources: Modelled deficits in a chalky aquifer, Geer Basin, Belgium. *Hydrogeology Journal*, 12, 123–134.
- Burges, S.J., Wigmosta, M.S., and Meena, J.M. 1998. Hydrological effects of land-use change in a zero-order catchment. *Journal of Hydrologic Engineering*, 3, 86–97.
- Burgess, S.S.O., Adams, M.A., Turner, N.C., and Ong, C.K. 1998. The redistribution of soil water by tree root systems. *Oecologica*, 115, 306–311.
- Burgess, S.S.O., Adams, M.A., Turner, N.C., White, D.A., and Ong, C.K. 2001. Tree roots: conduits for deep recharge of soil water. *Oecologia*, 126, 158–165.
- Canadell, J., Jackson, R.B., Ehleringer, J.R., Mooney, H.A., Sala, O.E., and Schulze, E.-D. 1996. Maximum rooting depth of vegetation types at the global scale, *Oecologia*, 108, 583–595.

- Carmon, N., Shamir, U., and Meiron-Pistiner, S. 1997. Water-sensitive urban planning: protecting groundwater. *Journal of Environmental Planning and Management*, 40, 413–434.
- Chapman, T. 1999. A comparison of algorithms for stream flow recession and baseflow separation. *Hydrological Processes*, 13, 701–714.
- Chapman, T.G. and Malone, R.W. 2002. Comparison of models for estimation of groundwater recharge, using data from a deep weighing lysimeter. *Mathematics and Computers in Simulation*, 59, 3–17.
- Charles, E.G., Behroozi, J., Schooley, J., and Hoffman, J.L. 1993. *A method for evaluating ground water recharge areas in New Jersey*. NJ Geological Survey Report GSR-32.
- Chen, Z., Grasby, S.E., and Osadetz, K.G. 2002. Predicting average annual groundwater levels from climatic variables: and empirical model. *Journal of Hydrology*, 260, 102–117.
- Chiew, F.H.S., McMahon, T.A., and O'Neill, I.C. 1992. Estimating groundwater recharge using an integrated surface and groundwater modelling approach. *Journal of Hydrology*, 131, 151–186.
- Collin, M.L. and Mellou, A.J. 2001. Combined land-use and environmental factors for sustainable groundwater management (Case Study). *Urban Water*, 3, 253–261.
- Colosimo, C. and Mendicino, G. 1996. GIS for distributed rainfall-runoff modeling. In: Singh, V.P. and M. Fiorentino. (eds.) *Geographical Information Systems in Hydrology*. Kluwer Academic Publishers, The Netherlands, pp. 195–235.
- Cook, P.G. and Kilty, S. 1992. A helicopter-borne electromagnetic survey to delineate groundwater recharge rates. *Water Resources Research*, 28, 2953–2961.
- Cooper, D.M., Wilkinson, W.B., and Arnell, N.W. 1995. The effects of climate changes on aquifer storage and river baseflow. *Hydrological Sciences Journal/Journal des Sciences Hydrologiques*, 40, 615–632.
- Coppola Jr., E.A., Duckstein, L., and Davis, D. 2002. Fuzzy rule-based methodology for estimating monthly groundwater recharge in a temperate watershed. *Journal of Hydrologic Engineering*, 7, 326–335.
- Croley II, T.E. and Luukkonen, C.L. 2003. Potential effects of climate change on ground water in Lansing, Michigan. *Journal of the American Water Resources Association*, 39, 149–163.
- Daniel, J.A. and Staricka, J.A. 2000. Frozen soil impact on ground water-surface water interaction. *Journal of the American Water Resources Association*, 36, 151–160.
- Dawes, W.R. and Hatton, T.J. 1993. *TOPOG_IRM, 1. Model Description*. Technical Memorandum 93(5), CSIRO Division of Water Resources, Canberra.
- Deng, Y., Flerchinger, G.N., and Cooley, K.R. 1994. Impacts of spatially and temporally varying snowmelt on subsurface flow in a mountainous watershed: 2. Subsurface processes. *Hydrological Sciences Journal/Journal des Sciences Hydrologiques*, 39, 521–534.
- Dingman, S.L. 1994. *Physical Hydrology*. Prentice Hall, Engelwood Cliffs, NJ.
- Eckhardt, K. and Ulbrich, U. 2003. Potential impacts of climate change on groundwater recharge and streamflow in a central European low mountain range. *Journal of Hydrology*, 284, 244–252.
- El-Kadi, A.I. 1989. Watershed models and their applicability to conjunctive use management. *Water Resources Bulletin*, 25, 125–137.
- Ella, V.B., Melvin, S.W., Kanwar, R.S., Jones, L.C., and Horton, R. 2002. Inverse three-dimensional groundwater modeling using the finite-difference method for recharge estimation in a glacial till aquitard. *Transactions of the ASAE*, 45, 703–715.
- Engelmark, H. 1988. Rates of infiltration into frozen and unfrozen fine sand. *Canadian Journal of Earth Sciences*, 25, 343–347.
- Erskine, A.D. and Papaioannou, A. 1997. The use of aquifer response rate in the assessment of groundwater resources. *Journal of Hydrology*, 202, 373–391.
- Fayer, M.J. and Jones, T.L. 1990. *UNSAT-H Version 2.0: Unsaturated Soil Water and Heat Flow Model*. PNL-6779, Pacific Northwest National Laboratory, Richland, WA.
- Fayer, M.J., Gee, G.W., Rockhold, M.L., Freshley, M.D., and Walters, T.B. 1996. Estimating recharge rates for a groundwater model using a GIS. *Journal of Environmental Quality*, 25, 510–518.

- Finch, J.W. 1998. Estimating direct groundwater recharge using a simple water balance model — sensitivity to land surface parameters. *Journal of Hydrology*, 211, 112–125.
- Finch, J.W. 2001. Estimating change in direct groundwater recharge using a spatially distributed soil water balance model. *Quarterly Journal of Engineering Geology and Hydrogeology*, 34, 71–83.
- Flint, A.L., Flint, L.E., Kwicklis, E.M., Fabryka-Martin, J.T., and Bodvarsson, G.S. 2002. Estimating recharge at Yucca Mountain, Nevada, USA: comparison of methods. *Hydrogeology Journal*, 10, 180–204.
- Foster, S.S.D. 1998. Groundwater recharge and pollution vulnerability of British aquifers: a critical overview. In: Robins, N.S. (ed.) *Groundwater Pollution, Aquifer Recharge and Vulnerability*. Geological Society, London, Special Publications, 130, pp. 7–22.
- Foster, S.S.D. 2001. The interdependence of groundwater and urbanisation in rapidly developing cities (Review). *Urban Water*, 3, 185–192.
- Freeze, R.A. 1971. Three-dimensional transient saturated-unsaturated flow in a groundwater basin. *Water Resources Research*, 7, 347–366.
- Freeze, R.A. and Banner, J. 1970. The mechanism of natural groundwater recharge and discharge; 2. Laboratory column experiments and field measurements. *Water Resources Research*, 6, 138–155.
- Freeze, R.A. and Harlan, R.L. 1969. Blueprint for a physically-based, digitally-simulated hydrologic response model. *Journal of Hydrology*, 9, 237–258.
- Freeze, R.A. and Witherspoon, P.A. 1968. Theoretical analysis of regional groundwater flow; 3. Quantitative interpretations. *Water Resources Research*, 4, 581–590.
- Futter, M. 1995. Integrated water resources management. *World Water and Environmental Engineering*, 18, 24.
- Gau, H.S. and Liu, C.W. 2000. Estimation of the effective precipitation recharge coefficient in an unconfined aquifer using stochastic analysis. *Hydrological Processes*, 14, 811–830.
- Gerber, R.E. and Howard, K. 2000. Recharge through a regional till aquitard: Three-dimensional flow model water balance approach. *Ground Water*, 38, 410–422.
- Gleick, P.H. 1989. Climate change, hydrology, and water resources. *Reviews of Geophysics*, 27, 329–344.
- Goddard, L., Mason, S.J., Zebiak, S.E., Ropelewski, C.F., Basher, R., and Cane, M.A. 2001. Current approaches to seasonal-to-interannual climate predictions. *International Journal of Climatology*, 21, 1111–1152.
- Graham, D.N. and Refsgaard, A. 2001. MIKE SHE: A distributed, physically based modeling system for surface water/groundwater interactions. In: Seo, H.S., E. Poeter, C. Zheng, and O. Poeter. (eds.) *Proceedings of Modflow 2001 and Other Modelling Odysseys*. International Ground Water Modeling Center (IGWMC), Colorado School of Mines and the U.S. Geological Survey, pp. 321–327.
- Granger, R.J., Gray, D.M., and Dyck, G.E. 1984. Snowmelt infiltration to frozen prairie soils. *Canadian Journal of Earth Sciences*, 21, 669–677.
- Green, W.H. and Ampt, G.A. 1911. Studies in soil physics, I. The flow of air and water through soils. *Journal of Agricultural Science*, 4, 1–24.
- Grindley, J. 1969. The calculation of evaporation and soil moisture deficit over specified catchment area. *Hydrological Memorandum 28*, Meteorological Office, Bracknell, U.K., pp. 10.
- Gureghian, A.B., DeWispelare, A.R., and Sagar, B. 1994. Sensitivity and probabilistic analyses of the impact of climatic conditions on the infiltration rate in a variably saturated multilayered geologic medium. In: Thomas L. Sanders (ed.) *Proceedings of the Fifth International Conference on High Level Radioactive Waste Management*. American Nuclear Society, 1622–1633.
- Habib, E., Krajewski, W.F., and Kruger, A. 2001. Sampling errors of tipping-bucket rain gauge measurements. *Journal of Hydrologic Engineering*, 6, 159–166.
- Halford, K.J. and Mayer, G.C. 2000. Problems associated with estimating ground water discharge and recharge from stream-discharge records. *Ground Water*, 38, 331–342.
- Hall, M.J. 1984. *Urban Hydrology*. Elsevier Applied Science, London.
- Harms, T.E. and Chanasyk, D.S. 1998. Variability of snowmelt runoff and soil moisture recharge. *Nordic Hydrology*, 29, 179–198.

- Holder, M.K., Brown, W., Thomas, J.C., Zabcik, D., and Murray, H.E. 1991. Capillary-wick unsaturated zone soil pore water sampler. *Soil Science Society of America Journal*, 55, 1195–1202.
- Holysh, S., Pitcher, J., and Boyd, D. 2000. *Regional Groundwater Mapping: An Assessment Tool for Incorporating Groundwater into the Planning Process*. Grand River Conservation Authority Report.
- Horton, R.E. 1940. An approach toward a physical interpretation of infiltration capacity. *Soil Science Society of America Proceedings*, 5, 399–417.
- Hough, M.N. and Jones, R.J.A. 1998. The United Kingdom Meteorological Office rainfall and evaporation calculation system: MORECS version 2.0 — an overview. *Hydrology and Earth System Science*, 1, 227–239.
- Huyakorn, P.S., Springer, E.P., Guvanasen, V., and Wadsworth, T.D. 1986. A three-dimensional finite-element model for simulating water flow in variably-saturated porous media. *Water Resources Research*, 22, 1790–1808.
- HydroGeoLogic, Inc. 2002. MODHMS — A MODFLOW based fully integrated Hydrologic-Water Quality numerical modeling system.
- Intergovernmental Panel on Climate Change (IPCC). 2001. *Climate Change 2001: The Scientific Basis*. Cambridge University Press, Cambridge, U.K.
- Jackson, R.B., Canadell, J., Ehleringer, J.R., Mooney, H.A., Sala, O.E., and Schulze, E-D. 1996. A global analysis of root distributions for terrestrial biomes. *Oecologia*, 108, 389–411.
- Javed, I. and Bonnell, R.B. 1999. Groundwater management in irrigated agriculture of Pakistan by inverse modeling: Model application. *Canadian Water Resources Journal/Revue Canadienne des Ressources Hydriques*, 24, 293–306.
- Johnsson, H. and Lundin, L. 1991. Surface runoff and soil water percolation as affected by snow and soil frost. *Journal of Hydrology*, 122, 141–159.
- Jyrkama, M.I. 1999. *Modelling Non-Equilibrium Dissolution and Multiphase Contaminant Transport in Freezing Porous Media*. MASc Thesis, Department of Civil Engineering, University of Waterloo, Waterloo, Ontario, Canada.
- Jyrkama, M.I., Sykes, J.F., and Normani, S.D. 2002. Recharge estimation for transient ground water modeling. *Ground Water*, 40, 638–648.
- Kane, D.L. and Stein, J. 1983. Water movement into seasonally frozen soils. *Water Resources Research*, 19, 1547–1557.
- Kennett-Smith, A., Narayan, K., and Walker, G. 1996. *Calibration of a groundwater model for the upper south east of South Australia*. Divisional Report 96(2), CSIRO Division of Water Resources, Canberra.
- Kirkham, R.R., Gee, G.W., and Jones, T.L. 1984. Weighing lysimeters for long-term water balance investigations at remote sites. *Soil Science Society of America Journal*, 48, 1203–1205.
- Kirshen, P.H. 2002. Potential impacts of global warming on groundwater in eastern Massachusetts. *Journal of Water Resources Planning and Management*, 128, 216–226.
- Knutsson, G. 1988. Humid and arid zone groundwater recharge — a comparative analysis. In: Simmers, I. (ed.) *Estimation of natural groundwater recharge*. NATO ASI Series C, Vol. 222 (Proceedings of the NATO advanced research workshop, Antalya, Turkey, March 1987), D. Reidel Publication Company, Dordrecht, 493–504.
- Kouwen, N. and Garland, G. 1989. Resolution considerations in using radar rainfall data for flood forecasting. *Canadian Journal of Civil Engineering*, 16, 279–289.
- Kung, K.-J.S. 1990. Preferential flow in a sandy vadose zone, 2, Mechanisms and implications. *Geoderma*, 46, 51–58.
- LaBolle, E.M., Ahmed, A.A., and Fogg, G.E. 2003. Review of the Integrated Groundwater and Surface-Water Model (IGSM). *Ground Water*, 41, 238–246.
- Larcher, W. 1983. *Physiological Plant Ecology*. Springer-Verlag, Berlin.
- Le Maitre, D.C., Scott, D.F., and Colvin, C. 1999. A review of information on interactions between vegetation and groundwater. *Water South Africa*, 25, 137–152.
- Lerner, D.N. 2002. Identifying and quantifying urban recharge: A review. *Hydrogeology Journal*, 10, 143–152.

- Lerner, D., Issar, A.S., and Simmers, I. (eds.) 1990. *Groundwater recharge: a guide to understanding and estimating natural recharge*. Heise, Hannover.
- Levine, J.B. and Salvucci, G.D. 1999. Equilibrium analysis of groundwater-vadose zone interactions and the resulting spatial distribution of hydrologic fluxes across a Canadian prairie. *Water Resources Research*, 35, 1369–1383.
- Lin, Y.-F. and Anderson, M.P. 2003. A digital procedure for ground water recharge and discharge pattern recognition and rate estimation. *Ground Water*, 41, 306–315.
- Lloyd, J.W. 1986. A review of aridity and groundwater. *Hydrological Processes*, 1, 63–78.
- Loaiciga, H.A. 2003. Climate change and ground water. *Annals of the Association of American Geographers*, 93, 30–41.
- Louie, M.J., Shelby, P.M., Smesrud, J.S., Gatchell, L.O., and Selker, J.S. 2000. Field evaluation of passive capillary samplers for estimating groundwater recharge. *Water Resources Research*, 36, 2407–2416.
- Lu, Z. and Danskin, W.R. 2001. InSAR analysis of natural recharge to define structure of a ground-water basin, San Bernardino, California. *Geophysical Research Letters*, 28, 2661–2664.
- Martin, P.J. and Frind, E.O. 1998. Modeling a complex multi-aquifer system: the Waterloo moraine. *Ground Water*, 36, 679–690.
- McDonald, M.D. and Harbaugh, A.W. 1996. *A Modular Three-Dimensional Finite Difference Groundwater Flow Model*. United States Geological Survey.
- McGuffie, K. and Henderson-Sellers, A. 2001. Forty Years of Numerical Climate Modelling. *International Journal of Climatology*, 21, 1067–1109.
- Molson, J.W., Frind, E.O., and Palmer, C.D. 1992. Thermal energy storage in an unconfined aquifer. 2. Model development, validation and application. *Water Resources Research*, 28, 2857–2867.
- Monteith, J.L. 1965. Evaporation and environment. *Symposium of the Society for Experimental Biology*, 19, 205–214.
- Montgomery Watson, Inc. 1993. *Integrated Groundwater and Surface Water Model (IGSW) Documentation and User Manual*. Montgomery Watson, Inc., Broomfield, Co.
- Mullins, J.A., Carsel, R.F., Sarbrough, J.E., and Ivery, A.M. 1993. *PRZM-2, A Model for Predicting Pesticide Fate in the Crop Root and Unsaturated Soil Zones, User's Manual for Release Version 2.0*. U.S. Environmental Protection Agency, Environmental Research Lab, Athens, GA.
- Natural Resources Conservation Service. 1986. *Urban Hydrology for Small Watersheds*. Technical Report 55, United States Department of Agriculture.
- Nichols, D.S. and Verry, E.S. 2001. Stream flow and ground water recharge from small forested watershed in north central Minnesota. *Journal of Hydrology*, 245, 89–103.
- Pearcy, R.W., Ehleringer, J.R., Mooneyand, H.A., and Rundel, P.W. (eds.) 1989. *Plant Physiological Ecology*. Chapman and Hall, London.
- Peck, A.J. 1978. Salinization of non-irrigated soils: a review. *Australian Journal of Soil Research*, 16, 157–168.
- Penman, H.L. 1948. Natural evaporation from open water, bare soil and grass. *Proceedings of the Royal Society of London, Serial A*, 193, 120–145.
- Philip, J.R. 1957. The theory of infiltration. 1. The infiltration equation and its solution. *Soil Science*, 83, 345.
- Pitman, A.J. 2003. The evolution of, and revolution in, land surface schemes designed for climate models. *International Journal of Climatology*, 23, 479–510.
- Prathapar, S.A., Meyer, W.S., Jain, S., and van der Leij, A. 1994. *SWAGSIM: A soil water and groundwater simulation model*. Divisional Report 94(3), CSIRO Division of Water Resources, Canberra, 1–38.
- Priestley, C.H.B. and Taylor, R.J. 1972. On the assessment of surface heat flux and evaporation using large scale parameters. *Monthly Weather Review*, 100, 81–92.
- Querner, E.P. 1997. Description and application of the combined surface and groundwater flow model MOGROW. *Journal of Hydrology*, 192, 158–188.
- Querner, E.P. 2000. The Effects of Human Intervention in the Water Regime. *Ground Water*, 38, 167–171.

- Reddi, L.N. and Danda, S.K.R. 1994. Unsaturated flow modeling — Exact solution to approximate problem? *Journal of Water Resources Planning and Management*, 120, 186–198.
- Richards, L.A. 1931. Capillary conduction of liquids in porous mediums. *Physics*, 1, 318.
- Richardson, C.W. and Wright, D.A. 1984. *WGEN: A model for generating daily weather variables*. ARS-8. Agricultural Research Service, US Department of Agriculture.
- Robins, N.S. (ed.) 1998. *Ground-water Pollution, Aquifer Recharge and Vulnerability*. Geological Society, London, Special Publications, 130, 107–115.
- Rorabaugh, M.I. 1964. Estimating changes in bank storage and groundwater contribution to streamflow. *International Association of Scientific Hydrology Publication*, 63, 432–441.
- Rose, S. and Peters, N.E. 2001. Effects of urbanization on streamflow in the Atlanta area (Georgia, USA): a comparative hydrological approach. *Hydrological Processes*, 15, 1441–1457.
- Rosenberg, N.J., Epstein, D.J., Wang, D., Vail, L., Srinivasan, R., and Arnold, J.G. 1999. Possible impacts of global warming on the hydrology of the Ogallala Aquifer Region. *Climatic Change*, 42, 677–692.
- Rushton, K.R. 1988. Numerical and conceptual models for recharge estimation in arid and semi-arid zones. In: Simmers, I. (ed.) *Estimation of natural groundwater recharge*. NATO ASI Series C, vol. 222 (Proceedings of the NATO advanced research workshop, Antalya, Turkey, March 1987), D. Reidel Publication Company, Dordrecht, pp. 223–238.
- Ruud, N., Harter, T., and Naugle, A. 2001. A conjunctive use model for the Tule River groundwater basin in the San Joaquin Valley, California. In: M.A. Marino and S.P. Simonovic. (eds.) *Proceedings of the Conference on Integrated Water Resources Management*. IAHS Publication No. 272, pp. 167–174.
- Salama, R. B., Tapley, I., Ishii, T., and Hawkes, G. 1994. Identification of areas of recharge and discharge using Landsat-TM satellite imagery and aerial photography mapping techniques. *Journal of Hydrology*, 162, 119–141.
- Sandstrom, K. 1995. Modeling the effects of rainfall variability on groundwater recharge in semi-arid Tanzania. *Nordic Hydrology*, 26, 313–330.
- Scanlon, B.R., Healy, R.W., and Cook, P.G. 2002. Choosing appropriate techniques for quantifying groundwater recharge. *Hydrogeology Journal*, 10, 18–39.
- Schroeder, P.R., Dozier, T.S., Zappi, P.A., McEnroe, B.M., Sjostrom, J.W., and Peyton, R.L. 1994. *The Hydrologic Evaluation of Landfill Performance (HELP) Model: Engineering Documentation for Version 3*. EPA/600/R-94/168b, U.S. Environmental Protection Agency Office of Research and Development, Washington, DC.
- Schuh, W.M., Meyer, R.F., Sweeney, M.D., and Gardner, J.C. 1993. Spatial variation of root-zone and shallow dose-zone drainage on a loamy glacial till in a sub-humid climate. *Journal of Hydrology*, 148, 27–60.
- Schulze, E.D., Caldwell, M.M., Canadell, J., Mooney, H.A., Jackson, R.B., Parson, D., Scholes, R., Sala, O.E., and Trimborn, P. 1998. Downward flux of water through roots (i.e. inverse hydraulic lift) in dry Kalahari sands. *Oecologia*, 115, 460–462.
- Selker J.S., Parlange, J.-Y., and Steenhuis, T. 1992. Fingered flow in two dimensions, 2, Predicting finger moisture profile. *Water Resources Research*, 28, 2523–2528.
- Shade, P.J. and Nichols, W.D. 1996. *Water budget and the effects of land-use changes on ground-water recharge, Oahu, Hawaii*. USGS Professional Paper 1412-C.
- Sharma, M.L. (ed.) 1989. *Groundwater Recharge*. Proceedings of the Symposium on Groundwater Recharge, Mandurah, Australia, 6–9 July 1987, A.A. Balkema, Rotterdam.
- Sililo, O.T.N. and Tellam, J.H. 2000. Fingered flow in unsaturated zone flow: a qualitative review with laboratory experiments on heterogeneous systems. *Ground Water*, 38, 864–871.
- Simmers, I. 1990. Part I: Aridity, groundwater recharge and water resources management. In: Lerner, D., A.S. Issar, and I. Simmers. (eds.). *Groundwater recharge: a guide to understanding and estimating natural recharge*. Heise, Hannover.
- Simmers, I. 1998. Groundwater recharge: an overview of estimation ‘problems’ and recent developments. In: Robins, N.S. (ed.) *Ground-water Pollution, Aquifer Recharge and Vulnerability*. Geological Society, London, Special Publications, 130, 107–115.

- Singh, V.P. 1997. Effect of spatial and temporal variability in rainfall and watershed characteristics on stream flow hydrograph. *Hydrological Processes*, 11, 1649–1669.
- Singh, V.P. and Woolhiser, D.A. 2002. Mathematical modeling of watershed hydrology. *Journal of Hydrologic Engineering*, 7: 270–292.
- Sinha, B.P.C. and Sharma, S.K. 1988. Natural groundwater recharge estimation methodologies in India. In: Simmers, I. (ed.) *Estimation of natural groundwater recharge*. NATO ASI Series C, vol. 222 (Proceedings of the NATO advanced research workshop, Antalya, Turkey, March 1987), D. Reidel Publication Company, Dordrecht, 301–311, 1988.
- Sophocleous, M. and Perkins, S.P. 2000. Methodology and application of combined watershed and ground-water models in Kansas. *Journal of Hydrology*, 236, 185–201.
- Srinivas, A., Venkateswara Rao, B.V., and Gurunadha Rao, V.V.S. 1999. Recharge process and aquifer models of a small watershed. *Hydrological Sciences Journal*, 44, 681–692.
- Stephens, D.B. 1996. *Vadose Zone Hydrology*. Lewis Publishers, Florida.
- Stewart, J.B., Engman, E.T., Feddes, R.A., and Kerr, Y.H. 1998. Scaling up in hydrology using remote sensing: Summary of a Workshop. *International Journal of Remote Sensing*, 19, 181–194.
- Stoertz, M.W. and Bradbury, K.R. 1989. Mapping recharge areas using a ground water flow model: A case study. *Ground Water*, 27, 220–228.
- Taniguchi, M. 2002. Estimations of the past groundwater recharge rate from deep borehole temperature data. *Catena*, 48, 39–51.
- Taniguchi, M. and Sharma, M.L. 1993. Determination of groundwater recharge using the change in soil temperature. *Journal of Hydrology*, 148, 219–229.
- Taniguchi, M., Tsujimura, M., and Tanaka, T. 1996. Significance of stemflow in groundwater recharge. 1: Evaluation of the stemflow contribution to recharge using a mass balance approach. *Hydrological Processes*, 10, 71–80.
- Tarasov, L. and Peltier, W.R. 2004. A geophysically constrained large ensemble analysis of the deglacial history of the North American ice-sheet complex. *Quaternary Science Reviews*, 23, 359–388.
- Therrien, R. and Sudicky, E.A. 1996. Three-dimensional analysis of variably-saturated flow and solute transport in discretely-fractured porous media. *Journal of Contaminant Hydrology*, 23, 1–44.
- Thornthwaite, C.W. and Mather, J.R. 1957. Instructions and tables for computing potential evapotranspiration and the water balance. *Publications in Climatology*, 10, 185–311.
- Toth, J. 1963. A theoretical analysis of groundwater flow in small drainage basins. *Journal of Geophysical Research*, 68, 4795–4812.
- Vaccaro, J.J. 1992. Sensitivity of groundwater recharge estimates to climate variability and change, Columbia Plateau, Washington. *Journal of Geophysical Research*, 97, 2821–2833.
- VanderKwaak, J.E. 1999. *Numerical Simulation of Flow and Chemical Transport in Integrated Surface-Subsurface Hydrologic Systems*. PhD thesis, Department of Earth Sciences, University of Waterloo, Waterloo, Ontario, Canada.
- VanderKwaak, J.E. and Sudicky, E.A. 2000. Application of a physically based numerical model of surface and subsurface flow and solute transport. In: Stauffer, F., W. Kinzelbach, K. Kovar, and E. Hoehn. (eds.) *Proceedings of ModelCARE 1999, Calibration and Reliability in Groundwater Modelling*. IAHS Publication No. 265, pp. 515–523.
- Varni, M.R. and Usunoff, E.J. 1999. Simulation of regional-scale groundwater flow in the Azul River basin, Buenos Aires Province, Argentina. *Hydrogeology Journal*, 7, 180–187.
- Vermulst, J.A.P. and De Lange, W.J. 1999. An analytic-based approach for coupling models for unsaturated and saturated groundwater flow at different scales. *Journal of Hydrology*, 226, 262–273.
- Walker, G.R., Zhang, L., Ellis, T.W., Hatton, T.J., and Petheram, C. 2002. Estimating impacts of changed land use on recharge: review of modelling and other approaches appropriate for management of dryland salinity. *Hydrogeology Journal*, 10, 68–90.
- Wilkinson, W.B. and Cooper, D.M. 1993. Response of idealized aquifer/river systems to climate change. *Hydrological Sciences Journal*, 38, 379–389.

- Williamson, D.R. 1990. *Salinity — An Old Environmental Problem (abstract)*. Technical Memorandum 90/7, CSIRO Division of Water Resources, Institute of Natural Resources and Environment.
- Winter, T.C. 1978. Numerical simulation of steady state three-dimensional groundwater flow near lakes. *Water Resources Research*, 14, 245–254.
- Winter, T.C. 1983. The interaction of lakes with variably saturated porous media. *Water Resources Research*, 19, 1203–1218.
- Wittenberg, H. and Sivapalan, M. 1999. Watershed groundwater balance estimation using streamflow recession analysis and baseflow separation. *Journal of Hydrology*, 219, 20–33.
- Woolhiser, D.A., Smith, R.E., and Gilardez, J.-V. 1966. Effects of spatial variability of saturated hydraulic conductivity on Hortonian overland flow. *Water Resources Research*, 32(3), 671–678.
- Wu, J., Zhang, R., and Yang, J. 1997. Estimating infiltration recharge using a response function model. *Journal of Hydrology*, 198, 124–139.
- Xiao, Q., Ustin, S.L., and Wallender, W.W. 1996. A spatial and temporal continuous surface-subsurface hydrologic model. *Journal of Geophysical Research*, 101, 29565–29584.
- Xu, Y. and van Tonder, G.J. 2001. Estimation of recharge using a revised CRD method. *Water South Africa*, 27, 341–344.
- Yan, J. and Smith, K.R. 1994. Simulation of integrated surface water and ground water systems — model formulation. *Water Resources Bulletin*, 30, 879–890.
- Yang, Y., Lerner, D.N., Barrett, M.H., and Tellam, J.H. 1999. Quantification of groundwater recharge in the city of Nottingham, UK. *Environmental Geology*, 38, 183–198.
- Yang, J., Li, B., and Liu, S. 2000. A large weighing lysimeter for evapotranspiration and soil-water-groundwater exchange studies. *Hydrological Processes*, 14, 1887–1897.
- Zhang, L. and Dawes, W.R. 1998. WAVES — *An Integrated Energy and Water Balance Model*. Technical Report 98(31), CSIRO Land and Water Division.
- Zhang, L., Dawes, W.R., Hatton, T.J., Reece, P.H., Beale, G.T.H., and Packer, I. 1999a. Estimation of soil moisture and groundwater recharge using the TOPOG_IRM model. *Water Resources Research*, 35, 149–161.
- Zhang, L., Hume, I.H., O'Connell, M.G., Mitchell, D.C., Milthorpe, P.L., and Yee, M.D.H. 1999b. Estimating episodic recharge under different crop/pasture rotations in the Mallee region. Part 1. Experiments and model calibration. *Agricultural Water Management*, 42, 219–235.
- Zhang, L., Dawes, W.R., Hatton, T.J., Hume, I.H., O'Connell, M.G., and Mitchell, D.M.Y. 1999c. Estimating episodic recharge under different crop/pasture rotations in the Mallee region. Part 2. Recharge control by agronomic practices. *Agricultural Water Management*, 42, 237–249.

29

Ecohydrology: Water, Carbon, and Nutrient Cycling in the Soil–Plant Atmosphere Continuum

29.1	Introduction.....	29-2
29.2	Atmospheric Controls on Photosynthesis and Transpiration.....	29-2
	Scalar Mass Balance and Turbulent Transport • Scalar Transfer from Leaf Stomata and Resistance Parameterization	
	• Photosynthesis and \bar{C}_i Parameterization • Radiative Transfer	
	• Parameterization of \bar{T}_i and \bar{q}_i : The Leaf Energy Budget	
	• Modeling the Flow Field inside Canopies •	
	Modeling Ecosystem Fluxes • Transpiration and Hydraulic Controls	
29.3	Soil Water Balance in the Vadose Zone	29-13
	Instantaneous Equation for Soil-Moisture Dynamics •	
	Soil-Moisture Dynamics at the Daily Time Scale •	
	Long-Term Water Balance • Heat Flow into the Soil	
29.4	Soil Moisture Deficit and Plant Water Stress.....	29-19
	Physiological Impact of Soil Moisture Deficit on Plants •	
	Modeling Plant Water Stress • Soil-Climate-Vegetation Control on Plant Water Stress and Plant Carbon Assimilation	
29.5	Soil Nutrient Dynamics	29-24
	Brief Review of Soil Carbon and Nitrogen Cycles • Model Structure	
	• System Behavior under Constant and Stochastic Soil Moisture Conditions • Soil Organic Matter and Soil Hydraulic Properties	
29.6	Summary and Conclusions	29-34
	List of Symbols	29-34
	Glossary	29-36
	References	29-37
	Further Information	29-42

Edoardo Daly,
Gabriel Katul, and
Amilcare Porporato
Duke University

29.1 Introduction

The vadose zone is located between the land surface and the water table within which the soil moisture content is below saturation (except in the capillary fringe) and the water pressure is less than atmospheric. In this zone, the soil pore space typically contains air and other gases (e.g., O₂, CO₂) that are central to the biogeochemistry of the soil-plant-atmosphere system. The dynamics of water and nutrients in this zone impact groundwater quality and quantity through intricate interconnections between physical, biological, and chemical processes attributed to the presence of vegetation, soil microorganism, and their interaction with local climate and above-ground micrometeorology.

The urgency of linking carbon-nutrient-water activities in the vadose zone to simultaneous plant activities and ground recharge cannot be underemphasized. Management strategies to reduce anthropogenic CO₂ from the atmosphere are now providing incentives for timber-product companies to increase the managed plantation acreage and shortening stand rotations. For example, in the Southeastern United States, managed pine plantations are projected to increase by up to 30% of the entire timberland area in 2050. Naturally, to increase aboveground net primary productivity, nitrogen fertilization is routinely employed with application rates well in excess of 10 g N m⁻² yr⁻¹ (natural N deposition rates in the Southeast are on the order of 0.5 g N m⁻² yr⁻¹). While nitrogen amendments can increase aboveground biomass and leaf area density as well as leaf photosynthetic capacity, the impact of such high N fertilization on groundwater quality cannot be ignored.

A detailed description of such dynamics requires taking into account the water, energy, carbon, and nutrients budgets in the soil-plant-atmosphere system, where fluxes and state and forcing variables fluctuate on a broad range of spatial and temporal scales, making the entire system extremely complex and high-dimensional. The modeling difficulty is compounded by the central role played by the biological components, going from the plant stomatal control of transpiration and photosynthesis to plant nutrient uptake, litter production, and the dynamics of soil organic matter (SOM) and microbial biomass.

The focus of this chapter is the mathematical modeling of carbon-nutrient-water cycling in the soil-plant-atmosphere system with particular attention to time scales most relevant to groundwater quality and quantity. Our review will be centered on the temporal dynamics, while spatial issues will not be directly considered.

At the finest scale, describing water vapor exchange (and transpiration) over vegetated surfaces commands simultaneous considerations of heat and CO₂ exchanges, as all these exchange processes are intertwined at the most fundamental level, the leaf stomata. Hence, a logical starting point of this chapter is to describe the basic formulations for computing maximum CO₂ uptake through the stomata and maximum H₂O loss out of the stomata based on hydro-meteorological conditions and photosynthetic demands (Section 29.2). We then progress toward a discussion of the impact of soil-water availability on these maximum rates and their implications to soil biogeochemistry and groundwater recharge. Section 29.3 presents a stochastic model for soil-moisture dynamics in the vadose zone, along with the consequent implications in the long-term water balance and the soil heat balance. The main relations between soil water availability and vegetation behaviors are described in Section 29.4, where plant water stress is modeled and its influence on vegetation carbon assimilation is analyzed. Section 29.5 is finally dedicated to soil biogeochemical cycles, and particularly to modeling soil carbon and nitrogen, whose dynamics is strongly dependent on soil-moisture level and variability.

29.2 Atmospheric Controls on Photosynthesis and Transpiration

This section investigates recent developments in multi-layer methods needed for computing distributions and strengths of H₂O, heat, and CO₂ (often referred to as scalars) sources and sinks within the above-ground canopy volume using basic radiative transfer schemes, turbulent transport theories, and physiological principles assuming water loss is entirely “slaved” to the carbon needs of leaf photosynthesis.

Hence, the resulting carbon-water rates from these models represent the maximum transpiration and photosynthesis for a set of physiological properties and microclimate conditions (e.g., radiative load, vapor pressure deficit, air temperature, etc.). The conservation equations and parameterization schemes used in such methods are reviewed next.

In multi-layer models, the canopy is represented by a sequence of finite layers (Figure 29.1), each of thickness dz , and encompassing leaf area density $a(z)$. The leaf area index (LAI) of the stand is given by $LAI = \int_0^h a(z)dz$, where h is the mean canopy height.

29.2.1 Scalar Mass Balance and Turbulent Transport

For a horizontally uniform and rigid canopy, the time and horizontally averaged one-dimensional scalar mass conservation equations for mean carbon dioxide (\bar{C}_a), water vapor (\bar{q}_a), and temperature (\bar{T}_a), within a layer of thickness dz (Figure 29.1) can be written as

$$\begin{aligned} \frac{\partial \bar{C}_a}{\partial t} + \frac{\partial F_c}{\partial z} &= S_c \\ \frac{\partial \bar{q}_a}{\partial t} + \frac{\partial F_q}{\partial z} &= S_q \\ \frac{\partial \bar{T}_a}{\partial t} + \frac{\partial F_T}{\partial z} &= S_T \end{aligned} \quad (29.1)$$

where t is time, z is height above the forest floor (this notation for z will reverse once we consider the vadose zone), F_c , F_q , and F_T are the turbulent vertical fluxes of CO₂, water vapor, and heat, and S_c , S_q , and S_T are the vegetation source (or sink) strength of CO₂, water vapor, and heat, respectively. Mean quantities are subject to both time and horizontal averaging as discussed in Raupach and Shaw (1982). Temporal averaging, often on the order of 1/2 h, is needed to average out turbulent eddies responsible for much of the mass transport to and from the canopy, and spatial averaging, often on the order of few canopy heights, is needed to include the effects of the canopy on the airflow. Equation 29.1 represents three equations with nine unknowns necessitating additional formulations to solve for \bar{C}_a , \bar{q}_a , \bar{T}_a , F_c , F_q , F_T , S_c , S_q , and S_T .

The required additional formulations can be derived from the relationship between source strength and mean concentration via basic fluid mechanics principles. A set of prognostic equations describing the

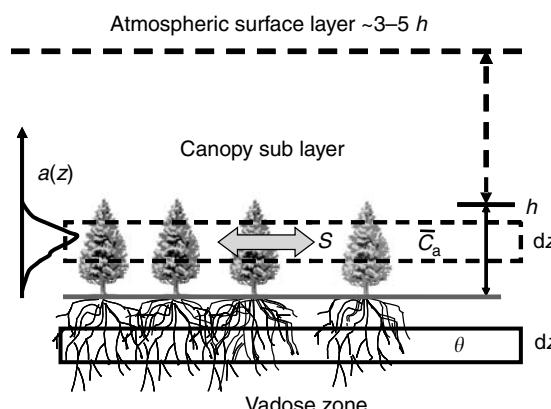


FIGURE 29.1 Schematic representation of a canopy with a mean height h and leaf area density $a(z)$ (foliage area per unit volume) showing an arbitrary layer of thickness dz . The leaves are assumed to exchange heat, water vapor, and CO₂ with the neighboring atmosphere via source terms S . The term $a(z)$ is analogous to the porosity distribution in groundwater flow. (Modified from Katul, G.G. and Siqueira, M., *Vadose Zone J.*, 1, 58, 2002.)

relationship between concentration and source (or sink) for CO₂, water vapor, and temperature can be formulated in the Lagrangian or Eulerian frame of reference. In the Lagrangian frame of reference, these equations take the form of an arbitrary scalar s (representing here either CO₂, H₂O, or temperature)

$$\bar{s} = \frac{1}{\rho_a} \iint P(z, t; z_0, t_0) S_s(z_0, t_0) dz dt \quad (29.2a)$$

where ρ_a is the mean air density, $P(z, t; z_0, t_0)$ is the transition probability density function determined from measured or modeled turbulent flow statistics within the canopy volume, and S_s is the source (or sink) strength of scalar s , and \bar{s} is the mean concentration. For practical estimation of $P(z, t; z_0, t_0)$, Lagrangian analytical models were developed and tested for a wide range of vegetation types (Raupach, 1989a,b; Katul et al., 1997; Warland and Thurtell, 2000). In particular, the Localized Near Field theory (LNF) proposed by Raupach (1989a,b) proved to be a parsimonious model for such applications (Gu et al., 1999; Denmead et al., 2000; Leuning, 2000; Siqueira et al., 2000; Katul and Siqueira, 2002) though its ability to explicitly deal with thermal stratification within the canopy remains absent.

In the Eulerian frame of reference, the time and horizontally averaged conservation equation for the vertical scalar flux budget s' ($= s - \bar{s}$) can be expressed as

$$\begin{aligned} \frac{\partial \overline{w' s'}}{\partial t} &= 0 = -\overline{w'^2} \frac{\partial \bar{s}}{\partial z} - \overline{\frac{\partial w' w' s'}{\partial z}} - \overline{s' \frac{\partial p'}{\partial z}} \\ 0 &= P_1 + P_2 + P_3 \end{aligned} \quad (29.2b)$$

where w' is the turbulent vertical velocity, and terms P_2 and P_3 are unknowns requiring closure approximations if the velocity statistics are known. Note that Equation 29.2b neglects buoyancy, scalar drag, and waving source production though they can be readily included in such formulations as discussed in Finnigan (1985) and Siqueira and Katul (2002). Using higher order closure models and detailed measurements within a pine forest, Siqueira and Katul (2002) showed that accounting for thermal stratification during day-time conditions is important for matching the sensible heat flux, and accounting for thermal stratification at night is important for modeling CO₂ storage dynamics within the canopy volume. Below, we show the near-neutral stratification for illustration purposes though the more general case is derived and discussed elsewhere (Siqueira and Katul, 2002). The flux transport term (P_2) can be modeled after Meyers and Paw U (1987) and the dissipation term (P_3) can be modeled after Finnigan (1985). These approximations often lead to standard closure schemes given by

$$\begin{aligned} \overline{w' w' s'} &= \frac{\tau}{C_8} \left[-\overline{w' w' w'} \frac{\partial \bar{s}}{\partial z} - \overline{w' s'} \frac{\partial \overline{w' w'}}{\partial z} - 2\overline{w' w'} \frac{\partial \overline{w' s'}}{\partial z} \right] \\ \overline{s' \frac{\partial p'}{\partial z}} &= C_4 \frac{\overline{w' s'}}{\tau} \end{aligned}$$

where C_4 and C_8 (the subscripts here are retained for consistency with standard nomenclature in turbulent transport theories) are closure constants and τ is an Eulerian relaxation time scale given by

$$\tau = \frac{\sigma_e^2}{\varepsilon},$$

where $\sigma_e = \sqrt{\overline{u'_i u'_i}}$ is a characteristic turbulent velocity (proportional to the turbulent kinetic energy), and ε is the mean rate of turbulent kinetic dissipation, and u_i are the velocity components in the x_1 (or x), x_2 (or y), and x_3 (or z) directions, respectively, with x_1 aligned along the mean wind direction so that $\overline{u_1} = 0$. The relaxation time τ is strictly dependent on the velocity statistics and can be computed from

momentum higher-order closure models (Katul and Albertson, 1998, 1999; Katul and Chang, 1999) or LES techniques.

Upon combining these closure approximations with the scalar budget equation, for example, Equation 29.2b, a second order ordinary differential equation (ODE) can be derived to describe the variations of the scalar turbulent flux with height (Siqueira et al., 2000; Katul and Albertson, 1998), as

$$A_1(z) \frac{\partial^2 \overline{w' s'}}{\partial z^2} + A_2(z) \frac{\partial \overline{w' s'}}{\partial z} + A_3(z) \overline{w' s'} = A_4(z)$$

where

$$\begin{aligned} A_1(z) &= \frac{2\tau}{C_8} \overline{w' w'} \\ A_2(z) &= \frac{\tau}{C_8} \frac{\partial \overline{w' w'}}{\partial z} + 2 \frac{\partial}{\partial z} \left(\frac{\tau}{C_8} \overline{w' w'} \right) \\ A_3(z) &= \frac{\partial}{\partial z} \left(\frac{\tau}{C_8} \frac{\partial \overline{w' w'}}{\partial z} \right) - C_4 \frac{1}{\tau} \\ A_4(z) &= \overline{w' w'} \frac{\partial \bar{s}}{\partial z} - \frac{\partial}{\partial z} \left(\frac{\tau}{C_8} \overline{w' w' w'} \right) \frac{\partial \bar{s}}{\partial z} - \left(\frac{\tau}{C_8} \overline{w' w' w'} \right) \frac{\partial^2 \bar{s}}{\partial z^2}. \end{aligned}$$

Combining Equation 29.1 and either Equation 29.2a or Equation 29.2b to estimate $\overline{C_a}$, $\overline{q_a}$, $\overline{T_a}$, F_c , F_q , F_T , S_c , S_q , and S_T reduces now the problem to six equations with nine unknowns and still remains unsolvable. To mathematically close this system of equations, three additional relations describing mass transfer from the leaves to the atmosphere are required. These are described next.

29.2.2 Scalar Transfer from Leaf Stomata and Resistance Parameterization

The scalar source strength can be related to the atmospheric concentration by a Fickian diffusion formulation through the stomatal cavity and the leaf boundary layer. These formulations lead to the following expressions:

$$S_c(z) = -\rho a(z) \frac{\overline{C_a}(z) - \overline{C_i}(z)}{r_b(z) + r_s(z)} \quad (29.3a)$$

$$S_q(z) = -\lambda_w a(z) \frac{\overline{q_a}(z) - \overline{q_i}(z)}{r_b(z) + r_s(z)} \quad (29.3b)$$

$$S_T(z) = -\rho C_p a(z) \frac{\overline{T_a}(z) - \overline{T_i}(z)}{r_b(z)} \quad (29.3c)$$

where subscripts “a” and “i” represent the “free” atmosphere and stomata sub-cavity space concentration values, λ_w is the latent heat of water vaporization, C_p is the specific heat of dry air under constant pressure, $a(z)$ is, as before, the leaf area density, and $r_b(z)$ and $r_s(z)$ are the mean boundary layer and stomatal resistances at a given z , respectively. For simplicity, all the symbols representing mean resistances (or conductance) are shown without overbar throughout. Combining Equation 29.1 to Equation 29.3, $\overline{C_a}$, $\overline{q_a}$, $\overline{T_a}$, F_c , F_q , F_T , S_c , S_q , and S_T can be solved if $r_b(z)$, $r_s(z)$, $\overline{C_i}$, $\overline{q_i}$ and $\overline{T_i}$ are modeled or parameterized. Modeling $r_b(z)$, $r_s(z)$, $\overline{C_i}$, $\overline{q_i}$ and $\overline{T_i}$ is discussed next.

The estimation of r_b is based on flat plate theory (Schuepp, 1993; Baldocchi and Meyers, 1998), and requires the mean wind speed within the canopy, which in turn is modeled from second order closure principles (and will be reviewed later).

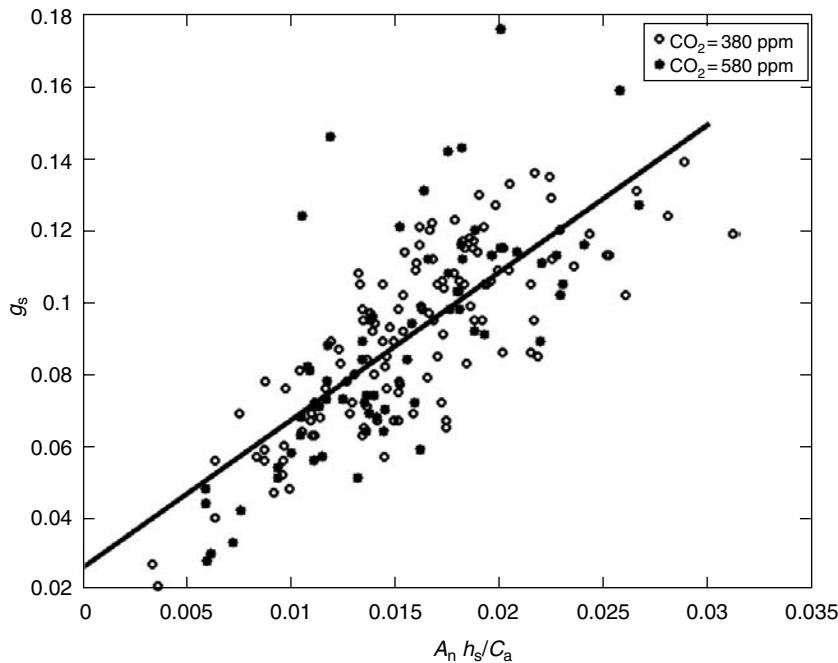


FIGURE 29.2 Comparison between measured and modeled leaf stomatal conductance (g_s) using Equation 29.4 for sunlight Loblolly pine needles under ambient (=380 ppm) and elevated (=580 ppm) atmospheric CO₂ concentration levels. The g_s , leaf photosynthesis (A_n), and the surface relative humidity (h_s) data are collected using leaf-level gas exchange porometry within the Free Air CO₂ Enrichment (FACE) facility at the Blackwood division of the Duke Forest, near Durham, NC.

Collatz et al. (1991) developed a semi-empirical physiological model to relate stomatal conductance $g_s = r_s^{-1}$ to leaf photosynthesis, given by

$$g_s = m \frac{A_n h_s}{C_a} + b \quad (29.4)$$

where m and b are species-specific parameters, often determined by gas-exchange measurements, A_n is the net leaf-scale assimilation rate, and h_s is the mean air relative humidity near the leaf surface (not to be confused with canopy height h). Equation 29.4 has been tested for a wide range of climatic conditions, including elevated atmospheric CO₂ as shown in Figure 29.2. The experiments in Figure 29.2 suggest the parameters m and b are robust to such wide range of variations in atmospheric CO₂ conditions and do reflect “internal” physiological properties. While there is some debate as to whether relative humidity or vapor pressure deficit ought to be used in such models (e.g., Leuning, 1995; Katul et al., 2000; Kumagai et al., 2004), the data-model comparison in Figure 29.2 suggests that Equation 29.4 (or variants on it) is sufficient to link leaf conductance to photosynthesis, at least for the purposes of modeling “maximum” leaf photosynthesis and subsequent carbon input to the vadose zone.

29.2.3 Photosynthesis and $\overline{C_i}$ Parameterization

According to Farquhar et al. (1980), the net photosynthetic rate at the leaf scale depends on light, CO₂, and leaf temperature and can be described as

$$A_n \approx \min \left\{ \frac{J_E}{J_c} \right\} - R_d, \quad (29.5)$$

where J_E and J_c are the assimilation rate restricted by light and ribulose bisphosphate (R_uBP) carboxylase (or Rubisco) respectively, and R_d is the respiration rate during daytime without photorespiration. The flux per unit leaf area, J_E , describes the light restriction on photosynthesis, given by

$$J_E = \alpha_l \times e_m \times Q_p \frac{\bar{C}_i - \Gamma_*}{\bar{C}_i + 2\Gamma_*}$$

where α_l is the leaf absorptivity for photosynthetically active radiation (PAR), e_m is the maximum quantum efficiency for CO₂ uptake, Q_p is the PAR irradiance on the leaf surface (and varies significantly inside the canopy), \bar{C}_i , as before, represents mean intercellular mean CO₂ concentration. The CO₂ compensation point, Γ_* , is the CO₂ concentration at which $A_n = 0$ in the absence of photorespiration, and is given by:

$$\Gamma_* = \frac{[O_2]}{2\omega}$$

where [O₂] is the oxygen concentration in air (≈ 210 mmol mol⁻¹), and ω is a ratio of kinetic parameters describing the partitioning of R_uBP to the carboxylase or oxygenase reactions of Rubisco (Campbell and Norman, 1998). The leaf flux J_c is the Rubisco-limited rate and is estimated from

$$J_c = \frac{V_{c\max}(\bar{C}_i - \Gamma_*)}{\bar{C}_i + K_c(1 + [O_2]/K_{o_2})}$$

where $V_{c\max}$ (μmol m⁻² sec⁻¹) is the maximum catalytic capacity of Rubisco per unit leaf area, K_c and K_{o_2} are the Michaelis constants for CO₂ fixation and O₂ inhibition with respect to CO₂, respectively. Note that J_c increases linearly with increasing \bar{C}_i , but approaches a maximum $V_{c\max}$ under high CO₂ concentration state, which is rarely encountered under natural conditions. Following Collatz et al. (1991), the respiration rate R_d can be estimated as $R_d = 0.015V_{c\max}$.

Temperature dependence of kinetic variables is computed following the equations in Campbell and Norman (1998). Five kinetic parameters are needed to adjust for temperature, namely K_c , K_{o_2} , ω , $V_{c\max}$ and R_d . For the first three parameters, an exponential temperature function is employed:

$$k = k_{25} \exp[y(T_s - 25)]$$

with k defined at leaf surface temperature T_s , and where k_{25} is the value of the parameter at 25°C, and y is the temperature coefficient for that parameter. In addition, $V_{c\max}$ and R_d are adjusted by a more complex function incorporating deactivation effects at extremely high temperatures (see Figure 29.3), such as

$$V_{c\max} = \frac{V_{c\max,25} \exp[0.031(T_s - 25)]}{1 + \exp[0.115(T_s - 41)]}$$

and

$$R_d = \frac{R_{d,25} \exp[0.069(T_s - 25)]}{1 + \exp[1.3(T_s - 55)]}$$

where $V_{c\max,25}$ and $R_{d,25}$ are the values of $V_{c\max}$ and R_d at 25°C, respectively. Again, this strong dependency of $V_{c\max}$ on leaf temperature necessitates simultaneous modeling of heat, water vapor, and CO₂ exchange.

Finally, at the leaf scale, \bar{C}_s , \bar{C}_i , and \bar{C}_a are related by $\bar{C}_i = \bar{C}_s - A_n/g_s$ and $\bar{C}_s = \bar{C}_a - A_n/g_b$.

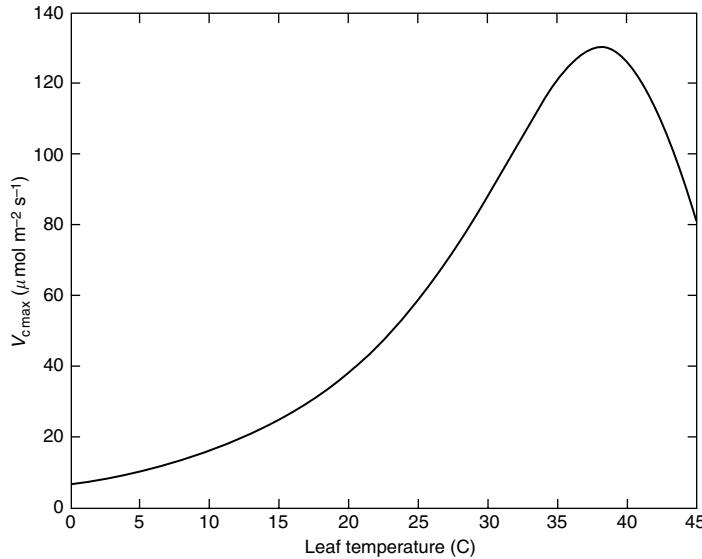


FIGURE 29.3 Typical relationship between $V_{c\max}$ and leaf temperature for a southern Loblolly pine stand. Note the deactivation effects at air temperatures in excess of 36°C.

29.2.4 Radiative Transfer

Radiation attenuation inside the canopy is critical to both the Farquhar photosynthesis model (through J_E) and the leaf energy budget. The light transmission model of Campbell and Norman (1998) is perhaps among the most parsimonious models and is briefly described below.

The solar (short-wave) radiation can be decomposed into direct beam, diffuse, and reflective radiation. For thermal or long-wave radiation, the atmosphere can be assumed to function as a gray body (e.g., Campbell and Norman, 1998). This permits using an averaged sky emissivity for calculating atmospheric thermal emittance. For simplicity, a constant emissivity of 0.97 can be employed.

Light transmission through the canopy is computed for sunlit and shaded portions separately to estimate PAR and near-infrared (NIR) irradiance absorbed at each canopy level. This waveband decomposition is necessary because leaf absorptivity is different for these two spectral bands (Monteith and Unsworth, 1990; Campbell and Norman, 1998).

The fraction $\tau_b(\psi)$ of incident beam radiation from a zenith angle ψ penetrating through the canopy is given by

$$\tau_b(\psi) = \exp(-\sqrt{\alpha_l} K_{be}(\psi) a_l \Omega), \quad (29.6)$$

where $K_{be}(\psi)$ is the extinction coefficient for an ellipsoidal leaf distribution (see Campbell and Norman, 1998), $a_l (= - \int_h^z a(z) dz)$ is the cumulative leaf area density integrated from the canopy top, and Ω is the clumping factor of leaf distribution ($\Omega = 1$ when leaves are randomly distributed in space).

29.2.5 Parameterization of \bar{T}_i and \bar{q}_i : The Leaf Energy Budget

The energy budget at the leaf surface is used to solve for surface variables (i.e., \bar{T}_i and \bar{q}_i) and absorbed radiation for estimating g_s . Due to the strong nonlinearity in the leaf energy budget, a Taylor series expansion along with a Penman approximation were used to derive an explicit algebraic equation for mean leaf surface temperature (\bar{T}_s), given by

$$\bar{T}_s = \bar{T}_a + (Q_{ab} - \lambda_w g_v D_v / p_a - \varepsilon_l \sigma \bar{T}_a^4) / (h_c + \lambda_w g_v \Delta_s / p_a + 4 \varepsilon_l \sigma \bar{T}_a^3)$$

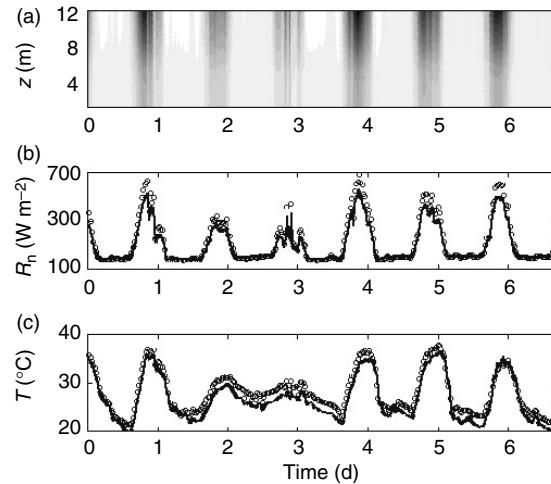


FIGURE 29.4 (a) Modeled radiation attenuation inside a 17-year-old pine canopy at the Blackwood division of the Duke Forest, near Durham, North Carolina. Regions of maximum radiation absorption are in dark. (b) Comparison between modeled radiation absorbed by the canopy and net radiation measured above canopy. (c) Comparison between remotely sensed, measured, and modeled skin temperature in the upper third of the canopy. (Revised from Lai, C.T., et al. *Bound. Lay. Meteorol.*, 95, 91, 2000.)

where Q_{ab} is the absorbed energy at each canopy layer, g_v is the water vapor conductance (derived from g_s and g_b for water vapor), D_v is the vapor pressure deficit, p_a is the atmospheric pressure, ε_l is the leaf emissivity, σ is the Stefan-Boltzmann constant, h_c is a convection coefficient (not to be confused with canopy height), and Δ_s is the slope of the curve relating saturation vapor pressure to temperature. A comparison between modeled absorbed radiation and measured net radiation, and measured and modeled remotely sensed skin temperature are shown in Figure 29.4. The agreement between modeled and measured surface temperature in the figure for the canopy morphology of a pine forest is encouraging despite the simplifications in the radiative transfer schemes.

29.2.6 Modeling the Flow Field inside Canopies

To predict the flow statistics for Equation 29.2, particularly the second-moment statistics such as $\sigma_w = (w'w')^{1/2}$, $\sigma_u = (\bar{u}'u')^{1/2}$, $\sigma_v = (v'v')^{1/2}$, $\sigma_e = (\sigma_u^2 + \sigma_v^2 + \sigma_w^2)^{1/2}$, and the mean velocity (\bar{u}) needed for the calculation of r_b , second-order closure principles can be used. One simple approach is to consider the analytical theory of Massman and Weil (1999), hereafter referred to as MW99, which assumes an exponential mean velocity profile given by

$$\frac{\bar{u}(z)}{\bar{u}(h)} = e^{-n\gamma(z)}$$

and results in

$$\frac{\overline{w'u'}(z)}{u_*^2} = e^{-2n(\gamma(z))}, \quad \frac{\sigma_u}{u_*} = A_u v_1 \frac{\sigma_e(z)}{u_*}, \quad \frac{\sigma_v}{u_*} = A_v v_1 \frac{\sigma_e(z)}{u_*}, \quad \frac{\sigma_w}{u_*} = A_w v_1 \frac{\sigma_e(z)}{u_*}$$

where

$$\frac{\sigma_e}{u_*} = [v_3 e^{-\Lambda \varsigma(h)\gamma(z)} + B_1 (e^{-3n\gamma(z)} - e^{-\Lambda \varsigma(h)\gamma(z)})]^{1/3}, \quad \gamma(z) = 1 - \frac{\varsigma(z)}{\varsigma(h)}, \quad \varsigma(z) = \int_0^z C_d a(z) dz$$

and the constants are given by

$$n = \frac{1}{2} \left(\frac{u_*}{\bar{u}(h)} \right)^{-2} \varsigma(h), \quad B_1 = \frac{-9u_*/\bar{u}(h)}{2\alpha v_1 [9/4 - \Lambda^2 u_*^4 / (\bar{u}(h))^4]}, \quad \Lambda^2 = \frac{3v_1}{\alpha^2}$$

$$v_1 = (A_u^2 + A_v^2 + A_w^2)^{-1/2}, \quad v_3 = (A_u^2 + A_v^2 + A_w^2)^{3/2}, \quad \text{and} \quad v_2 = \frac{v_3}{6} - \frac{A_w^2}{2v_1}$$

where, $u_*^2 = -\overline{u'w'}$ defined above the canopy (and assumed to be independent of z up to the atmospheric surface layer, see Figure 29.1).

The MW99 model computes the zero displacement height (d) from the centroid of the momentum sink using

$$\frac{d}{h} = 1 - \int_0^1 \left(\frac{\overline{u'w'}(r)}{u_*^2} \right) d(r)$$

where, $r = z/h$. Above the canopy, the flow is assumed to attain its atmospheric surface layer state (see Figure 29.1) with

$$\frac{\overline{u'w'}}{u_*^2} = -1, \quad \frac{\bar{u}}{u_*} = \frac{1}{\kappa} \log \left[\frac{z-d}{z_0} \right], \quad \frac{\sigma_u}{u_*} = A_u, \quad \frac{\sigma_v}{u_*} = A_v, \quad \frac{\sigma_w}{u_*} = A_w$$

where $z_0/h = (1 - d/h)e^{-k_v(\bar{u}_h/u_*)}$, and $\kappa = 0.4$ is the von-Karman constant, and z_0 is the momentum roughness length. This model has been tested for a wide range of canopies ranging from uniform rice fields to hardwood forest, as shown in Figure 29.5.

In Equation 29.2b, a simplified gradient-diffusion model for $\overline{w'w'w'} = -C_6 \sigma_e^2 \tau (d\overline{w'w'}/dz)$ can be used. Common values for the constants $A_u = 2.1\text{--}2.4$, $A_v = 1.8\text{--}2.2$, and $A_w = 1.1\text{--}1.3$.

29.2.7 Modeling Ecosystem Fluxes

When combining the scalar continuity equations, the turbulent flux budget equations, the radiation balance, the leaf energy balance, and the Farquhar model, it is possible to solve for $\overline{C_a}$, $\overline{q_a}$, $\overline{T_a}$, F_c , F_q , F_T , S_c , S_q , and S_T . The concentration and fluxes within and above the canopy can be compared to measurements for well-water conditions. For example, Figure 29.6 compares eddy-covariance measured and modeled scalar fluxes above a six-year-old pine canopy for an ambient and a fertilized stand at the South-Eastern Tree Research and Education Site (SETRES) in North Carolina, where the fertilization increased LAI by a factor of two, the respiring biomass by a factor of 1.5, and $V_{c,\max}$ by 20%. In both, ambient and N-fertilized stands, the formulation here reproduced well the eddy-covariance measured fluxes of heat, water vapor, and CO_2 .

29.2.8 Transpiration and Hydraulic Controls

The method described so far can determine plant transpiration for conditions in which soil moisture does not limit stomatal conductance using

$$T_{r,\max}(t) = \int_0^h S_q(t, z) dz$$

Because S_q and the leaf-conductance in Equation 29.4 do not include any hydraulic limitations (e.g., leaf pressure), the model cannot account for cavitation and other hydraulic-related phenomena leading to stomatal shutdown as drought progresses. This shutdown is crucial in shallow-rooted plants because

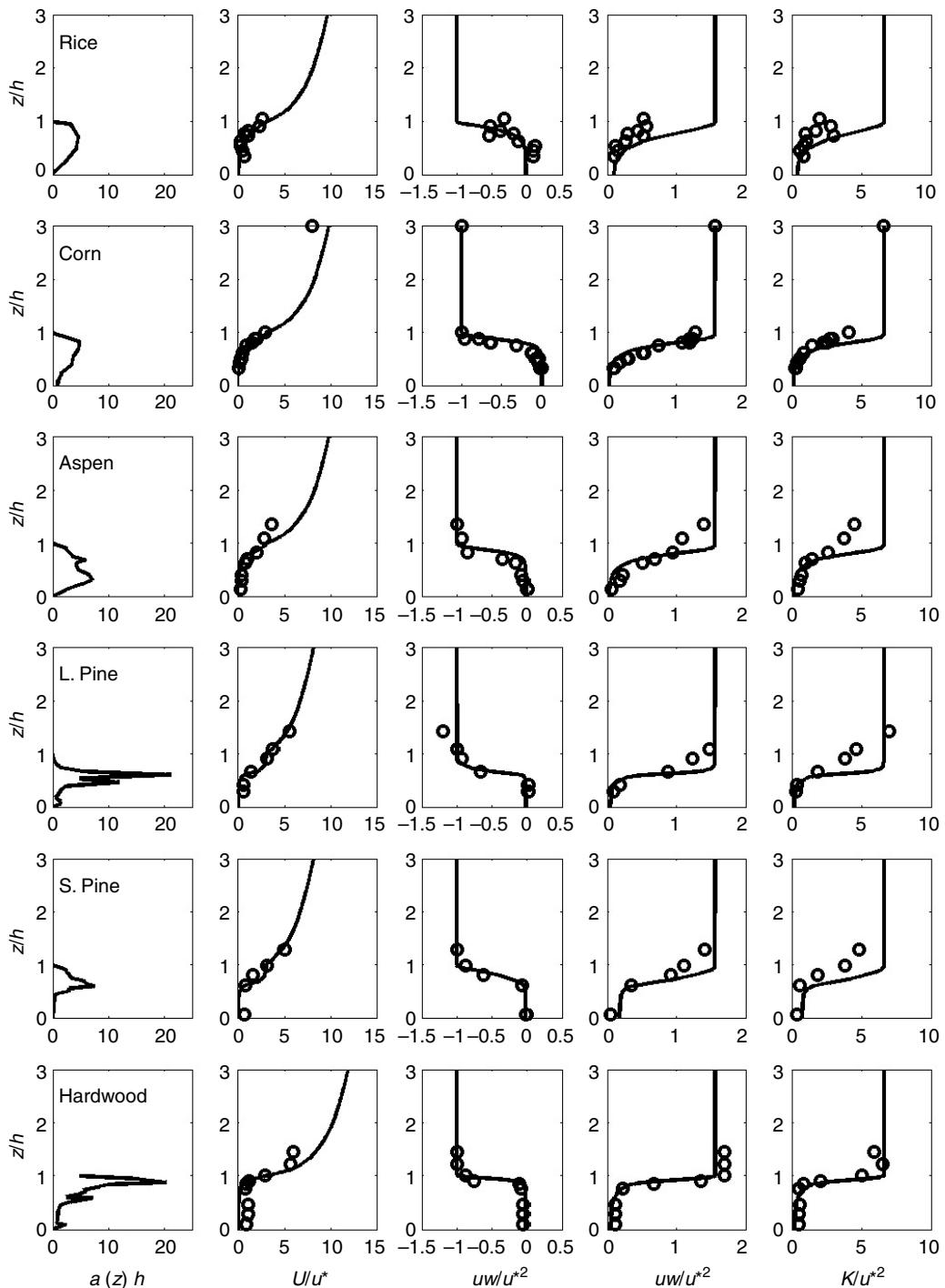


FIGURE 29.5 Comparison between measured and MW99 modeled velocity statistics for a wide range of canopy morphologies, including a rice- and a cornfield, an aspen forest, a southern Loblolly pine stand, a Scots pine forest, and a southern hardwood forest (see Nathan and Katul, 2005 for further details). Here $K = \frac{1}{2}\sigma_e^2$ is the turbulent kinetic energy, U is the mean wind speed, and uw and ww represent the momentum flux and the vertical velocity variance.

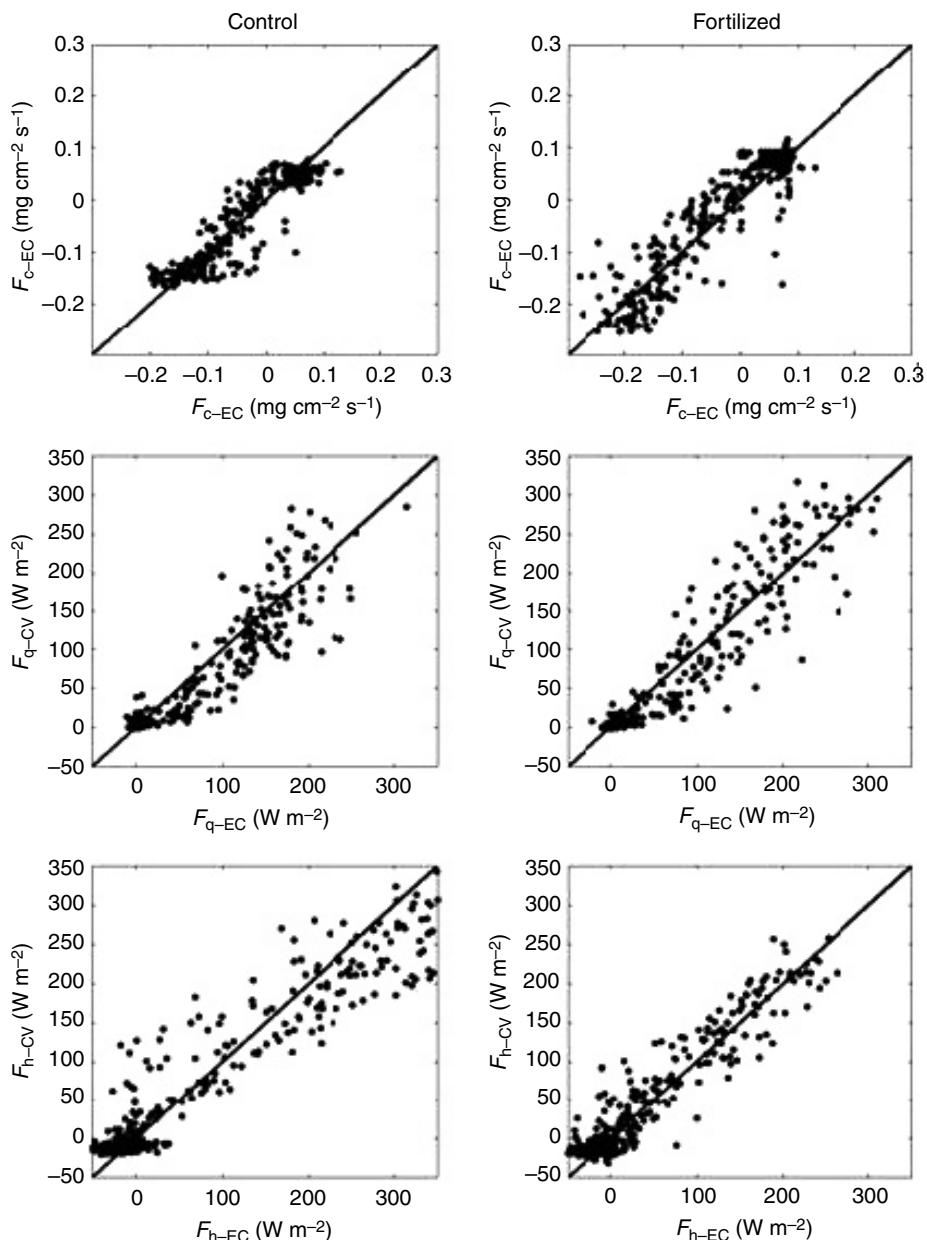


FIGURE 29.6 Comparison between 1/2-hourly measured (subscript EC for eddy-covariance) and modeled (subscript CV for canopy – vertically resolved) scalar fluxes of CO₂ (top), water vapor (middle), and sensible heat (bottom) for the 6-year-old ambient and fertilized pine stands at SETRES for control and fertilized plots. (From Lai, C.T., et al. *Plant Cell Environ.*, 25, 1095, 2002, with permission.)

the soil moisture reservoir is rather limited and can be depleted rapidly. Using sapflow and soil moisture measurements for the entire root zone (=30 cm), Oren et al. (1998) reported a reduction in transpiration of more than 60% in a southeastern Loblolly pine forest as the root-zone soil moisture content is reduced from 0.20 to 0.13. The actual transpiration is thus a function of the $T_{r,\max}$, that is, of the evaporative demand, and soil-water content (e.g., Guswa, 2005). For the sake of simplicity, in the following, the effect of evaporative demand on actual evapotranspiration will be neglected with respect to the influence of

water availability. Hence, for soil biogeochemical models, running on time-scales of months to decades, it is possible to represent the transpiration rate as

$$T_r = T_{r,\max} f(\theta)$$

where $T_{r,\max}$ is the maximum loss of water that can occur if the plant is permitted to satisfy its carbon uptake demand for a given set of light and micrometeorological conditions above the canopy (as determined from the model earlier described), and $f(\theta)$ is the soil moisture reduction function that accounts for stomatal shutdown to prevent runaway cavitation (e.g., Bohrer et al., 2005 for additional developments taking into account plant architecture). This is the basic transpiration-assimilation model used in the coupled C-N-H₂O described next.

29.3 Soil Water Balance in the Vadose Zone

As noted earlier, a correct description of soil-moisture dynamics is essential to predict plant uptake and evapotranspiration (see Section 29.2.8) and the related plant water stress, the soil energy balance and deep percolation of water and nutrients to groundwater, as well as the decomposition of SOM and the nutrient availability to plants. The main processes involved in the soil water balance are schematically described in Figure 29.7.

We will start with the balance equation of soil water at a point in space and time and then proceed with formal averaging procedures to describe the soil water balance at the daily time scale and vertically averaged over the root zone to provide a simplified but realistic equation to couple with the stochastic rainfall variability.

29.3.1 Instantaneous Equation for Soil-Moisture Dynamics

The three-dimensional continuity equation for soil moisture is given by

$$\frac{\partial \theta}{\partial t} = -\frac{\partial q_i}{\partial x_i} - S_r(\theta) \quad (29.7)$$

where θ is the volumetric soil moisture content at some arbitrary point x_i in the soil, q_i is the water flux in direction x_i , x_i ($i = 1, 2, 3$) being the Cartesian coordinates for the longitudinal ($x_1 = x$), the lateral ($x_2 = y$), and the vertical ($x_3 = z$, directed into the soil) direction, and S_r now represents the local sinks of water by the rooting system. For a planar homogeneous flow $\partial(\cdot)/\partial x_1 = \partial(\cdot)/\partial x_2 = 0$, and the continuity equation in the vertical direction reduces to

$$\frac{\partial \theta}{\partial t} = -\frac{\partial q_3}{\partial x_3} - S_r(\theta)$$

Upon depth averaging this equation from the soil surface ($x_3 = z = 0$) to the root-zone depth ($z = Z_r$), we obtain

$$\int_0^{Z_r} \frac{\partial \theta}{\partial t} dz = q(0) - q(Z_r) - \int_0^{Z_r} S_r(\theta) dz$$

where $q(0)$ is the amount of water infiltrating or evaporating from the soil system, $q(Z_r)$ is the drainage loss from the rooting system (that percolates to groundwater), and $\int_0^{Z_r} S_r(\theta) dz$ is the actual plant-water

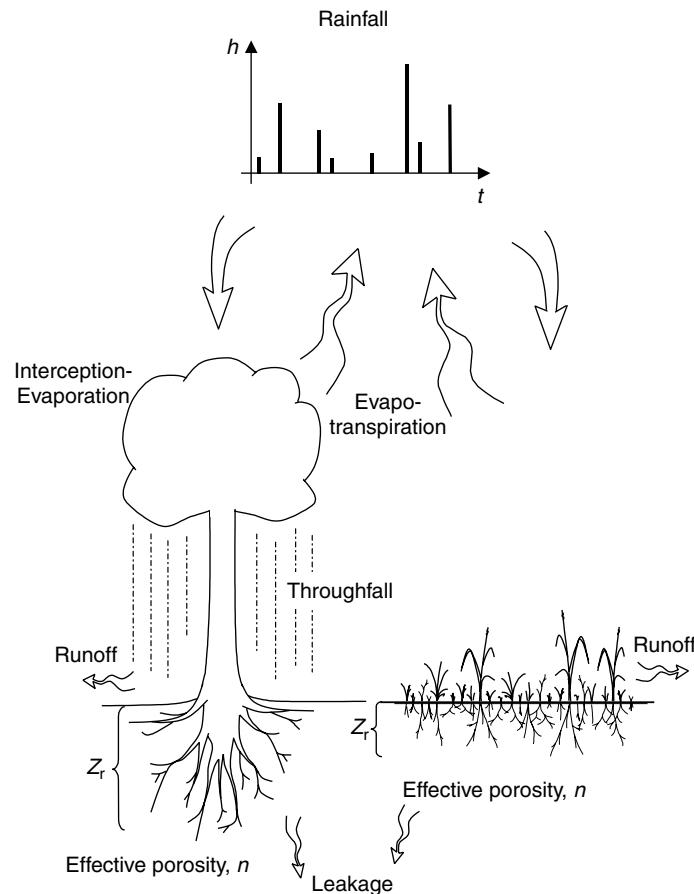


FIGURE 29.7 Schematic representation of the key processes controlling the soil water balance. (From Laio, F., et al. *Adv. Water Resour.*, 24, 707, 2001a, with permission.)

uptake or transpiration, T_r . Defining the normalized depth-averaged soil moisture content by

$$s_m(t) = \int_0^{Z_r} \frac{1}{Z_r(t)} \frac{\theta(t, z)}{n(t, z)} dz,$$

where $n(t, z)$ is the soil porosity, and using Leibniz's theorem, leads to the simplified hydrologic balance

$$nZ_r \frac{ds_m}{dt} + s_m \frac{d(nZ_r)}{dt} = q(0) - q(Z_r) - T_r$$

In the following, n and Z_r will be assumed not to change appreciably with t so that the term $s_m d(nZ_r)/dt$ is neglected compared to $nZ_r ds_m/dt$.

29.3.2 Soil-Moisture Dynamics at the Daily Time Scale

For hydrologic models, the logical time scale for resolving the soil moisture dry-down period is daily. Hence, it becomes necessary to apply an additional averaging operator to the previous formulation that

reduces the daily hydrologic balance to

$$nZ_r \frac{d\bar{s}_m}{dt} = \overline{q(0)} - \overline{q(Z_r)} - \overline{T_r}$$

where the overbar represents time averaging over a day (rather than over turbulent time scales as earlier defined).

For prognostic (vis-à-vis diagnostic) models of subsurface flow, the daily rainfall depth, which determines $\overline{q(0)}$, is not known; only its statistical properties are available from historical records (or climate forecasts in the future). Consequently, it is reasonable to treat $\overline{q(0)}$ as a stochastic process and consider the probabilistic treatment of soil moisture in relation to these precipitation statistics (the evaporation component of $\overline{q(0)}$ is described later).

This approach, pioneered by Eagleson (1978), gives rise to several simplified models accounting for rainfall stochasticity (e.g., Cordova and Bras, 1981; Hosking and Clarke, 1990; Milly, 1993; Rodriguez-Iturbe et al., 1999; Laio et al., 2001a; Porporato et al., 2004; Daly et al., 2004b). Although, it is not our intent to provide a detailed or chronological review of all the work already reported on this topic, as it is well beyond the scope of a single chapter, the salient points alone are repeated.

Using a simplified scheme, Milly (1993) derived an analytical expression of the asymptotic soil moisture probability distribution (pdf). Milly's model, however, lacks the dependence of evapotranspiration on the soil water level and therefore cannot account for the impact of soil-moisture dynamics on vegetation (Milly, 2001). This problem was addressed by Rodriguez-Iturbe et al. (1999) and Laio et al. (2001a) and further pursued by Daly et al. (2004a,b) and Porporato et al. (2004), who modeled the soil water losses from drainage and evapotranspiration using a nonlinear function of soil moisture (e.g., Federer, 1979; Cordova and Bras, 1981; Gollan et al., 1985; Crago and Brutsaert, 1992), aimed at reproducing the effect of reduced soil water potential on water movement through the plant and the highly nonlinear form of deep infiltration (Daly et al., 2004a).

In what follows, we will briefly present the approach of Rodriguez-Iturbe et al. (1999), Laio et al. (2001a), and Porporato et al. (2004) as it combines a simplified approach that allows analytical solutions to a realistic description of rainfall stochasticity and nonlinearity in soil water losses.

According to this approach, the arrival of rainfall events is described as a Poisson process, assuming that each event carries a random amount of water extracted from an exponential distribution with mean α . During rainfall events, the soil accumulates water by infiltration to successively lose it by means of evapotranspiration and leakage. Hence, for simplicity, the water balance of the so-called hydrologically active soil is re-written as

$$nZ_r \frac{ds_m}{dt} = I[s_m(t), t] - E[s_m(t), t] - L[s_m(t)] \quad (29.8)$$

where $I[s_m(t), t]$ is infiltration from rainfall given by the rainfall rate minus runoff and canopy interception ($\overline{q(0)}$), $E[s_m(t), t]$ is evapotranspiration, and $L[s_m(t)]$ is deep infiltration or drainage ($\overline{q(Z_r)}$). Since $I[s_m(t), t]$ depends on soil moisture and rainfall variability, it represents the state-dependent stochastic forcing of the system, while evapotranspiration and drainage are nonlinear function of soil moisture content.

Infiltration depth into the soil is taken to be the minimum of the rainfall amount, h_r , and $nZ_r(1 - s_m)$, to reflect the fact that only a fraction of h_r can infiltrate when the rainfall amount exceeds the storage capacity of the soil column. Excess rainfall is considered to produce runoff according to the mechanism of saturation from below, which is a reasonable assumption in vegetated areas (no soil crust) with negligible-to-mild topography.

Canopy interception, which is responsible for the capture and subsequent direct evaporation of some of the precipitation prior to its arrival at the soil surface, is included in the model by assuming a threshold of rainfall depth, Δ (not to be confused with the slope of the saturation vapor pressure Δ_s) below which no water effectively penetrates the canopy. Then, analytically, the only change necessary is to modify the rate

of storm arrivals, λ (not to be confused with the latent heat of water vaporization, λ_w), to $\lambda' = \lambda e^{-\Delta/\alpha}$ (see Rodriguez-Iturbe et al. (1999) for details).

The link between daily evapotranspiration and soil moisture can be obtained by upscaling the fast dynamics of evapotranspiration (e.g., Daly et al., 2004a). The resulting behavior can be approximated by three regimes: a linear increase from $E = 0$ at the hygroscopic point, s_h , to E_w , at the wilting point, s_w (soil evaporation regime). For larger values of s , $E(s)$ has a linear rise (stressed evapotranspiration regime) from E_w at s_w to E_{max} at s^* , where s^* is the soil moisture level at which the plant begins to close stomata in response to water stress and E_{max} is the climate- and vegetation-dependent maximum daily evapotranspiration rate. For values of s exceeding s^* , evapotranspiration is decoupled from soil moisture and remains constant at E_{max} (unstressed evapotranspiration regime), which represents the average daily evapotranspiration rate during the growing season under well-watered conditions. Recall that the maximum transpiration contribution E_{max} can be theoretically determined by “slaving” the water loss to the biochemical carbon demand through leaf photosynthesis as earlier described. Assuming no interaction with the underlying soil layers and water table and the absence of significant topographic gradients, $L(s)$ represents vertical percolation forced by a unit gradient and is equal in magnitude to the unsaturated soil hydraulic conductivity. For reasons of analytical tractability, Laio et al. (2001a) adopted an exponential function (over the customary power law) increasing from 0 at field capacity, s_{fc} , to the saturated hydraulic conductivity at $s = 1$. The sum of the evapotranspiration and drainage losses is shown in Figure 29.8; its behavior is similar to the ones employed in previous studies (Cordova and Bras, 1981) and measured in field experiments (Salvucci, 2001). The values of s_h , s_w , s^* , and s_{fc} are related to the corresponding soil matric potentials through the soil–water retention curves.

The model allows one to obtain analytically the steady state (i.e., under the assumption of a homogeneous growing season) soil moisture pdf as a function of the parameters defining climatic conditions, soil type, and vegetation characteristics. As an example, Figure 29.9 shows an analysis of the role of changes in the frequency of storm events λ on the soil moisture pdf. A coarser soil texture corresponds to a consistent shift of the pdf toward drier conditions. The shape of the pdf also undergoes marked changes and resembles the one empirically measured elsewhere (e.g., Salvucci, 2001). The broadest pdfs are found for shallower soils.

Several applications demonstrate the good adaptability of the model to very different environmental conditions, from semiarid, temperate, to tropical (Salvucci, 2001; Laio et al., 2001b; Fernandez-Illescas et al., 2001; Porporato et al., 2003a). Other applications of the probabilistic approach to soil-moisture dynamics have regarded the seasonal water balance (Laio et al., 2002), as well as the role of inter-annual climate variability (D’Odorico et al., 2000; Porporato et al., 2003a). Guswa et al. (2002) also explored the importance of the assumption of vertically averaging the soil moisture equation, noting that the main differences are in the transpiration temporal dynamics in case of deep-rooted plants with little active control on their water uptake from different soil layers (see also Guswa et al. (2004) and Puma et al. (2005)).

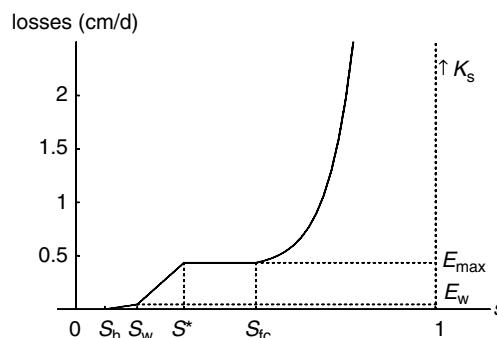


FIGURE 29.8 Soil water losses as a function of relative soil moisture content. (From Laio, F., et al. *Adv. Water Resour.*, 24, 707, 2001a, with permission.)

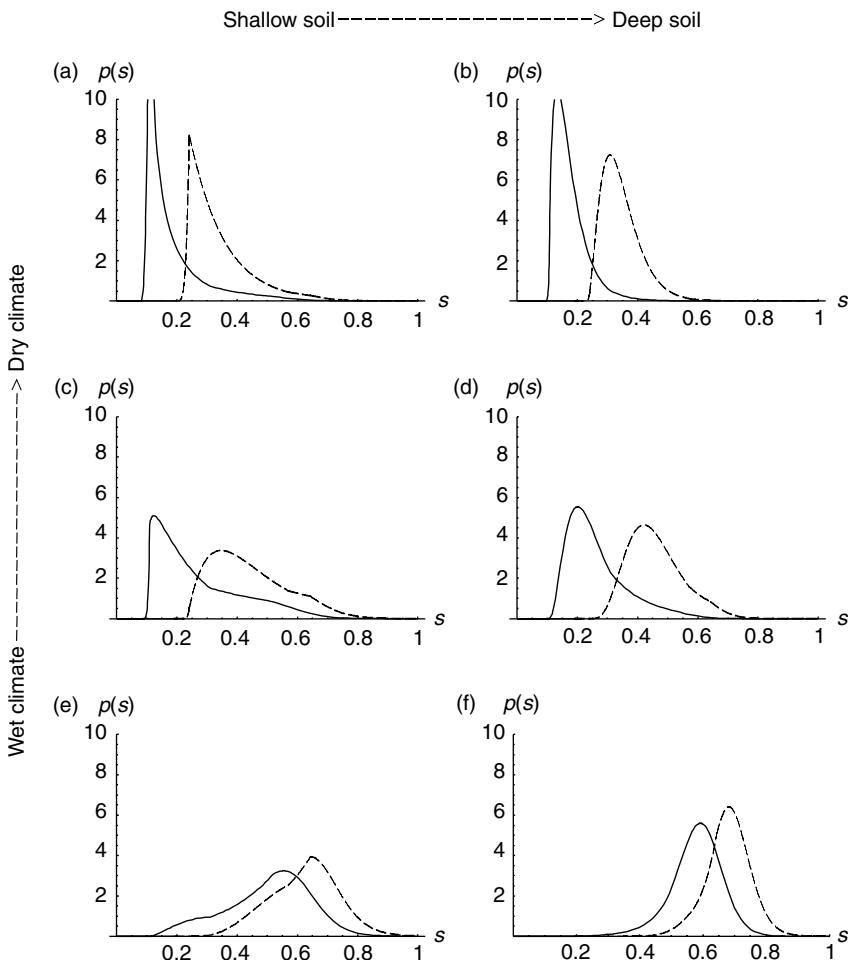


FIGURE 29.9 Examples of pdf's of relative soil moisture for different type of soil, soil depth, and mean rainfall rate. Continuous lines refer to loamy sand, dashed lines to loam. Left column corresponds to $Z_r = 30\text{ cm}$, right column to 90 cm . Top, center, and bottom graphs have a mean rainfall rate λ of 0.1 , 0.2 , and 0.5 d^{-1} respectively. Common parameters to all graphs are $\alpha = 1.5\text{ cm}$, $\Delta = 0\text{ cm}$, $E_w = 0.01\text{ cm/d}$, and $E_{\max} = 0.45\text{ cm/d}$. (From Laio, F., et al. *Adv. Water Resour.*, 24, 707, 2001a, with permission.)

29.3.3 Long-Term Water Balance

The different terms of the long-term water balance are the averages of the respective components of the soil-moisture dynamics. Rainfall is first partitioned into interception, runoff, and infiltration. Infiltration is then divided into leakage and evapotranspiration. For purposes of describing vegetation conditions, the amount of water transpired can be further divided into water transpired under stressed conditions and water transpired under non-stressed conditions (e.g., for soil moisture below and above s^* , respectively; see Section 29.4).

The long-term water balance of the stochastic model presented above can be obtained analytically by multiplying the terms of Equation 29.8 by the steady state soil moisture pdf and then integrating between 0 and 1 (e.g., Rodriguez-Iturbe et al., 1999). Figure 29.10 presents examples of the dependence of the water-balance components — normalized by the mean rainfall rate — on some characteristics of rainfall regime, soil, and vegetation. The influence of the rainfall rate λ is shown in Figure 29.10a for the case of a shallow loam. As interception is a linear function of the rainfall rate, it is not surprising that the fraction

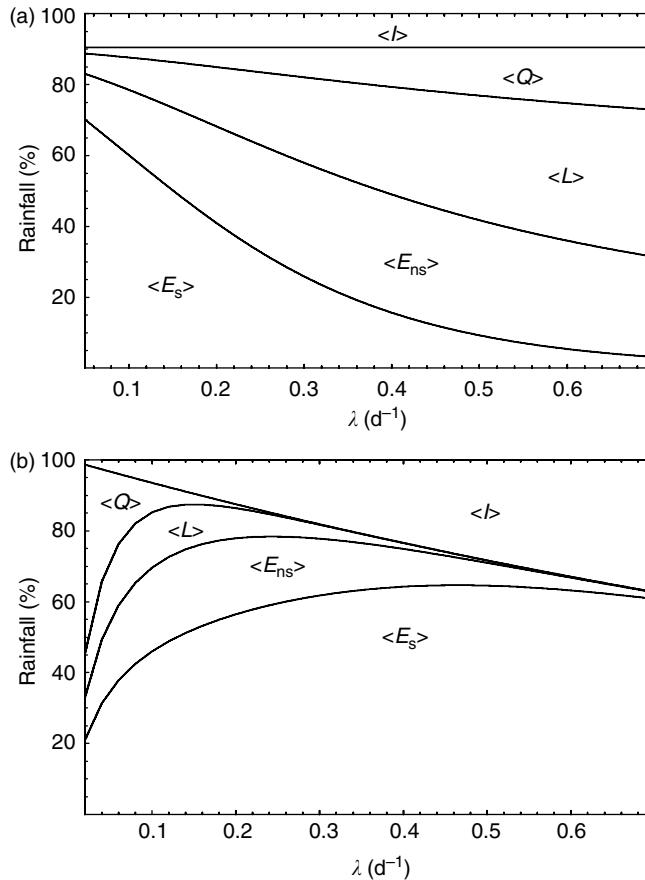


FIGURE 29.10 Components of the water balance normalized by the total rainfall for a shallow loamy sand. (a) Water balance as a function of the frequency of rainfall events, λ ($Z_r = 30$ cm, $\alpha = 2$ cm). (b) Water balance as a function of λ for a constant mean total rainfall during a growing season of 60 cm. Parameters are $s_h = 0.08$, $s_w = 0.11$, $s^* = 0.31$, $s_{fc} = 0.52$, $E_w = 0.01$ cm/d, $E_{max} = 0.45$ cm/d, and $\Delta = 0.2$ cm. $\langle I \rangle$, $\langle Q \rangle$, $\langle L \rangle$, $\langle E_{ns} \rangle$, and $\langle E_s \rangle$ are the long-term averaged interception, runoff, leakage, evapotranspiration in non-stressed and stressed conditions, respectively. (From Laio, F., et al. *Adv. Water Resour.*, 24, 707, 2001a, with permission.)

of water intercepted remains constant in the normalized equation. The percentage of runoff, however, increases almost linearly. More interesting is the interplay between leakage and the two components of evapotranspiration. The fraction of water transpired under stressed conditions rapidly decreases, while the evapotranspiration under non-stressed conditions evolves in a much more gradual manner. As discussed by Porporato et al. (2001), this last aspect has interesting implications for vegetation productivity. It is clear that in semiarid conditions most of the water that actually reaches the soil is lost by evapotranspiration (in particular transpiration), consistent with many field experiments. Figure 29.10b shows the terms of the water balance when both the frequency, λ , and the average storm depth, α , change but the total amount of rainfall in a growing season (i.e., $\alpha\lambda$) is kept constant. The result is interesting, due to the existence of two opposite mechanisms regulating the water balance. On the one hand, for a given seasonal amount of rain, runoff production strongly depends on the ratio between soil depth and the mean depth of rainfall events. On the other hand interception increases almost linearly with λ . The interplay between these two mechanisms determines a maximum of both leakage and evapotranspiration at moderate values of λ (of course, the position of the maxima changes according to the parameters used). This interplay is particularly important to ecosystem productivity given the intimate linkage between transpiration,

photosynthesis, and carbon gain (Kramer and Boyer, 1995). Not only the amount but also the frequency of rainfall is important for the existence of an optimum transpiration (a surrogate for productivity).

29.3.4 Heat Flow into the Soil

Given that all the kinetic rates and respiration terms in the soil carbon–nitrogen dynamics, which will be described in detail in Section 29.5, depend on soil temperature (in addition to soil moisture), we end this section with a brief summary on modeling soil temperature dynamics in depth and time.

The heat flow into the soil is highly impacted by soil moisture and the incident radiation and leaf area density. The heat flow conduction in a soil is given by

$$q_{h,i}(t, x_i) = -K_h \frac{\partial T_t(t, z)}{\partial x_i}$$

where $q_{h,i}$ is the heat flux along direction x_i ($i = 1, 2, 3$), T_t is the soil temperature, and K_h is the thermal conductivity of the soil, which varies with the mineral type and soil moisture. The energy conservation in a soil layer can be expressed as

$$\rho_t c_m \frac{\partial T_t}{\partial t} = -\frac{\partial q_{h,i}}{\partial x_i}$$

with ρ_t soil density, and c_m the soil specific heat capacity per unit mass. In the case of vertical heat flow the balance reduces to

$$\frac{\partial T_t}{\partial t} = \frac{1}{\rho c_m} \frac{\partial}{\partial z} \left(K_h \frac{\partial T_t}{\partial z} \right)$$

where z in this case is directed upward.

The simplest method to determine “effective” thermal properties of the bulk soil is from its constitutive components using a volume-based aggregation scheme,

$$\begin{aligned} K_h &= f_a K_{h,a} + \theta K_{h,w} + (1 - f_a - \theta) K_{h,s} \\ \rho c_m &= f_a (\rho_a c_{m,a}) + \theta (\rho_w c_{m,w}) + (1 - f_a - \theta) (\rho_s c_{m,s}) \end{aligned}$$

where subscripts a, w, and s represent the air, water, and solid components of the soil matrix, respectively, and f_a is the air-filled porosity. Roughly, the value of the parameters is $K_{h,s} = 2.9 \text{ W m}^{-1} \text{ K}^{-1}$, $K_{h,w} = 0.57 \text{ W m}^{-1} \text{ K}^{-1}$, and $K_{h,a} = 0.025 \text{ W m}^{-1} \text{ K}^{-1}$, $\rho_a c_{m,a} = 1.25 \times 10^3 \text{ J m}^{-3} \text{ K}^{-1}$, $\rho_w c_{m,w} = 4.2 \times 10^6 \text{ J m}^{-3} \text{ K}^{-1}$, and $\rho_s c_{m,s} = 2.0 \times 10^6 \text{ J m}^{-3} \text{ K}^{-1}$. A volume-based parameter averaging may not be valid for the effective thermal properties in a heterogeneous soil with dense rooting system. This, in part, can be attributed to the temperature difference between the precipitation water and the soil and the rooting system.

We tested this aggregation methodology in a grass-covered forest clearing in which the soil heat flux at the ground surface was measured by soil heat flux plates every 20 min, and soil temperature and soil moisture were measured at depths of $-2, -5, -10, -15$, and -30 cm . Figure 29.11 shows the measured soil heat flux and the measured soil temperature, and Figure 29.12 compares measured and modeled soil temperatures using the volume-based aggregation scheme (the porosity of the soil also varied with z and was taken as the maximum measured θ in the 1-year time series) using the one-dimensional version of the soil heat flow model. The agreement between measured and modeled fluxes is rather encouraging — at least for the purposes of modeling microbial activity and respiration rates.

29.4 Soil Moisture Deficit and Plant Water Stress

In many ecosystems, and especially in arid and semiarid climates, soil moisture deficit is often the most important stress factor for vegetation. In fact, other important sources of stress such as nutrient limitation

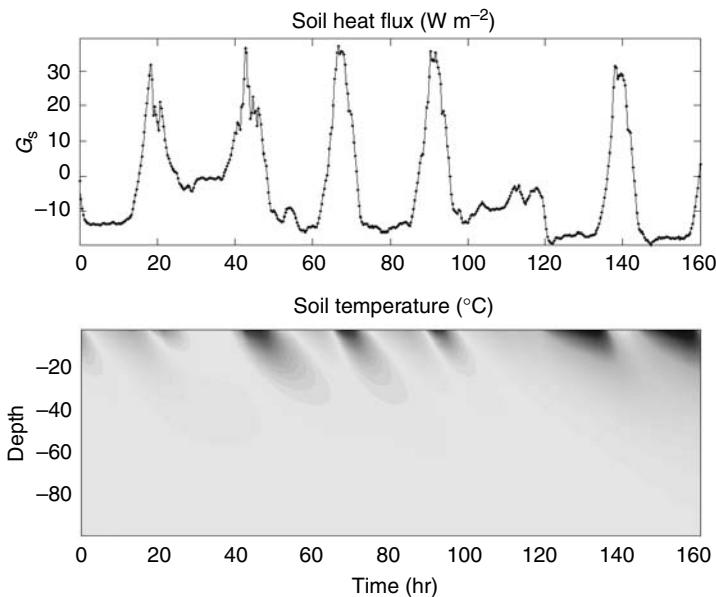


FIGURE 29.11 (See color insert following page 23-52) Measured soil heat flux (G_s) near the land-atmosphere interface and measured soil temperature (ranging from 12, black, to 18°C, white) within a grass-covered surface illustrating the warming and cooling heat waves.

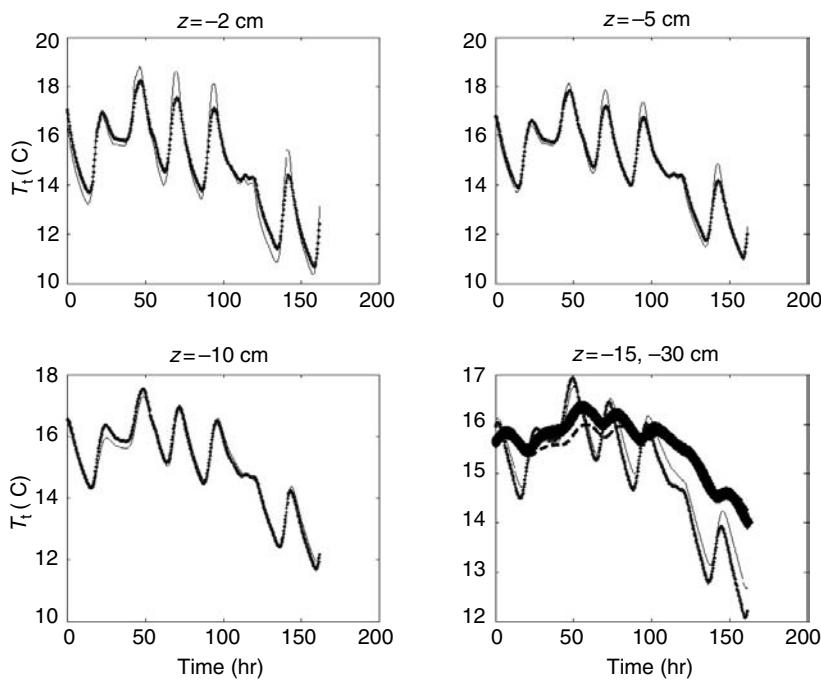


FIGURE 29.12 Comparison between measured (symbols) and modeled (line) soil temperature at different soil layer depths (z) for a grass-covered forest clearing at the Blackwood division of the Duke Forest, near Durham, NC. The “forcing” input is the measured soil heat flux at $z = 0$ (shown in Figure 29.11). The locally measured soil moisture content at each depth was used in the volume aggregation scheme of model parameters.

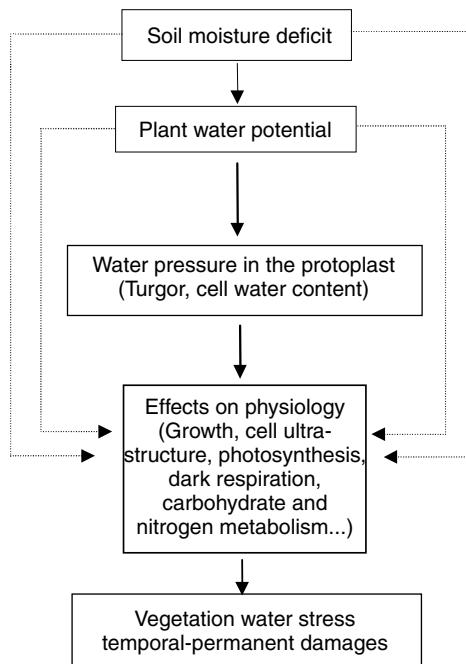


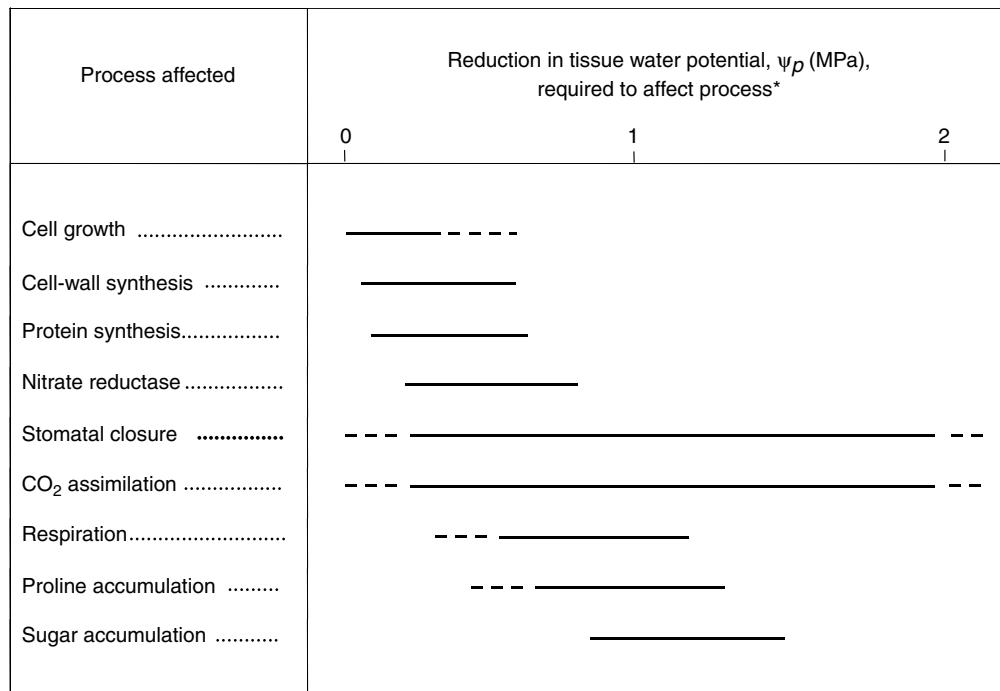
FIGURE 29.13 Schematic representation of the process linking soil moisture deficit to plant water stress. (Modified from Porporato, A., et al. *Adv. Water Resour.*, 24, 725, 2001.)

are frequently initiated by the occurrence of water stress (e.g., Larcher, 1995; Nilsen and Orcutt, 1998). A large number of studies on physiological plant ecology has dealt with the problem of plant water stress (e.g., Hsiao, 1973; Lange et al., 1976; Levitt, 1980; Bradford and Hsiao, 1982; Smith and Griffith, 1993; Larcher, 1995; Ingram and Bartels, 1996; Nilsen and Orcutt, 1998). Porporato et al. (2001) provided a simplified mathematical description of plant water stress and linked it to the model of stochastic soil-moisture dynamics previously described.

29.4.1 Physiological Impact of Soil Moisture Deficit on Plants

As noted earlier, plants need to maintain an adequate level of hydration in their tissues to ensure both growth and survival; they also require a continuous flux of water from the soil to perform vital processes such as photosynthesis and nutrient uptake. The reduction of soil moisture content during droughts lowers the plant water potential and leads to a decrease in transpiration. This in turn causes a reduction of cell turgor and plant water content, which brings about a sequence of damages of increasing seriousness. Figure 29.13 presents a simplified scheme of the steps linking soil moisture deficit to plant water stress.

Figure 29.14 reproduces a typical sequence of effects caused by a decrease in cell water potential. As can be seen, during drought conditions this decrease triggers a series of harmful events on plant physiology whose number and seriousness grows with the intensity and duration of water deficit. As the plant water potential is reduced, the first effect is a reduction of cell growth and wall-cell synthesis; then nitrogen uptake from soil is diminished due to the sharp decrease in nitrate reductase, an enzyme that catalyzes the reduction of nitrate to nitrite (which is the first internal step for nitrogen assimilation after nitrate has been taken up by the plant). The stomatal closure and the related reduction in CO₂ assimilation also follows: at first, stomatal closure is mostly a way to reduce water losses while maintaining carbon uptake and does not produce permanent damage to the plant. Soon after, however, respiration is affected. Changes in the patterns of resource allocation are then induced, including the accumulation of amino acid proline, whose increase is a characteristic disturbance in protein metabolism. Sugar accumulation, along with possible



* With ψ_p of well-watered plants under mild evaporative demand as the reference point

FIGURE 29.14 Typical sequences of plant physiological effects caused by a decrease in cell water potential. Redrawn after Hsiao, T.C. *Ann. Rev. Plant Physiol.*, 24, 519, 1973.

flowering reduction, inhibition of seed production, and fruit abortion are possible further effects of higher levels of water stress (Nilsen and Orcutt, 1998). At the same time, reduction in transpiration inhibits the cooling effect and radiation stress more likely. Finally, at very low water potentials, complete stomatal closure occurs followed by turgor loss and wilting (Bradford and Hsiao 1982, p. 309) as well as widespread cavitation.

The interplay between cavitation in the xylem system and turgor pressure remains a subject of active research. This, in part, is due to the fact that stomatal closure often occurs at leaf pressures much less negative than the pressure of turgor loss. Recent studies on plant hydraulics seem to indicate that evolution of vascular plants was associated with some loss of “lower plants” to survive desiccation (Sperry, 2000), as processes associated with the evolution of a larger and more complex plant body became more limiting to gas exchange than those occurring within individual cells (Sperry, 2000; Katul et al., 2003). The vascular tissue is a vital whole-plant process that can be impaired by water supply to the photosynthetic tissue. As soil water deficit develops, events in plant and soil make it increasingly difficult to keep the hydraulic pipelines between soil and leaf hydraulically intact. The transport characteristics of this network of pipelines impose physical limits on the rate at which water can be supplied to the leaves, and on the potential rate of transpiration allowed by stomata (Tyree and Sperry, 1989; Bohrer et al., 2005). In a recent review, Sperry (2000) suggested that much of the stomatal closure observed during drought is a result of a decline in plant hydraulic conductance in the xylem of the root system because roots can be substantially more susceptible to cavitation than shoots. This explains why stomatal conductance shuts down at suction levels well below the turgor suction. On re-watering, there is also the potential for considerable hysteresis in the recovery of conductance to “pre-drought” values. Recovery can occur by refilling of cavitated vessels by root pressure (short time scale), and by growth of new roots (intermediate time scale).

Stomatal closure is thus an important process for the description of plant response to water deficit. It is controlled by the pressure of the guard cells surrounding the stomata: high guard-cell turgor produces

stomatal opening, while low turgor induces stomatal closure (e.g., Salisbury and Ross, 1992; Larcher, 1995). Stomatal closure and xylem cavitation are the main mechanisms that lead to a nonlinear relationship between bulk canopy conductance (and thus transpiration) and soil water content. Furthermore, this relationship need not be unique at a given soil moisture state because of this hysteresis — which is driven by soil moisture history and carbon uptake.

The sequence of events shown in Figure 29.14 is related to the intensity of the soil moisture deficit without specifically considering its duration. It is clear that this “static” description of the physiological effects is not sufficient to describe the development of plant suffering; water stress is a gradually intensifying process in which the time dimension is particularly important. This is especially true for semiarid ecosystems, where drought is often a prolonged and frequent phenomenon, more than an isolated event. Duration and frequency of periods of water stress are therefore as important as the reduction in water potential to model plant water stress.

29.4.2 Modeling Plant Water Stress

Porporato et al. (2001) used a “static” vegetation water stress index to relate the actual value of soil moisture to two levels of the same variable associated with important changes in the physiological activities of the plant, namely the point at which transpiration (and thus photosynthetic activity) starts being reduced, s^* , and the point at which plants begin to wilt, s_w . These two levels correspond to the point of incipient and complete stomatal closure, respectively. Static stress is assumed to be equal to zero when soil moisture is above s^* and equal to one when soil moisture is at or below s_w . In between these soil moisture levels, the static stress is given by

$$\zeta = \left(\frac{s - s_w}{s^* - s_w} \right)^q$$

where q is a measure of the nonlinearity of the effects of soil moisture deficit on plant conditions (not to be confused with the water flux q_i). This simple relationship between ζ and s permits the derivation of the pdf of static stress as a derived distribution of the steady state pdf of soil moisture. Using this relationship, the mean value of water stress $\bar{\zeta}$, given that the plant is under stress, can be derived (Porporato et al., 2001).

Besides the information given by $\bar{\zeta}$, the dynamical aspects of the water deficit process, namely the duration and frequency of the stress periods, should also be taken into account to describe the plant stress conditions. To this purpose, Porporato et al. (2004) derived the expressions for the mean duration, \bar{T}_ξ , of an excursion below an arbitrary soil moisture threshold, ξ , as well as the mean number, \bar{n}_ξ , of such intervals during the growing season. Using s^* as a meaningful threshold to mark the onset of water stress, the stress intensity, its mean duration, and its frequency of occurrence may be combined to define an overall indicator of plant conditions under given edaphic and climatic factors, called the dynamic plant water stress, as

$$\bar{\vartheta} = \begin{cases} \left(\frac{\bar{\zeta} \bar{T}_{s^*}}{k T_s} \right)^{\bar{n}_{s^*}^{-r}} & \text{if } \bar{\zeta} \bar{T}_{s^*} > k T_{\text{seas}} \\ 1 & \text{otherwise} \end{cases} \quad (29.9)$$

where k is an indicator of plant resistance to water stress and T_{seas} is the duration of the growing season and r is a nonlinearity coefficient to account for the importance of sequence of stress events. The reader is referred to Porporato et al. (2001) for the derivation and discussion of this dynamic water stress function.

29.4.3 Soil-Climate-Vegetation Control on Plant Water Stress and Plant Carbon Assimilation

The above formulation of the total dynamic water stress permits a quantitative study of its dependence on climate, soil, and vegetation. As an example, Figure 29.15 shows the interplay between the timing and amount of rainfall, similar to Figure 29.10 where the storm frequency λ is varied while fixing the

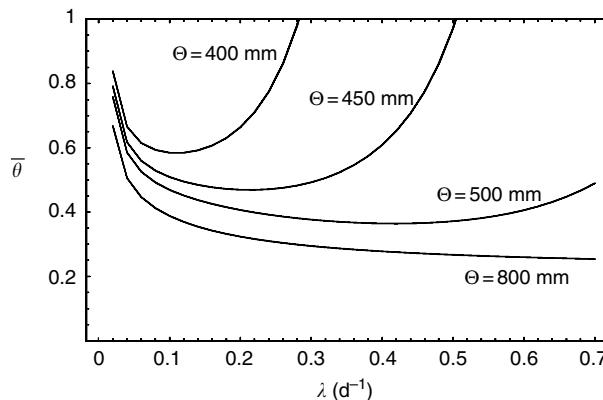


FIGURE 29.15 Impact of timing and amount of rainfall on dynamic water stress. Each curve corresponds to a constant total rainfall, Θ , during the growing season, $q = 2$, $k = 0.5$, and $r = 0.5$. Parameters: $T_{\text{seas}} = 200$ d, $Z_r = 60$ cm, $E_{\max} = 0.45 \text{ cm d}^{-1}$, $s_w = 0.24$, $s^* = 0.57$, $s_c = 0.65$. (Modified from Porporato, A., et al. *Adv. Water Resour.*, 24, 725, 2001.)

total precipitation per growing season. Except for very wet climates, there exists a clear optimal trade-off between timing and amount of rainfall, which provides the best condition for vegetation (minimum of the dynamic water stress). The minima decrease and move toward the right-hand side of the diagram for larger values of total rainfall. Such evidence provides interesting connotation to the concept of effective rainfall (i.e., an event that is intense enough to stimulate biological processes, particularly growth and reproduction), and whose importance is well known in the ecology of arid and semiarid ecosystems (e.g., Noy-Meir, 1973). The dynamic water stress defined above has been used in Porporato et al. (2003a) to describe the tree-grass coexistence in the Kalahari transect.

Porporato et al. (2004) and Daly et al. (2004b) have extended it to include plant carbon assimilation. In particular, Figure 29.16, shows a comparison of the experimental results with the theoretical mean carbon assimilation as a function of the frequency of rainfall events for fixed total rainfall during a growing season in mesic grassland in Kansas (Knapp et al., 2002). The 20% decrease in measured net assimilation for the altered rainfall pattern (when total rainfall is the same but concentrated in fewer events) is well reproduced. As shown by the effective relative soil moisture pdfs, the shift in the rainfall frequency changes qualitatively the soil moisture pdf and thus the grassland water balance. The figure also shows that in such an ecosystem, the impact on carbon assimilation of a decrease in total rainfall is more pronounced when such a decrease is accompanied by a reduction in the frequency of rainfall events.

29.5 Soil Nutrient Dynamics

Having completed the precipitation and vegetation controls on soil-moisture dynamics, we proceed next to soil biogeochemical cycles. While the former is critical to modeling groundwater recharge, the latter exerts significant controls on groundwater quality. Soil biogeochemical cycles are characterized by complex dynamics, acting at different spatial and temporal scales, and remain crucial for modeling groundwater recharge (on multiple time scales) and quality. They are impacted by vegetation and hydro-meteorological forcing and, in turn, exert various feedbacks on the ecosystems, atmosphere, and climate dynamics. In particular, the hydro-climatic variability and soil-moisture dynamics regulate the sequence of fluxes between different components of soil nutrient cycles and determine the temporal dynamics of the system state variables at different time scales. Moreover, the soil moisture and temperature regimes control decomposition, leaching, and plant uptake and, indirectly, influence vegetation growth and composition of plant residues.

Various models have been proposed in the past to investigate these dynamical processes for a wide range of ecosystems and climatic conditions (e.g., Parton et al., 1988; Hunt et al., 1991; Melillo et al., 1995;

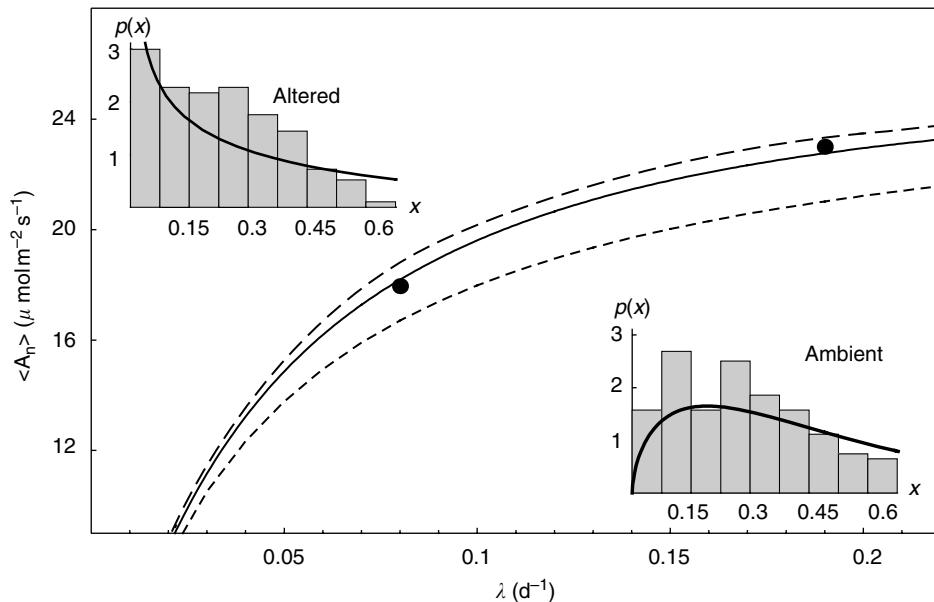


FIGURE 29.16 Mean daily carbon assimilation rate as a function of the frequency of rainfall events for constant total amount of precipitation during a growing season. The lines are the theoretical curves derived from the soil moisture probability density function, while the two points are field data published by Knapp et al. (2002). The point on the right corresponds to the ambient conditions ($\lambda = 0.19 \text{ d}^{-1}$), and the other point to artificially modified conditions ($\lambda = 0.08 \text{ d}^{-1}$) while keeping the total rainfall the same. The continuous line is for mean total rainfall during a growing season of 507 mm, the dashed line for 600 mm, and the dotted line for 400 mm. The two insets show observed and theoretical soil moisture probability density functions for ambient and altered conditions ($E_{\max} = 0.63 \text{ cm d}^{-1}$, $s_h = s_w = 0.25$, $s^* = 0.65$, $s_f = 0.8$, $n = 0.55$, $Z_r = 30 \text{ cm}$). Revised from Porporato, A., Daly, E., and Rodriguez-Iturbe, I. *Am. Nat.*, 164, 625, 2004.

Gusman and Marino, 1999; Benbi and Richter, 2002; Porporato et al., 2003b; Schimel and Weintraub, 2003; Schimel and Bennett, 2004, and references therein).

29.5.1 Brief Review of Soil Carbon and Nitrogen Cycles

The carbon cycle in the soil-plant system is dominated by a sequence of biochemical processes including assimilation (photosynthesis), residues production and deposition, and decomposition (litter mineralization), which releases carbon as CO_2 back to the atmosphere (Figure 29.17). Assuming a long-term equilibrium condition of ecosystems, the total respiration is generally balanced by the input of plant residues while at shorter time scales (e.g., seasonal-to-interannual) the carbon content is subject to fluctuations induced by hydro-climatic and other environmental forcing. Carbon is stored in the SOM, which is a complex and varied mixture of organic substances, with three main components: plant residues, microbial biomass, and humus. Plant residues accumulate on the soil surface where they are degraded by both physical and biological agents. The organic compounds are then partially assimilated and partially oxidized by the microbial biomass in the soil. The biomass provides a high turnover rate of C and N in the SOM pool, with release of stable mineralized products of the catabolic processes (e.g., CO_2) at each turnover (Molina and Smith, 1998).

The carbon assimilation and respiration processes are affected by the composition of organic compounds and by climatic conditions. The presence of water and elevated soil temperatures are key factors to provide a fast decomposition and biomass growth. The relation between biomass activity and soil water content is strongly nonlinear (Brady and Weil, 1996). In particular, low soil moisture levels reduce diffusion of nutrients, limiting their availability to microbes, and lower hydration, and enzymatic

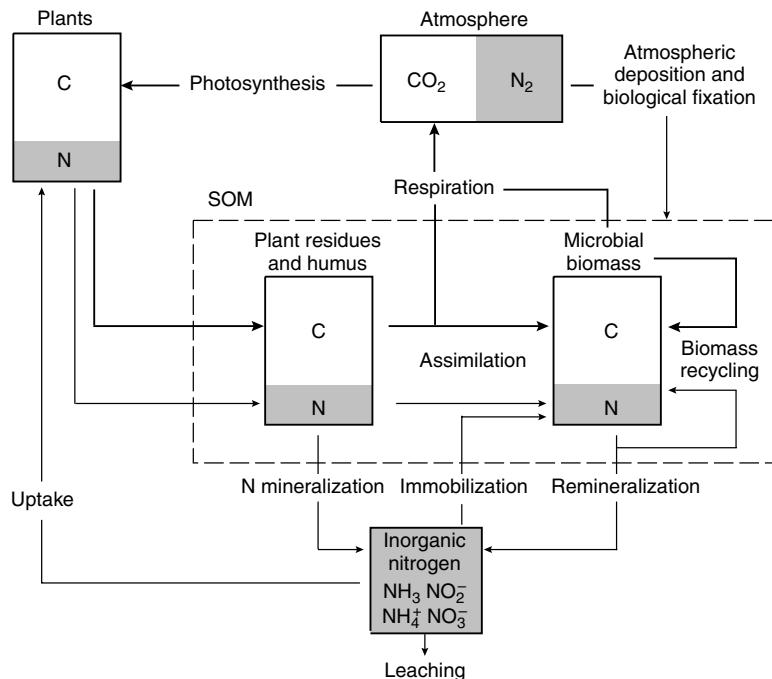


FIGURE 29.17 Schematic representations of the soil carbon and nitrogen cycles, with particular attention to the interactions among the three main components of the SOM: plant litter, humus, and biomass. Thick lines: C fluxes; thin lines: N fluxes. (Modified after Mary, B., et al. *Plant Soil*, 181, 71, 1996.)

activity of microbes causing water stress. Interestingly, the water stress of soil microbes has characteristics similar to plant water stress. The dependence of decomposition on soil temperature is also nonlinear, rising from zero at low temperatures ($\sim -5^\circ\text{C}$) to a maximum at around 40°C (although these values tend to be site-specific). When the conditions are favorable, nitrification is quite rapid. As a result, in hot and dry environments, sudden water availability can cause a flush of soil nitrate production (see Figure 29.18), which may greatly influence the growth patterns of natural vegetation (Cui and Caldwell, 1997; Fierer and Schimel, 2002). Depletion of soil mineral nitrogen, through plant uptake and deep percolation (leaching), also depends nonlinearly on soil moisture. As described in Porporato et al. (2001), plant nitrogen uptake is typically a nonlinear function of soil moisture with some similarities with the transpiration function. Leaching is an increasing nonlinear function of soil water content (related to soil hydraulic conductivity) and occurs intermittently following intense rainfall events.

The so-called Mineralization-Immobilization Turnover (MIT) and the C/N ratio of microbial biomass are also very important for the entire soil biogeochemistry. To maintain a constant C/N ratio during its growth, the microbial biomass needs to assimilate proportional amounts of carbon and nitrogen, independently of the composition of the organic compounds of the substrate (Brady and Weil, 1996). As a consequence, if the nitrogen content of the decomposed compounds is high, mineralization proceeds unrestricted and mineral components in excess are released into the soil, while if such compounds are nitrogen poor, the microbes immobilize mineral nitrogen for their growth (Figure 29.18). If mineral nitrogen is not available, immobilization is halted. The MIT is thus characterized by an extremely strong nonlinearity that acts as a threshold-process and determines the conditions of when the growth of soil microbial biomass is carbon or nitrogen limited.

29.5.2 Model Structure

Various models of different complexity have been proposed in the literature and a full review would be outside of the scope of this chapter. Here, we provide a brief description of the “low-dimensional” model

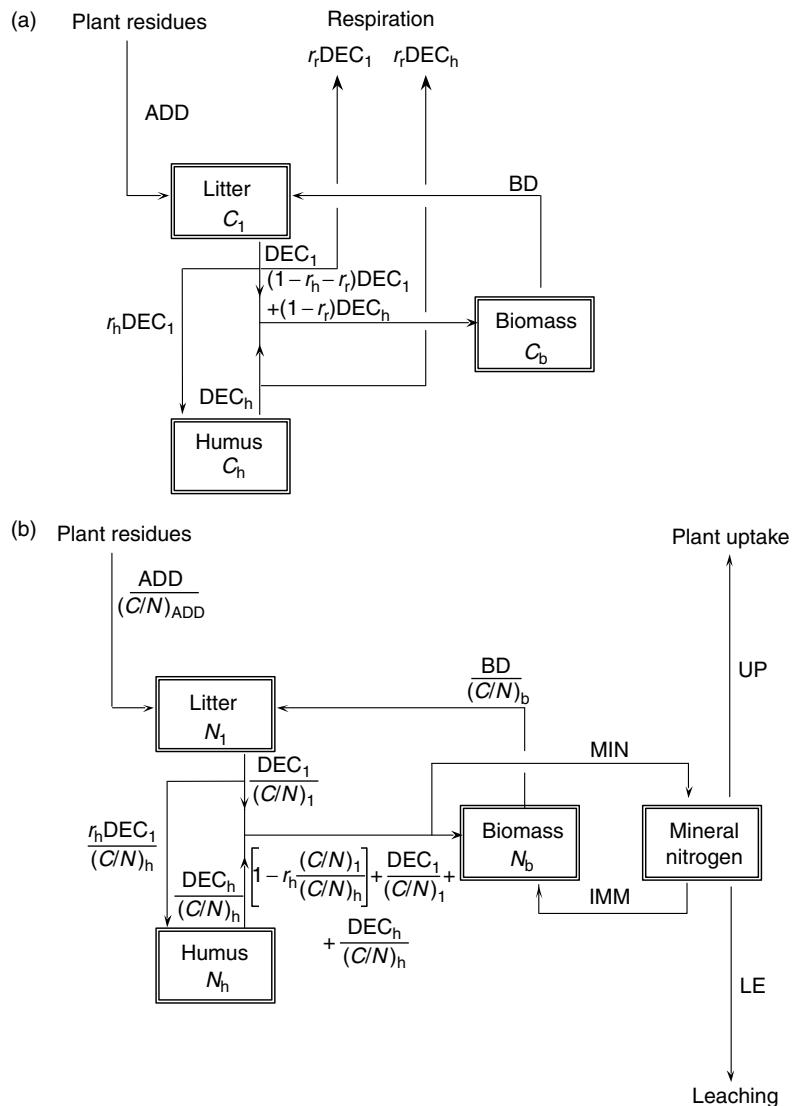


FIGURE 29.18 Schematic representation of the main components of the model. (a) Soil carbon cycle; (b) soil nitrogen cycle. (Modified from Porporato, A., et al. *Adv. Water Res.*, 26, 45, 2003b.)

proposed by Porporato et al. (2003b) because of its simple yet realistic structure of the soil C-N cycles for vertically averaged concentrations over the active soil depth, Z_t .

The carbon and nitrogen cycles are modeled using four different pools, one for each main component of the system (Figure 29.18). The only external input of carbon and nitrogen to the system is through vegetation litter. Respiration is the only carbon output, while the nitrogen losses are leaching and plant uptake. Fluxes such as ammonium adsorption and desorption, deposition, and volatilization are neglected, being of lesser importance for the soil C-N balances at the daily-to-seasonal time scale (Mary et al., 1996; Baisden and Amundson, 2003; D'Odorico et al., 2003). SOM is divided into three compartments, representing litter, humus, and microbial biomass. For simplicity, we lump ammonium and nitrate into a single mineral nitrogen pool.

Once plant residues enter the litter pool, they move partially to humus and partially to the biomass pool, losing a respiration fraction during the decomposition process. The approach here differs from the widely

used CENTURY model (Parton et al., 1988) because decomposition is explicitly considered to account for the fact that it also depends on the amount of biomass, in agreement with Schimel and Weintraub (2003). The model requires seven state variables, each in terms of mass per unit volume of soil (e.g., g m^{-3}), six of them to describe carbon and nitrogen concentrations in the three SOM pools, and one for the soil mineral nitrogen (see the scheme in Figure 29.18): C_l , carbon concentration in the litter pool; C_h , carbon concentration in the humus pool; C_b , carbon concentration in the biomass pool; N_l , organic nitrogen concentration in the litter pool; N_h , organic nitrogen concentration in the humus pool; N_b , organic nitrogen concentration in the biomass pool; N , mineral nitrogen concentration in the soil.

The temporal dynamics of such variables is controlled by a system of seven coupled differential equations that describe the balance of carbon and nitrogen in the various pools and the fluxes among them, in terms of mass per unit volume per unit time (e.g., $\text{g m}^{-3}\text{d}^{-1}$). As many of the fluxes are controlled by environmental fluctuations, their modeling explicitly accounts for soil moisture and temperature conditions.

The system of first order nonlinear differential equations in describing these processes is given by

$$\begin{aligned}\frac{dC_l}{dt} &= \text{ADD} + \text{BD} - \text{DEC}_l \\ \frac{dN_l}{dt} &= \frac{\text{ADD}}{(C/N)_{\text{ADD}}} + \frac{\text{BD}}{(C/N)_b} - \frac{\text{DEC}_l}{(C/N)_l} \\ \frac{dC_h}{dt} &= r_h \text{DEC}_l - \text{DEC}_h \\ \frac{dN_h}{dt} &= r_h \frac{\text{DEC}_l}{(C/N)_h} - \frac{\text{DEC}_h}{(C/N)_h} \\ \frac{dC_b}{dt} &= (1 - r_h - r_r) \text{DEC}_l + (1 - r_r) \text{DEC}_h \\ \frac{dN_b}{dt} &= \left[1 - r_h \frac{(C/N)_l}{(C/N)_h} \right] \frac{\text{DEC}_l}{(C/N)_l} + \frac{\text{DEC}_h}{(C/N)_h} - \frac{\text{BD}}{(C/N)_b} - \Phi \\ \frac{dN}{dt} &= \Phi - \text{LE} - \text{UP}\end{aligned}$$

The first six equations represent the balances of litter (l subscript), humus (h subscript), and biomass (b subscript) pools, while the last one is the balance equation for the mineral nitrogen, N . A brief description of the terms is given next.

The term ADD is the external input into the system, representing the rate at which carbon in plant residues is added into the soil and made available to the microbial colonies. Here ADD is assumed to be constant in time. The term BD is the rate at which carbon returns to the litter pool due to the death of microbial biomass. Porporato et al. (2003b) simply used a linear dependence on the amount of microbial biomass, that is, $\text{BD} = k_d C_b$.

DEC_l represents the carbon output due to microbial decomposition, modeled using a first order kinetics with respect to the carbon concentration in the litter pool, C_l , as well as to the carbon in the biomass pool, C_b ,

$$\text{DEC}_l [\varphi k_l f(s, T) C_b] C_l$$

where the coefficient φ is a nondimensional factor accounting for a possible reduction of the decomposition rate when the litter is very poor in nitrogen and immobilization is not sufficient to integrate the nitrogen required by bacteria. The constant k_l defines the rate of decomposition for the litter pool as a weighted average of the decomposition rates of the different organic compounds in the plant residues. Its average value is usually much higher than the corresponding value for humus, k_h . The close relationship between the decomposition rate and the microbial biomass is due to the role of exoenzymes in the depolymerization process. The factor $f(s, T)$ describes soil moisture (s) and soil temperature (T) effects on decomposition. The dependence of microbial activity on soil temperature is described by a

quadratic relation, while the soil moisture control on aerobic microbial activity and decomposition is modeled via a linear increase from s_b up to field capacity, s_{fc} , and a hyperbolic decrease from there to soil saturation. The parameter s_b represents a sort of permanent “wilting” point for soil microbial biomass and its relationship with the plant wilting point, s_w , that defines the level at which transpiration and passive nitrogen uptake are stopped, provides a way to account for the impact of water stress on the competition for nutrients between plants and soil microbial biomass (Kaye and Hart, 1997).

The second equation in the system represents the nitrogen balance in the litter pool. This is similar to that of carbon, with each term divided by the C/N ratio of its respective pool. $(C/N)_{ADD}$ is the C/N ratio of added plant residues, whose variability can produce pronounced changes in the C/N ratio of the litter pool and on the MIT turnover. The balance equation for carbon in the humus pool has a single input flux, represented by the fraction r_h of the decomposed litter undergoing humification. The input of carbon in the biomass pool is represented by the fraction of organic matter incorporated by the microorganisms from litter and humus decomposition. The constant r_r defines the fraction of decomposed organic carbon that goes into respiration, usually estimated to be in the interval 0.6 to 0.8 (Brady and Weil, 1996), while Φ is the net mineralization defined so to ensure that $(C/N)_b$ is constant. The explicit modeling of mineralization/immobilization dynamics is quite complicated and the reader is referred to Porporato et al. (2003b) for details.

The last equation of the system describes the balance of mineral nitrogen, in which mineralization is the only input, while plant uptake (UP) and leaching (LE) are the two losses. The latter is simply modeled as proportional to the deep infiltration $L(s)$ losses through a solubility coefficient (see Porporato et al., 2003b), while the plant uptake involves a passive and an active process, which can be regarded as additive processes. The passive uptake is assumed to be proportional to the plant transpiration rate and to the nitrogen concentration in the soil solution. The active uptake is assumed to help compensate the nitrogen deficit only if the passive uptake is lower than a given plant demand.

29.5.3 System Behavior under Constant and Stochastic Soil Moisture Conditions

In steady-state conditions, the SOM carbon and nitrogen concentrations can be analytically determined from the full system (e.g., Manzoni et al., 2004). The steady solutions for humus and biomass nitrogen are simply proportional to the humus and biomass carbon, according to their C/N ratios. At steady state, C_b is the only state variable that depends on the climate forcing, while the other SOM pools accumulate carbon and nitrogen when the decomposition efficiency is low and lose them when the decomposition efficiency is high (i.e., near field capacity) and under elevated soil temperature conditions.

Figure 29.19 shows the dynamics of carbon biomass toward equilibrium as a function of soil moisture and under constant temperature conditions (for reference). The trajectories are obtained by numerically integrating the system using a standard ordinary differential equation solver. At very low soil moisture, decomposition is not sustained and equilibrium is degenerate. Following the usual classification of the different equilibrium points of dynamical systems, equilibrium is a stable focus with damped oscillation to equilibrium. If the soil moisture approaches field capacity, a dynamic bifurcation takes place and in the region where the decomposition is more efficient, the equilibrium point becomes a stable node. The situation is inverted at high soil moisture values where damped oscillations leading to a stable focus reappear. The damped oscillations are determined by the dead biomass recycling in the litter pool, which introduces a feedback in the system, and by the biomass control on decomposition. The duration of these oscillations ranges between 1 and 8 years, depending on soil moisture level and system initial conditions.

D’Odorico et al. (2003) analyzed the C-N dynamics under the more realistic conditions of stochastic soil moisture forcing using the model described in the previous section for the case study of the water-limited savanna of Nysvley in South Africa. As shown in Figure 29.20, depending on the inertia of the various pools and on the degree of dependence on soil moisture, the random fluctuations imposed by precipitation are filtered by the temporal dynamics of the state variables in a very interesting manner. Some variables (e.g., NO_3^-) preserve much of the high-frequency variability imposed by the random

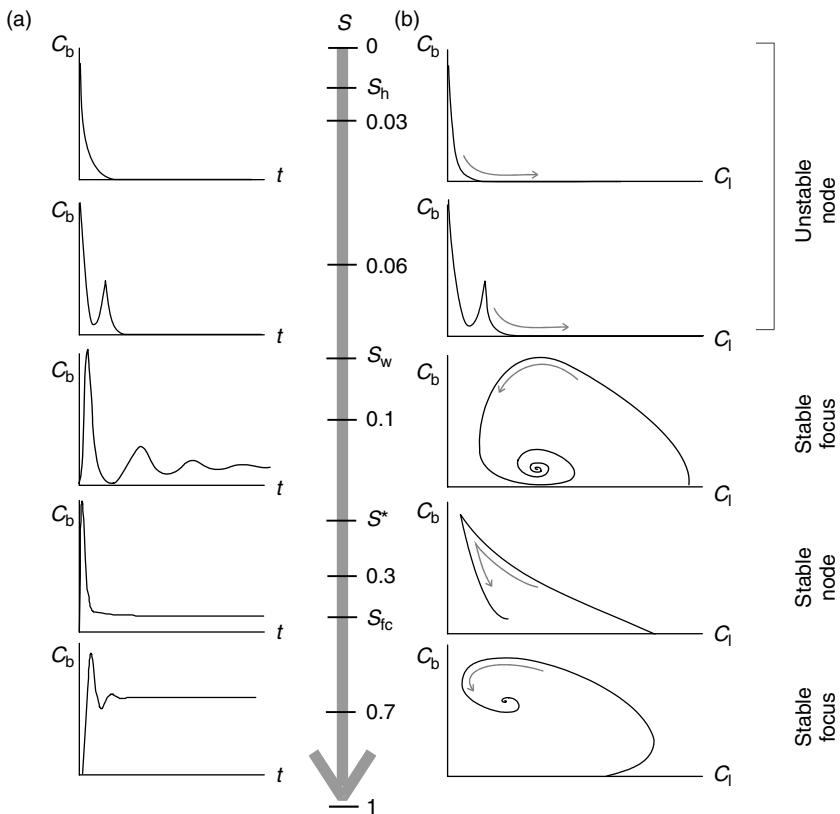


FIGURE 29.19 Qualitative representation of the temporal dynamics of biomass (plots in the left column) and the biomass-carbon litter trajectories (plots in the right column) for different soil moisture values.

forcing of precipitation, while some others (C_h , C_l , and C_b) show smoother fluctuations (Figure 29.20 and Figure 29.21). Nitrate dynamics is the final product of a number of intertwined processes in which both high- and low-frequency components interact. In particular, the high-frequency component of NO_3^- fluctuations (period of days to weeks) can be linked to the direct dependence of mineralization and nitrification on soil moisture (panels e and f), which transfers the random fluctuations of the rainfall forcing to the budget of nitrate. On the other hand, the low-frequency variability (period of seasons to years) resembles the one of organic matter (in particular of microbial biomass; e.g., panel d) and depends on the inertia imposed on the dynamics by the dimension of the soil carbon and nitrogen pools, which is very large compared to the fluxes (see Figure 29.21). Notice that the same low frequency component also characterizes the litter dynamics, which is negatively correlated to C_b (see Figure 29.21b) as the growth of one of these pools occurs at the expense of the other one.

The different time scales occurring in the carbon and nitrogen cycles were analyzed by means of the power spectra of the different variables. Figure 29.22 shows the logarithmic plot of the normalized spectral density of s , C_b , and NO_3^- . As expected, the energy associated to the high frequencies is higher for soil moisture than for nitrate and microbial biomass, while the converse is true for the low frequencies. The crossing between the C_b , (or NO_3^-) and s power spectra is located at frequencies corresponding to periods of seasons to years. The model thus reproduces the change in time scales occurring from the soil-moisture dynamics to the nutrient dynamics observed by Schimel et al. (1997). This indicates that, while in semiarid ecosystems soil moisture is not able by itself to provide memory to the system at scales larger than one year (due to the complete depletion of soil water content by the end of the dry season, for example (Nicholson, 2000)), nutrient and vegetation dynamics may have a much longer memory responsible for

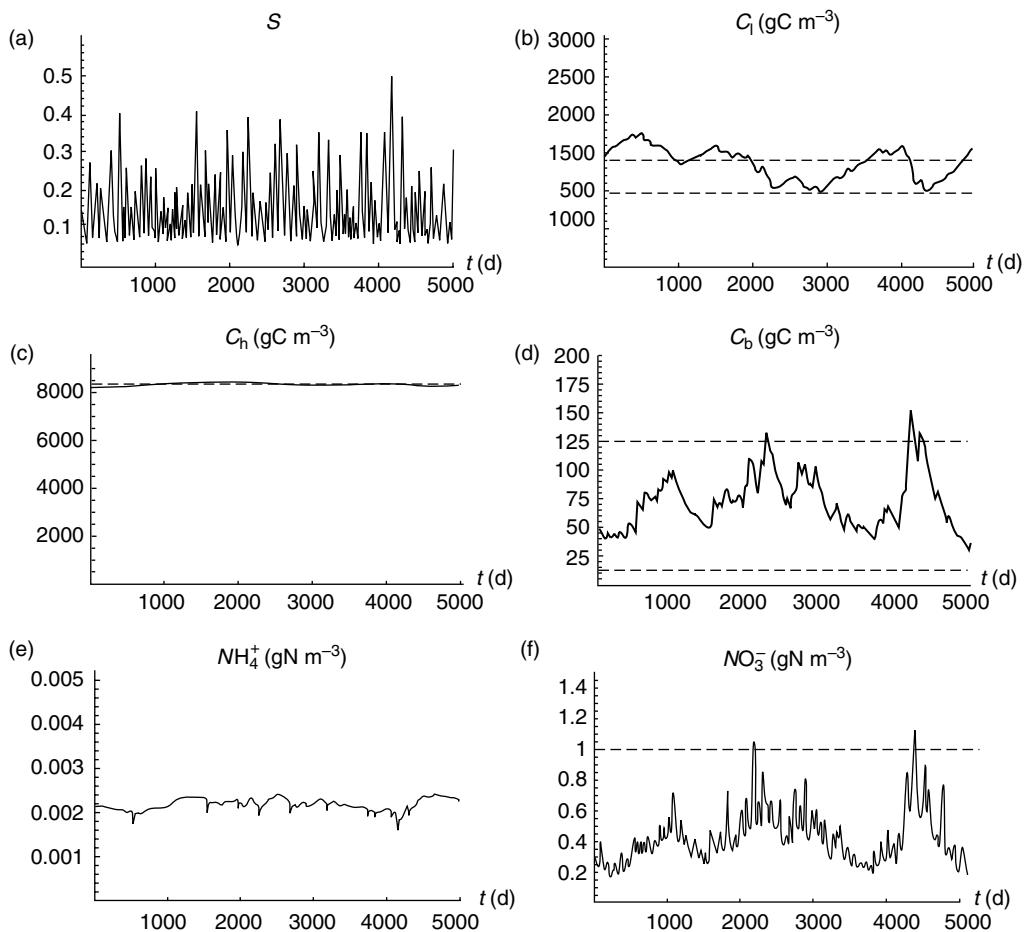


FIGURE 29.20 Temporal dynamics of relative soil moisture (a), carbon litter (b), carbon humus (c), carbon biomass (d), ammonium (e), and nitrate (f) simulated for the case of the broad-leaved savanna at Nylsvley. The broken lines represent the average values or the range of values observed at Nylsvley except in (f), where the broken line is an upper limit for nitrate concentration in the nitrogen-poor savanna. The broken lines are reported whenever observations are available. Parameters: $n = 0.4$, $s_{\text{fc}} = 0.3$, $s_h = 0.02$, $K_s = 1.1 \text{ m d}^{-1}$, $Z_r = 0.80 \text{ m}$, $\alpha = 11 \text{ mm}$, $\lambda = 0.23 \text{ d}^{-1}$, $E_{\text{max}} = 4.5 \text{ mm d}^{-1}$, $\text{Add} = 1.5 \text{ gC m}^{-2} \text{ d}^{-1}$, $s^* = 0.17$, $s_w = 0.065$, $(C/N)_{\text{add}} = 58$, $(C/N)_{\text{bio}} = 11.5$, $(C/N)_h = 22$. (From D'Odorico, P., et al. *Adv. Water Res.*, 26, 59, 2003, with permission.)

the interannual persistency observed both in hydro-climatic and ecosystem processes (Parton et al., 1988; Schimel et al., 1996, 1997).

A closer inspection of the time series of soil moisture and nitrate reveals that such a phenomenon is due to the presence of sudden flushes of nitrate, following a prolonged wet period after a drought (see Figure 29.23). In these conditions, first the dry soil hinders decomposition and favors SOM accumulation, then the subsequent wet period elicits biomass growth and enhances mineralization (e.g., Cui and Caldwell, 1997; Scholes et al., 1997). Episodic changes in nitrate levels greatly influence plant growth, because plants' response to increased availability of nitrogen tends to be very quick (e.g., Brady and Weil, 1996). These pulses of nitrate are therefore of considerable importance for natural ecosystems.

From the previous results, it is clear that the high frequency variability of soil moisture and the low frequency variability of SOM combine to produce complex temporal dynamics of soil mineral nitrogen, while the nonlinear mutual interactions between the processes may either enhance or reduce the effect of changes in the climatic regime. Moreover, the occurrences of exceptional hydrologic conditions, such as

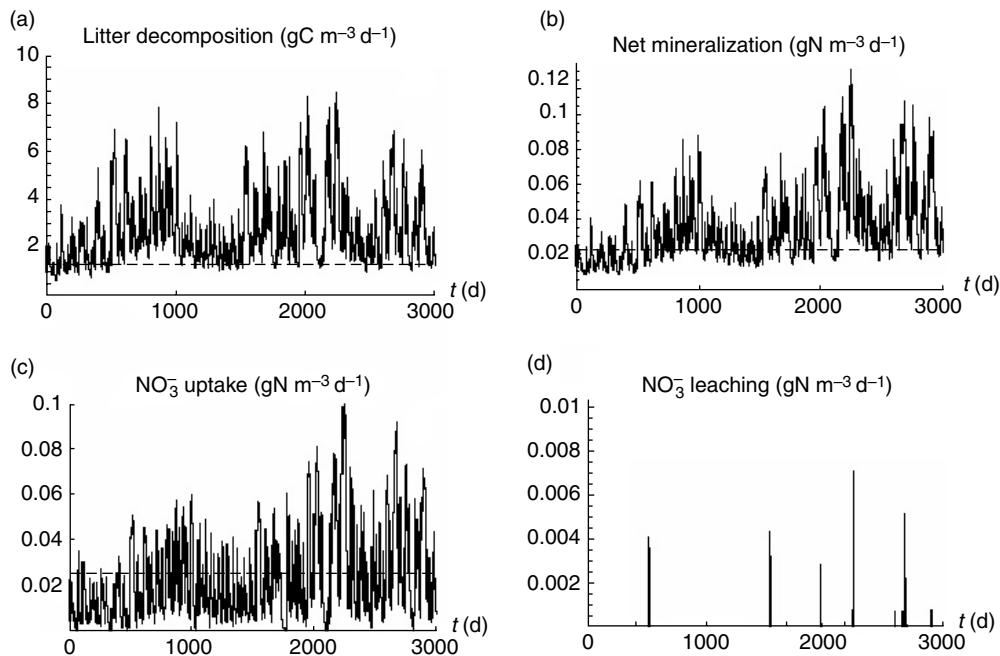


FIGURE 29.21 Simulated rates of litter decomposition (a), net mineralization (b), nitrate uptake (c) and nitrate leaching (d) in the broad-leaved savanna at Nylsvley. The broken lines represent the average values observed at Nylsvley. The nitrate leaching (d) is reported to be negligible at Nylsvley. (From D’Odorico, P., et al. *Adv. Water Res.*, 26, 59, 2003, with permission.) Parameters as in Figure 29.20.

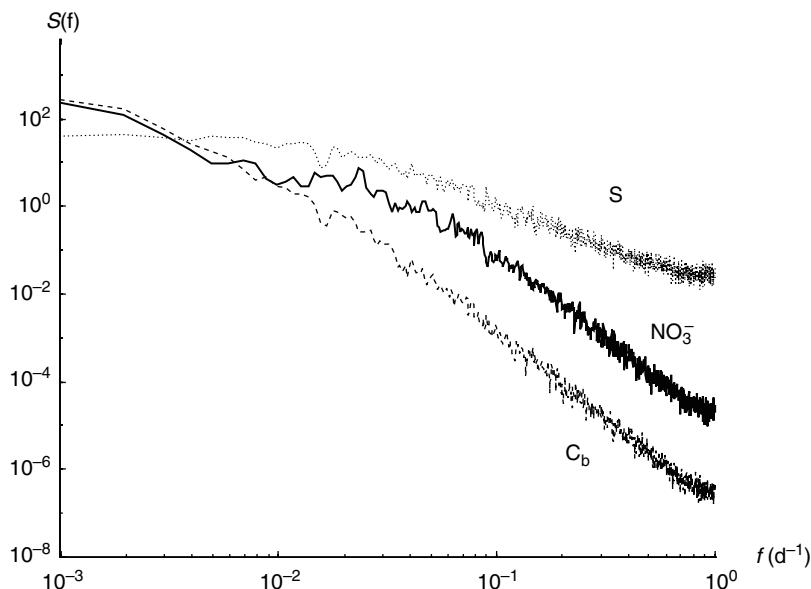


FIGURE 29.22 Power spectra of soil moisture, nitrate and microbial biomass. (From D’Odorico, P., et al. *Adv. Water Res.*, 26, 59, 2003, with permission.) Parameters as in Figure 29.20.

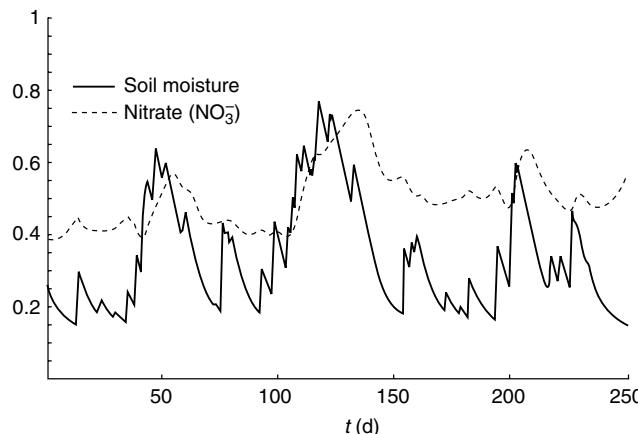


FIGURE 29.23 Effects of a prolonged wet period on the soil organic matter and nitrate dynamics. (From D’Odorico, P., et al. *Adv. Water Res.*, 26, 59, 2003, with permission.) Parameters as in Figure 29.20.

prolonged rainfall or drought, are capable of affecting the dynamics of the slow-varying pools inducing long-lasting effects.

29.5.4 Soil Organic Matter and Soil Hydraulic Properties

In the soil biogeochemical cycle models, the soil hydraulic properties are assumed to be static. However, there are significant (yet indirect) effects of SOM dynamics on the hydrological processes. The most crucial of them is the role of SOM (especially humus and soil microbial biomass) on changes in hydraulic properties such as porosity, hydraulic conductivity, and soil water retention functions. Despite intense experimental and theoretical research, a clear link between results from field and laboratory experiments and detailed dynamical description of SOM evolution and interaction with the soil matrix remains elusive and the long-term projections of soil biogeochemistry models still rely on semi-empirical relationships whose range of applicability and physical underpinning are un-assessed.

The interaction between SOM and soil-moisture dynamics depends on the relationship between (1) hydraulic properties and geometric structure of the soil matrix, (2) organic matter and soil matrix, and (3) soil water content and microbial activity in the SOM.

The hydraulic properties of soils are inseparable from their physical/geometrical characteristics. Thus the link between Pore Size Distribution (PSD) functions and Water Retention Curves (WRC) has received increasing attention in soil physics also in relation to pedo-transfer functions (e.g., Wosten et al., 2001; Aubertin et al., 2003; Rawls et al., 2003), stochastic and fractal models of soil structure (Tyler and Weatherford, 1990; Rieu and Sposito, 1991; Hueckel et al., 1997; Chan and Govindaraju, 2003), and percolation theory (e.g., Hunt and Gee, 2002).

SOM plays a significant role in determining the chemical and physical properties of soils because of its biotic and abiotic shares (Sollins et al., 1996; Neff et al., 2003). From the recent review by Six et al. (2004), it appears that the conceptual advances in understanding the soil matrix structure (microaggregate theory, aggregate hierarchy theory, flocculation, etc.) were not equally accompanied by an improved description of SOM–soil matrix interaction. On the other hand, the literature in soil bioremediation and soil sciences has shown how SOM and microorganisms growth determines substantial changes in the PSD and permeability of saturated soils. Using micro-models of triangular pore networks, Kim and Fogler (2000) found steep decreases in relative permeability with increasing SOM content. In similar experiments, Stewart and Kim (2004) found that microorganisms cluster thereby inducing periodic dynamics, alternating proliferation and starvation, hence causing periodic clogging of soil microporosity. Two mechanisms for bioclogging of saturated porous media have been investigated in Thullner et al. (2002): biofilm and colony. The results showed that colony-type development has more important consequences on

the PSD than the biofilm-type. Furthermore, they found that increased pore-size heterogeneity yields faster bioclogging. Exoenzymes produced by bacteria appear to be responsible for the increase in microbial volume contributing to bioclogging (Schimel and Weintraub, 2003). Relatively simple models have been proposed to describe the effects of bacterial microfilm and aggregates on soil hydraulic properties (e.g., Taylor et al., 1990; Vandevivere, 1992).

In unsaturated soils, Rawls et al. (2003) have shown that SOM somewhat tempers the soil hydraulic behavior, increasing the WRC in sandy soils and lowering it in fine-textured soils. As the WRC is sensitive to the PSD, it is inferred that SOM affects the structural organization of the pore network. Indirectly, also infiltration, evaporation, deep percolation, and plant water and nutrient uptake depend on the SOM dynamics with important consequences on the soil water balance and related processes. The effects of SOM in unsaturated soils range from modifying fluid properties such as viscosity, surface tension, and wetting characteristics, to reducing the effective porosity and PSD via microfilms and aggregates (Rockhold et al., 2002). Thus a comprehensive model requires a dynamic microscopic description of SOM growth and decay and its physical properties coupled with the detailed representation of the complex geometry of the soil matrix. As pointed out by Rockhold et al. (2002) modeling of these processes in unsaturated soils is still in its infancy, and is an emerging frontier in biogeoscience research.

29.6 Summary and Conclusions

The compass of this chapter was on the description and the mathematical formulation of temporal dynamics of water and carbon inputs from the surface and of their regulation of percolation and biogeochemistry of the soil-plant system within the vadose zone. Our intent was not to cover all the above issues in detail, but we aimed at providing a mathematical framework that can guide groundwater hydrologists and hydrogeologists on various approaches used to quantify the carbon-nitrogen-water fluxes through the soil-plant-atmosphere system.

The chapter commenced with the mathematical formulation for water-carbon cycling above ground as a means of determining the maximum CO_2 uptake and hence maximum water loss by plants (i.e., water is slaved to the carbon needs of the plant) based on leaf area density, physiological, drag, and radiative properties of the plants, and forced by standard meteorological drivers, such as air temperature, mean wind speed above the canopy, vapor pressure deficit, incident shortwave radiation (or photosynthetically active radiation), and atmospheric CO_2 concentration. This framework is timely, given the recent advances in remote sensing products (e.g., MODIS platform) to map leaf area over large spatial scales and at time scales relevant to plant phenology (e.g., weeks), and the availability of higher resolution re-analysis meteorological data. We then explored how soil moisture controls reduce these rates (i.e., the carbon uptake now becomes “slaved” to the soil-plant hydraulics). We showed how to upscale these processes in time “prognostically” by assuming precipitation as a random forcing with known statistics. This approach is much more suited to questions relevant to long-term carbon sequestration in future climate as the precise time series of precipitation, the key driver of the hydrological cycle, is unknown. We then examined how the basic elements of the soil biogeochemical cycle, including nitrate leaching (relevant to groundwater quality), is impacted by the soil-plant characteristics and precipitation statistics. We also discussed (more accurately speculated on) some long-term feedbacks from the soil-biogeochemical cycles on soil (and possibly plant) hydraulics, rarely investigated in current biogeochemistry or hydrology models. The latter feedbacks remain a knowledge gap and are likely to gain broader attention in the near future.

List of Symbols

$a(z)$	$\text{m}^2 \text{ leaf m}^{-2} \text{ ground m}^{-1}$	leaf area density
a_l	$\text{m}^2 \text{ leaf m}^{-2} \text{ ground}$	cumulative leaf area density
ADD	$\text{g m}^{-3} \text{ d}^{-1}$	soil carbon input rate
A_n	$\text{mol m}^{-2} \text{ s}^{-1}$	net leaf-scale assimilation

BD	$\text{g m}^{-3}\text{d}^{-1}$	soil carbon input rate due to microbial death
$\overline{C_a}$	$\text{mol CO}_2 \text{ mol}^{-1} \text{ air}$	mean atmospheric carbon dioxide concentration
C_b	g m^{-3}	biomass carbon concentration
C_h	g m^{-3}	humus carbon concentration
$\overline{C_i}$	$\text{mol CO}_2 \text{ mol}^{-1} \text{ air}$	intercellular stomata carbon dioxide concentration
C_l	g m^{-3}	litter carbon concentration
c_m	$\text{J kg}^{-1} \text{ K}^{-1}$	soil specific heat capacity
C_p	$\text{J kg}^{-1} \text{ K}^{-1}$	specific heat of air
DEC	$\text{g m}^{-3}\text{d}^{-1}$	soil carbon output rate
D_v	Pa	vapor pressure deficit
d	m	Zero-plane displacement height
e_m		maximum quantum efficiency for CO_2 uptake
E	cm d^{-1}	Evapotranspiration rate
E_{\max}	cm d^{-1}	potential evapotranspiration rate
E_w	cm d^{-1}	soil evaporation at s_w rate
f_a		air filled soil porosity
F_x	$[x] \text{ t}^{-1} \text{ m}$	turbulent vertical fluxes of x
g_b	$\text{mol m}^{-2}\text{s}^{-1}$	mean boundary layer conductance
G_s	W m^{-2}	Soil heat flux at the surface
g_s	$\text{mol m}^{-2}\text{s}^{-1}$	mean stomatal conductance
g_v	$\text{kg m}^{-2} \text{ s}^{-1}$	water vapor conductance
h	m	canopy height
h_c	$\text{W m}^{-2} \text{ K}^{-1}$	convection coefficient
h_r	cm	rainfall amount
h_s	$\text{g water g}^{-1} \text{ air}$	mean air relative humidity near the leaf surface
J_c	$\text{mol m}^{-2} \text{ s}^{-1}$	assimilation rate limited by Rubisco
J_E	$\text{mol m}^{-2} \text{ s}^{-1}$	assimilation rate limited by light
k		plant resistance to water stress
K_{be}		leaf extinction coefficient
K_h	$\text{W m}^{-2} \text{ K}^{-1}$	soil thermal conductivity
k_x		rate of decomposition for x pool
LAI	$\text{m}^2 \text{ leaf m}^{-2} \text{ ground}$	leaf area index
LE	$\text{g m}^{-3}\text{d}^{-1}$	leaching
n		soil porosity
\bar{n}_ξ		mean number of excursions below ξ in a growing season
N_b	g m^{-3}	biomass organic nitrogen concentration
N_h	g m^{-3}	humus organic nitrogen concentration
N_l	g m^{-3}	litter organic nitrogen concentration
p_a	Pa	atmospheric pressure
Q_{ab}	W m^{-2}	canopy absorbed energy
Q_p	$\text{mol m}^{-2} \text{ s}^{-1}$	leaf PAR irradiance
q		static water stress parameter
r_b	$\text{m}^2 \text{ s mol}^{-1}$	mean boundary layer resistance
R_d	$\text{mol m}^{-2} \text{ s}^{-1}$	leaf respiration rate
r_h		decomposed litter fraction undergoing humidification
r_r		organic carbon fraction going into respiration
r_s	$\text{m}^2 \text{ s mol}^{-1}$	Mean stomatal resistance
\bar{q}_a	$\text{g water g}^{-1} \text{ air}$	Mean atmospheric water vapor
s'		turbulent concentration fluctuation for an arbitrary scalar
s_m		relative soil moisture
s^*		value of s_m below which water stress occurs

s_{fc}		s_m at field capacity
s_h		s_m at the hygroscopic point
s_w		s_m at the wilting point
S_x	$[x] t^{-1}$	sink/source of x
T_a	K	mean air temperature
T_ξ	d	mean time duration of an excursion below ξ in a growing season
T_r	$cm d^{-1}$	actual transpiration rate
$T_{r,max}$	$cm d^{-1}$	potential transpiration rate
T_s	K	leaf surface temperature
T_t	$^{\circ}C$	Soil temperature
T_{seas}	d	growing season duration
U, v, w	$m s^{-1}$	wind speed components along x, y , and z directions
u_*	$m s^{-1}$	friction velocity
UP	$g m^{-3}d^{-1}$	nitrogen plant uptake
V_{cmax}	$mol m^{-2} s^{-1}$	maximum Rubisco capacity
Z_r	cm	soil root depth
z_o	m	Momentum roughness length
α	cm	mean depth of rainfall events
α_l		leaf absorptance for PAR
ε	$m^2 s^{-3}$	turbulent kinetic dissipation
ε_l		leaf emissivity
κ		Von Karman constant ($=0.4$)
ζ		static plant water stress coefficient
ψ		zenith angle
θ		volumetric soil moisture
$\bar{\vartheta}$		dynamic plant water stress
λ	d^{-1}	mean rainfall frequency
λ_w	$J kg^{-1}$	latent heat of water vaporization
ρ_a	$kg m^{-3}$	air density
ρ_t	$kg m^{-3}$	soil density
σ	$W m^{-2} K^{-4}$	Stefan-Boltzman constant
σ_e	$m s^{-1}$	Velocity scale related to the turbulent kinetic energy
$\sigma_{u,v,w}$	$m s^{-1}$	Longitudinal, lateral, and vertical velocity standard deviations
τ	s	relaxation time
τ_b		incident beam radiation
ω		CO_2/O_2 specification ratio
Γ_*	$mol CO_2 mol^{-1} air$	compensation point
Δ	Cm	threshold for rain interception
Δ_s	$Pa K^{-1}$	slope of saturation vapor pressure-temperature
Φ	$g m^{-3}d^{-1}$	net mineralization rate
Θ	cm	total rainfall amount in a growing season
Ω		clumping factor of leaf distribution

Glossary

Available nutrient The portion of any element in the soil that can be adsorbed and assimilated by growing plants.

Available water The portion of water that can be freely up-taken by plant roots. Therefore, it is the amount of water between field capacity and wilting point.

Autotroph An organism able to use carbon dioxide or carbonates to obtain energy for life processes from the oxidation of inorganic elements.

Biomass The total mass of living material of a specified type in a given environment.

Carbon cycle The sequence of transformations through which carbon dioxide is fixed in living organisms by photosynthesis or chemosynthesis, is produced by respiration and by the death and decomposition of living organisms, is used by heterotrophic species and is returned to its original state.

Compensation point It is the carbon dioxide concentration at which plant carbon assimilation vanishes in the absence of photorespiration.

Decomposition Chemical breakdown of organic matter into simpler compounds, often accomplished by the aid of microorganisms.

Field capacity Water content remaining in a soil after it has been saturated and after free drainage is practically ceased.

Fixation Process or processes by which elements essential for vegetation growth are converted from a soluble or exchangeable form to a less soluble or non-exchangeable form. In particular, nitrogen is chemically combined with hydrogen to produce ammonia.

Groundwater Subsurface water in the saturated soil free to move under the influence of gravity.

Heterotroph Organism capable of deriving energy for life processes only from the decomposition of organic compounds.

Humus The partially stable fraction of the soil organic matter remaining after the major portions of added plant and animal residues have decomposed.

Immobilization The conversion of an element from the inorganic to the organic form in microbial or plant tissues, thus making the element not available to other organisms or plants.

Leaching The removal of materials in solution from the soil by percolating water.

Mineralization The conversion of an element from an organic form to an inorganic state as a result of microbial decomposition.

Nitrogen cycle The sequence of changes undergone by nitrogen as it moves from the atmosphere into water, soil, and living organisms and upon death of such organisms is recycled through a part or all of the entire process.

Runoff The portion of the precipitation on an area that is discharged from the area through stream channels. In soil science it usually only refers to the water lost by surface flow (e.g., surface runoff).

Vadose zone The aerated region of soil above the permanent water table.

Wilting point The soil-moisture content at which plants wilt.

References

- Aubertin, M., et al. A model to predict the water retention curve from basic geotechnical properties, *Can. Geotech. J.*, 40, 1104, 2003.
- Baisden, W.T. and Amundson, R. An analytical approach to ecosystem biogeochemistry modeling, *Ecol. Appl.*, 13, 649, 2003.
- Baldocchi D. and Meyers, T. On using eco-physiological, micrometeorological and biogeochemical theory to evaluate carbon dioxide, water vapor and trace gas fluxes over vegetation: A perspective, *Agric. Forest Meteorol.*, 90, 1, 1998.
- Benbi, D.K. and Richter, J. A critical review of some approaches to modeling nitrogen mineralization, *Bio. Fertil. Soils*, 35, 168, 2002.
- Bohrer, G., Mourad, H., Laursen, T.A., Drewry, D., Avissar R., Poggi, D., Oren, T., and Katul, G. Finite element tree crown hydrodynamics model (FETCH) using porous media flow within branching elements: A new representation of tree hydrodynamics, *Water Resour. Res.*, 41, W11404, doi:10.1029, 2005.
- Bonan, G., *Ecological Climatology: Concepts and Applications*, Cambridge University Press, Cambridge, 2002.

- Bradford, K.J. and Hsiao, T.C. Physiological responses to moderate water stress, in *Physiological Plant Ecology II, Water Relations and Carbon Assimilation*, O. L. Lange, P. S. Nobel, C. B. Osmond, and H. Ziegler (eds.), Springer-Verlag, New York, 1982.
- Brady, N.C. and Weil, R.R. *The Nature and Properties of Soil*, 11th ed., Prentice Hall, Upper Saddle River, NJ, 1996.
- Brutsaert, W. *Evaporation into the Atmosphere: Theory, History, and Applications*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1982.
- Campbell, S.G. and Norman, J.M. *An Introduction to Environmental Biology*, Springer-Verlag, New York, 1998.
- Chan, T.P. and Govindaraju, R.S. A new model for soil hydraulic properties based on a stochastic conceptualization of porous media, *Water Resour. Res.*, 39, Art. No. 1195, 2003.
- Collatz, J.G., Ball, J.T., Grivet, C., and Berry, J.A. Physiological and environmental regulation of stomatal conductance, photosynthesis and transpiration: A model that includes a laminar boundary layer, *Agric. For. Meteorol.*, 54, 107, 1991.
- Cordova, J.R. and Bras, R.L. Physically based probabilistic models of infiltration, soil moisture, and actual evapotranspiration, *Water Resour. Res.*, 17, 93, 1981.
- Crago, R.D. and Brutsaert, W. A comparison of several evaporation equations, *Water Resour. Res.*, 28, 951, 1992.
- Cui, M. and Caldwell, M.M. A large ephemeral release of nitrogen upon wetting soil and corresponding root responses in the field, *Plant Soil*, 191, 291, 1997.
- Daly, E., Porporato, A., and Rodriguez-Iturbe, I. Coupled dynamics of photosynthesis, transpiration, and soil water balance. Part I: Upscaling from hourly to daily level, *J. Hydrometeorol.*, 5, 546, 2004a.
- Daly, E., Porporato, A., and Rodriguez-Iturbe, I. Coupled dynamics of photosynthesis, transpiration, and soil water balance. II. Stochastic dynamics and ecohydrological significance, *J. Hydromet.*, 5, 559, 2004b.
- Denmead, O.T., Harper, L.A., and Sharpe, R.R. Identifying sources and sinks of scalars in a corn canopy with inverse Lagrangian dispersion analysis: I. Heat, *Agric. For. Meteorol.*, 104, 67, 2000.
- Dingman, S.L. *Physical Hydrology*, Prentice Hall, Upper Saddle River, NJ, 1994.
- D'Odorico, P., et al. Preferential states of seasonal soil moisture: The impact of climate fluctuations, *Water Resour. Res.*, 36, 2209, 2000.
- D'Odorico, P., et al. Hydrologic control on soil carbon and nitrogen cycles. II: A case study, *Adv. Water Res.*, 26, 59, 2003.
- Eagleson, P. S. Climate, soil, and vegetation. 1. Introduction to water balance dynamics, *Water Resour. Res.*, 14, 705, 1978.
- Farquhar, G.D., von Cammerer, S., and Berry, J.A. A biochemical model of photosynthetic CO₂ assimilation in leaves of C₃ species, *Planta*, 149, 78, 1980.
- Federer, C.A. A soil-plant-atmosphere model for transpiration and availability of soil water, *Water Resour. Res.*, 15, 555, 1979.
- Fernandez-Illescas, C., et al. The role of soil texture in water controlled ecosystems, *Water Resour. Res.*, 37, 2663, 2001.
- Fierer, N. and Schimel, J.P. Effects of drying-rewetting frequency on soil carbon and nitrogen transformations, *Soil Biol. Biochem.*, 34, 777, 2002.
- Finnigan, J. Turbulent transport in plant canopies, in *The Forest-Atmosphere Interactions*, Hutchinson, B.A. and Hicks, B.B. (eds.) D. Reidel, Norwell, 1985, pp. 443.
- Finnigan, J. Turbulence in plant canopies, *Annu. Rev. Fluid Mech.*, 32, 519–571, 2000.
- Gollan, T., Turner, N.C., and Schulze, E.D. The responses of stomata and leaf gas exchange to vapour pressure deficit and soil water content, III. In the sclerophyllous woody species Nerium Oleander, *Oecologia*, 65, 356, 1985.
- Gu, L.H., et al. Micrometeorology, biophysical exchange and nee decomposition in a two-story boreal forest – development and test of an integrated model, *Agric. For. Meteorol.*, 94, 123, 1999.

- Gusman, A.J. and Marino, M.A. Analytical modeling of nitrogen dynamics in soils and ground water, *J. Irrig. Drain. Eng.*, 125, 330, 1999.
- Guswa, A.J. Soil-moisture limits on plant uptake: An upscaled relationship for water-limited ecosystems, *Adv. Water Resour.*, 28, 543–552, 2005.
- Guswa, A.J., Celia, M.A., and Rodriguez-Iturbe, I. Effect of vertical resolution on predictions of transpiration in water-limited ecosystems, *Adv. Water Resour.*, 27, 467, 2004.
- Guswa, A.J., Celia, M.A., and Rodriguez-Iturbe, I. Models of soil-moisture dynamics in ecohydrology: A comparative study, *Water Resour. Res.*, 38, Art. No. 1166, 2002.
- Hillel, D. *Environmental Soil Physics*, Academic Press, San Diego, 1998.
- Hosking, J.R.M. and Clarke, R.T. Rainfall-runoff relations derived from the probability-theory of storage, *Water Resour. Res.*, 26, 1455, 1990.
- Hsiao, T.C. Plant responses to water stress, *Ann. Rev. Plant Physiol.*, 24, 519, 1973.
- Hueckel, T., Kaczmarek, M., and Caramuscio, P. Theoretical assessment of fabric and permeability changes in clays affected by organic contaminants, *Can. Geotech. J.*, 34, 588, 1997.
- Hunt, A.G. and Gee, G.W. Application of critical path analysis to fractal porous media: Comparison with examples from the Hanford site, *Adv. Water Resour.*, 25, 129, 2002.
- Hunt, H.W., et al. Simulation model for the effects of climate change on temperate grasslands ecosystems, *Ecol. Modell.*, 53, 205, 1991.
- Ingram, J. and Bartels, D. The molecular basis of dehydration tolerance in plants, *Ann. Rev. Plant Physiol. Plant Mol. Biol.*, 47, 377, 1996.
- Jones, H.G. *Plants and Microclimate: A Quantitative Approach to Environmental Plant Physiology*, Cambridge University Press, Cambridge, 1992.
- Katul, G.G. and Albertson, J.D. An investigation of higher-order closure models for a forested canopy, *Bound.-Lay. Meteorol.*, 89, 47, 1998.
- Katul, G.G. and Albertson, J.D. Modeling CO₂ sources, sinks and fluxes within a forest canopy, *J. Geophys. Res.*, 104, 6081, 1999.
- Katul, G.G. and Chang, W.-H. Principal length scales in second-order closure models for canopy turbulence, *J. Appl. Meteorol.*, 38, 1631, 1999.
- Katul, G.G. and Siqueira, M. Modeling heat, water vapor, and carbon dioxide flux distribution inside canopies using turbulent transport theories, *Vadose Zone J.*, 1, 58, 2002.
- Katul, G.G., et al. A Lagrangian dispersion model for predicting CO₂ sources, sinks, and fluxes in a uniform loblolly pine (*Pinus taeda* L.) stand, *J. Geophys. Res.*, 102, 9309, 1997.
- Katul, G.G., Ellsworth, D., and Lai, C.T. Modeling assimilation and intercellular CO₂ from measured conductance: A synthesis of approaches, *Plant Cell Environ.*, 23, 1313, 2000.
- Katul, G., Leuning, R., and Oren, R. Relationship between plant hydraulic and biochemical properties derived from a steady-state coupled water and carbon transport model, *Plant Cell Environ.*, 26, 339, 2003.
- Kaye, J.P. and Hart, S.C. Competition for nitrogen between plants and soil microorganisms, *Trends Ecol. Evol.*, 12, 139, 1997.
- Kim, D.S. and Fogler, H.S. Biomass evolution in porous media and its effects on permeability under starvation conditions, *Biotechnol. Bioeng.*, 69, 47, 2000.
- Knapp, A. K., et al. Rainfall variability, carbon cycling, and plant species diversity in a mesic grassland, *Science*, 298, 2202, 2002.
- Kramer, P.J. and Boyer, J.S. *Water Relations of Plants and Soils*, Academic Press, London, 1995.
- Kumagai, T., et al. Water cycling in a Bornean tropical rainforest under current and projected precipitation scenarios, *Water Resour. Res.*, 27, 1135, 2004.
- Lai, C.T., et al. Modeling vegetation-atmosphere CO₂ exchange by a coupled Eulerian-Lagrangian approach, *Bound. Lay. Meteorol.*, 95, 91, 2000.
- Lai, C.T., et al. Modeling the limits on the response of net carbon exchange to fertilization in a southeastern pine forest, *Plant Cell Environ.*, 25, 1095, 2002.

- Laio, F., et al. Plants in water-controlled ecosystems: Active role in hydrological processes and response to water stress. II. Probabilistic soil-moisture dynamics, *Adv. Water Resour.*, 24, 707, 2001a.
- Laio, F., et al. Plants in water controlled ecosystems: Active role in hydrological processes and response to water stress. IV. Applications to real cases, *Adv. Water Res.*, 24, 745, 2001b.
- Laio, F., et al. On the seasonal dynamics of mean soil moisture, *J. Geophys. Res.*, 107, 10.1029/2001JD001252, 2002.
- Lange, O.L., Kappen, L., and Schulze, E.D. *Water and Plant Life: Problems and Modern Approaches*, Springer-Verlag, Berlin, 1976.
- Larcher, W. *Physiological Plant Ecology*, 3rd ed., Springer-Verlag, Berlin, 1995.
- Leuning, R. A critical appraisal of a combined stomatal-photosynthesis model for C₃ plants, *Plant Cell Environ.*, 18, 339, 1995.
- Leuning, R. Estimation of scalar source/sink distributions in plant canopies using Lagrangian dispersion analysis: Corrections for atmospheric stability and comparison with a multilayer canopy model, *Bound.-Lay. Meteorol.*, 96, 293, 2000.
- Levitt, J. *Responses of Plants to Environmental Stresses*, vol. I, 2nd ed., Academic Press, New York, 1980.
- Manzoni, S., et al. Soil nutrient cycles as a nonlinear dynamical system, *Nonlinear Proc. Geophys.*, 11, 589, 2004.
- Mary, B., et al. Interactions between decomposition of plant residues and nitrogen cycling in soil, *Plant Soil*, 181, 71, 1996.
- Massman, W.J. and Weil, J.C. An analytical one dimensional second-order closure model of turbulence statistics and the Lagrangian time scale within and above plant canopies of arbitrary structure, *Bound.-Lay. Meteorol.*, 91, 81, 1999.
- Melillo, J.M., Borchers, J., and Chaney, J. Vegetation ecosystem modeling and analysis project—comparing biogeography and biogeochemistry models in a continental-scale study of terrestrial ecosystem responses to climate change and CO₂ doubling, *Global Biogeochem. Cycles*, 9, 407, 1995.
- Meyers, T. and Paw U, K.T. Modeling the plant canopy micrometeorology with higher- order closure principles, *Agric. For. Meteorol.*, 41, 143, 1987.
- Milly, P.C.D. An analytical solution of the stochastic storage problem applicable to soil water, *Water Resour. Res.*, 29, 3755, 1993.
- Milly, P.C.D. A minimalistic probabilistic description of root zone soil water, *Water Resour. Res.*, 37, 457, 2001.
- Molina, J.A. and Smith, P. Modeling carbon and nitrogen processes in soils, *Adv. Agron.*, 62, 253, 1998.
- Monteith, J.L. and Unsworth, M. *Principles of Environmental Physics*, 2nd ed., Butterworth-Heinemann Publishers, 1990.
- Nathan, R. and Katul, G.G. Foliage shedding in deciduous forests lifts up long-distance seed dispersal by wind, *Proc. Natl. Acad. Sci.*, 102, 8251–8256, 2005.
- Neff, J.C., Chapin III, F. S., and Vitousek, P. Breaks in the cycle: Dissolved organic nitrogen in terrestrial ecosystems, *Front Ecol. Environ.*, 1, 205, 2003.
- Nicholson, S. Land surface processes and Sahel climate, *Rev. Geophys.*, 38, 117, 2000.
- Nilsen, E.T. and Orcutt, D.M. *Physiology of Plants Under Stress: Abiotic Factors*, Wiley, New York, 1998.
- Nobel, P.S. *Physicochemical and Environmental Plant Physiology*, Academic Press, San Diego, CA, 1999.
- Noy-Meir, I. Desert ecosystems: Environment and producers, *Annu. Rev. Ecol. Syst.*, 4, 25, 1973.
- Oren, R., et al. Water balance delineates the soil layer in which moisture affects canopy conductance, *Ecol. Appl.*, 8, 990, 1998.
- Parton, W.J., Steward, J., and Cole, C. Dynamics of C, N and S in grassland soils: A model, *Biogeochemistry*, 5, 109, 1988.
- Peixoto, J.P. and Oort, A.H. *Physics of Climate*, Springer-Verlag, New York, 1992.
- Porporato, A., et al. Plants in water-controlled ecosystems: active role in hydrologic processes and response to water stress. III. Vegetation water stress, *Adv. Water Resour.*, 24, 725, 2001.
- Porporato, A., et al. Soil moisture and plant stress dynamics along the Kalahari precipitation gradient, *J. Geophys. Res.–Atmos.*, 108, 4127, 2003a.

- Porporato, A., et al. Hydrologic controls on soil carbon and nitrogen cycle. I: Modelling scheme, *Adv. Water Res.*, 26, 45, 2003b.
- Porporato, A., Daly, E., and Rodriguez-Iturbe, I. Soil water balance and ecosystem response to climate change, *Am. Nat.*, 164, 625, 2004.
- Puma, M.J., et al. Functional relationship to describe temporal statistics of soil moisture averaged over different depths, *Adv. Water Resour.*, 28, 553, 2005.
- Raupach, M.R. Applying Lagrangian fluid mechanics to infer scalar source distributions from concentration profiles in plant canopies, *Agric. For. Meteorol.*, 47, 85, 1989a.
- Raupach, M.R. A practical Lagrangian method for relating scalar concentrations to source distributions in vegetation canopies, *Q. J. Roy. Meteorol. Soc.*, 115, 609, 1989b.
- Raupach, M.R. and Shaw, R.H. Averaging procedures for flow within vegetation canopies, *Bound.-Lay. Meteorol.*, 22, 79, 1982.
- Rawls, W.J., et al. Effect of soil organic carbon on soil water retention, *Geoderma*, 116, 61, 2003.
- Rieu, M. and Sposito, G. Fractal fragmentation, soil porosity, and soil-water properties. 1. Theory, *Soil Sci. Soc. Am. J.*, 55, 1231, 1991.
- Rockhold, M.L., et al. Considerations for modeling bacterial-induced changes in hydraulic properties of variably saturated porous media, *Adv. Water Resour.*, 25, 477, 2002.
- Rodriguez-Iturbe, I. and Porporato, A. *Ecohydrology of Water-Controlled Ecosystems*, Cambridge University Press, Cambridge, 2004.
- Rodriguez-Iturbe, I., et al. Probabilistic modeling of water balance at a point: The role of climate, soil and vegetation, *Proc. R. Soc. Lond. A*, 455, 3789, 1999.
- Salisbury, F.B. and Ross, C.W. *Plant Physiology*, 4th ed., Wadsworth Publishing Company, Belmont, 1992.
- Salvucci, G.D. Estimating the moisture dependence of root zone water loss using conditionally averaged precipitation, *Water Resour. Res.*, 37, 1357, 2001.
- Schimel, J.P. and Bennett, J. Nitrogen mineralization: challenges of a changing paradigm, *Ecology*, 85, 591, 2004.
- Schimel, J.P. and Weintraub, M.N. The implications of exoenzyme activity on microbial carbon and nitrogen limitation in soil: A theoretical model, *Soil Biol. Biochem.*, 35, 549, 2003.
- Schimel, D.S., et al. Climate and nitrogen controls on the geography and timescales of terrestrial biogeochemical cycling, *Global Biogeochem. Cycles*, 10, 677, 1996.
- Schimel, D.S., Braswell, B.H., and Parton, W.J. Equilibration of terrestrial water, nitrogen and carbon cycles, *Proc. Natl. Acad. Sci. USA*, 94, 8280, 1997.
- Scholes, M.C., et al. NO and N₂O emissions from savanna soils following the first simulated rains of the season, *Nutr. Cycl. Agroecosys.*, 48, 115, 1997.
- Schuepp, P.H. Tansley Review n. 59, Leaf boundary-layers, *New Phytologist*, 125, 477, 1993.
- Siqueira, M. and Katul, G. Estimating heat sources and fluxes in thermally stratified canopy flows using higher-order closure models, *Bound. Lay. Meteorology*, 103, 125, 2002.
- Siqueira, M., Lai, C.T., and Katul, G. Estimating scalar sources, sinks and fluxes in a forest canopy using Lagrangian, Eulerian and hybrid inverse models, *J. Geophys. Res.*, 105, 19475, 2000.
- Six, J., et al. A history of research on the link between (micro)aggregates, soil biota, and soil organic matter dynamics, *Soil Till. Res.*, 79, 7, 2004.
- Smith, J.A.C. and Griffith, H. *Water Deficits: Plant Responses from Cell to Community*, Bios Scientific Publishers, Oxford, 1993.
- Sollins, P., Homann, P., and Caldwell, B.A. Stabilization and destabilization of soil organic matter: mechanisms and controls, *Geoderma*, 74, 65, 1996.
- Sperry, J.S. Hydraulic constraints on plant gas exchange, *Agric. For. Meteorol.*, 104, 13, 2000.
- Stewart, T.L. and Kim, D.S. Modeling of biomass-plug development and propagation in porous media, *Biochem. Eng. J.*, 17(2), 107–119, 2004.
- Stull, R.B., *An Introduction to Boundary Layer Meteorology*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1988.

- Taylor, S.W., Milly, P.C.D., and Jaffe, P.R. Biofilm growth and the related changes in the physical properties of a porous medium. 2. Permeability, *Water Resour. Res.*, 26, 2161, 1990.
- Thullner, M., Zeyer, J., and Kinzelbach, W. Influence of microbial growth on hydraulic properties of pore networks, *Transport Porous Med.*, 49, 99, 2002.
- Tyler, S.W. and Wheatcraft, S.W. Fractal processes in soil-water retention, *Water Resour. Res.*, 26, 1047, 1990.
- Tyree, M.T. and Sperry, J.S. Vulnerability of xylem to cavitation and embolism, *Annu. Rev. Plant Physiol. Plant Mol. Biol.*, 40, 19, 1989.
- Vandevivere, P. Bacterial clogging of porous media: A new modeling approach, *Biofouling*, 8, 281, 1992.
- Warland, J.S. and Thurtell, G.W. A Lagrangian solution to the relationship between a distributed source and concentration profile, *Bound.-Lay. Meteorol.*, 96, 453, 2000.
- Wosten, J.H.M., Pachepsky, Ya.A., and Rawls, W.J. Pedotransfer functions: Bridging the gap between available basic soil data and missing soil hydraulic characteristics, *J. Hydrol.*, 251, 123, 2001.

Further Information

A general introduction on plant-atmosphere exchange and its role in climate science can be found in Peixoto and Oort (1992), Dingman (1994), and Bonan (2002). Campbell and Norman (1998) focus on modeling radiative and eco-physiological processes related to various components of the soil-plant-atmosphere system.

A detailed description of boundary-layer dynamics along with Reynolds-Averaged turbulence closure models can be found in Stull (1988) or Brutsaert (1982). Finnigan (2000) provides a review of the structure of turbulence inside canopies.

A number of books treat hydrological and physiological features of vegetation and their relationship to microclimate . We suggest three of them:

- Jones (1992) provides a description of the interactions between plants and their aerial environment accompanied by explanations of the state-of-the-art models now adopted.
- Larcher (1995) is an accurate synthesis of the vast number of processes linking plants and their environment.
- Nobel (1999) provides a mathematical treatment of the physical and chemical processes occurring inside plant cells and tissues.

Regarding the physical properties of soils, two good references are Hillel (1998) and Brady and Weil (1996); both give an exceptionally well-documented description of the soil properties and its role in the terrestrial energy exchange and water cycle.

Finally, the material related to stochastic soil-moisture models and soil nutrient dynamics is extensively presented in the book by Rodriguez-Iturbe and Porporato (2004).

30

The Role of Geographical Information Systems in Groundwater Engineering

30.1	Introduction to Geographic Information Systems	30-1
	Overview of GIS	
30.2	GIS Components.....	30-2
	Software • Hardware • Data	
30.3	Data Representation	30-3
30.4	Analysis Capabilities of GIS	30-3
30.5	Web GIS Services.....	30-5
30.6	Application of GIS to Groundwater Engineering.....	30-6
	Mapping the Occurrence of Groundwater and Groundwater Features • Managing Spatial Data Aspects of Groundwater Projects • Spatial Data Analyses Using Spatial Statistics • Surface Fitting and Interpolation • Visualization • Modeling Water Vulnerability Using Spatial Data • Modeling Groundwater Movement • Implementation of Policies Involving Spatial Dimensions	
	Glossary.....	30-14
	References	30-15
	Further Information	30-17

Bernard Engel and
Kyoung Jae Lim
Purdue University

Kumar C.S. Navulur
Digital Globe

30.1 Introduction to Geographic Information Systems

Geographic Information Systems (GIS) have become important tools in efficiently solving many problems in which spatial data is important. Natural resources and environmental concerns, including groundwater, have benefited greatly from the use of GIS. This chapter provides a brief introduction to GIS and some of its applications in addressing groundwater issues.

30.1.1 Overview of GIS

GIS have evolved rapidly in the last decade, becoming powerful computer tools for varied applications ranging from sophisticated analysis and modeling of spatial data to simple inventory and management. The rapid growth, diversity, and commercial orientation of GIS have resulted in multiple definitions of the term GIS. GIS can be described as a collection of computer software and hardware for storing, manipulating, analyzing, and displaying spatial or geographically referenced data. GIS can also be described as a combination of computer tools: cartographic and tools for displaying data and making maps, a database for storing spatial data, and an analysis tool for manipulating data. As a visualization tool, GIS allows graphical display of maps, tabular information, statistical summaries, and modeling solutions. As a database, GIS can store, maintain, and update spatial data and associated descriptive information. GIS supports most database functions such as browsing, tabular queries, updating data, and functions supported by commonly used computer software such as spreadsheets and statistical packages. The most significant difference between GIS and other information systems and databases is the spatial nature of the data in a GIS. The analysis functions in a GIS allow manipulation of multiple themes of spatial data to perform operations such as overlays, buffering, and arithmetic operations on the data.

The application of GIS has grown dramatically since the 1980s with yearly expenditures on GIS experiencing double digit growth. Estimated expenditures within the US on GIS were approximately \$6 billion in 1995 and are estimated at \$75 to \$95 billion in 2005 [1,2]. The ability of GIS to perform analysis on spatial data and corresponding attribute data and to integrate different types of data at high speeds are unmatched by manual methods. There is an increasing trend within organizations and public agencies at all levels to develop GIS for the management and analysis of spatial data. The ability of a GIS to perform complex spatial analyses rapidly provides a quantitative, as well as qualitative, advantage in planning scenarios, decision models, change detection and analyses, and other problems requiring refinements to successive analyses. Digital databases like SSURGO (Soil Survey Geographic), NASIS (National Soil Information System), land use data, satellite imagery, Digital Elevation Model (DEM) data, and NHD (National Hydrography Dataset) river reach data are now being distributed in GIS formats, and efforts are ongoing to integrate more data into GIS formats.

30.2 GIS Components

GIS, like other information systems, are integral parts of applications to address various problems and issues. To create successful GIS applications, there must be an organization of people, facilities, equipment, and data to successfully implement and maintain the GIS. Four essential elements that constitute a GIS are computer software, hardware, data, and personnel.

30.2.1 Software

GIS software are extremely diverse in functionality, database structure, and hardware requirements. A GIS user today is presented with a wide variety of commercial GIS software. Popular GIS software includes ArcGIS, ArcView, Arc Internet Map Server (IMS), MapServer, MapInfo, MGE (Modular GIS Environment), Genasys, ERDAS, GRASS, IDRISI, GeoMedia, and Atlas. These software tools are sophisticated and have a variety of functionalities for GIS applications and analyses. Most of these software tools have capabilities to handle data in various formats, thereby facilitating data transfer between GIS software. Some of the above software are hardware specific and others run on a variety of platforms. The marketing of GIS software today is increasingly oriented to an open system or platform-independent approach.

30.2.2 Hardware

The large memory and disk requirements of GIS software and data resulted in a trend toward workstations running the UNIX operating system in the late 1980s and early 1990s. However, rapidly declining computer hardware costs and increasing capabilities of personal computers have made GIS software available on

a wide variety of computer platforms including personal computers, high performance workstations, Windows CE devices, and PDAs. GIS applications often require the use of special peripherals such as digitizers and scanners for developing digital data. Global Positioning System (GPS) units with hand-held and mobile devices running the ArcPad or similar programs have become widely used to locate and collect points and define lines and boundaries for developing, analyzing, and displaying spatial data in GIS formats. Other hardware requirements for GIS applications include plotters and printers for output of the final GIS products, including hard copy maps and summaries.

30.2.3 Data

Even though hardware and software may be available, GIS users cannot perform analysis, visualization, or modeling without data. Non-spatial data with spatial indexing as well as spatial data are used in GIS projects. In the early stage of the GIS era, a GIS user had to develop spatial digital data by digitizing or scanning paper maps, which was an extremely time-consuming job and often introduced error in developing the digital data. However, many national, state and local level GIS projects have now been completed to provide digital datasets. Many of these GIS data have been extracted from airborne or remotely sensed image data with digital image processing techniques. The GIS data model has evolved with the introduction of relational and spatial database management systems. Maintenance of metadata (data describing other data) is a critical aspect of successful GIS projects.

30.3 Data Representation

Data are the most important resource in a GIS. Digital data are usually very expensive to develop and require large quantities of storage space. GIS data are commonly stored in one of two forms: (1) vector or (2) raster. Vector data are the set of points, lines, and polygons that are used to represent map feature locations (see Figure 30.1). The spatial relationships between various features must be explicitly coded by the GIS software for vector data. A topological data structure may be used to store the relationships between objects. Well locations, streams, and state boundaries are some of the examples of point, line, and polygon vector features, respectively. Vector data can be represented in various formats such as shape files, DXF, DLG, MOSS, and SDTS. GIS software is generally designed to accept and export data in one or more of these formats.

Raster data can be described as a matrix of columns and rows or as a regular grid of square or rectangular data. Location in raster data is defined by row and column numbers and size of each cell or pixel. The value assigned to a grid or cell indicates the value of the attribute it represents. Figure 30.1 shows a comparison of vector and raster representation of GIS data sets. Airborne and satellite data are typical examples of raster data. Various graphic interchange formats for raster data are GIF, TIFF, BIL, BIP, BSQ, JPEG, DEM, ASCII, and SDTS. The choice between vector and raster GIS is usually driven by application requirements.

The most important aspects of data quality are accuracy, time, scale, and completeness of data sets. The error associated with each of the data sets will affect the accuracy of the final product and therefore defines the usefulness of the final products. Hence, GIS data developers are encouraged to document the data or create metadata to facilitate the sharing of spatial information with other GIS users. Metadata for a GIS data layer includes the scale of the data, projection, how the data were developed, and many other items of information useful in understanding the GIS data layer. Metadata is typically stored in either ASCII text file format or Extensible Markup Language (XML) format. Metadata in an XML format has become widely accepted and used because of its ease for data sharing. The XML format has also become the standard for data sharing through the Internet.

30.4 Analysis Capabilities of GIS

Analysis of spatial and attribute data in a GIS can be classified into five main types of procedures: (1) data transformation and restructuring, (2) data retrieval, classification, and measurement, (3) overlay,

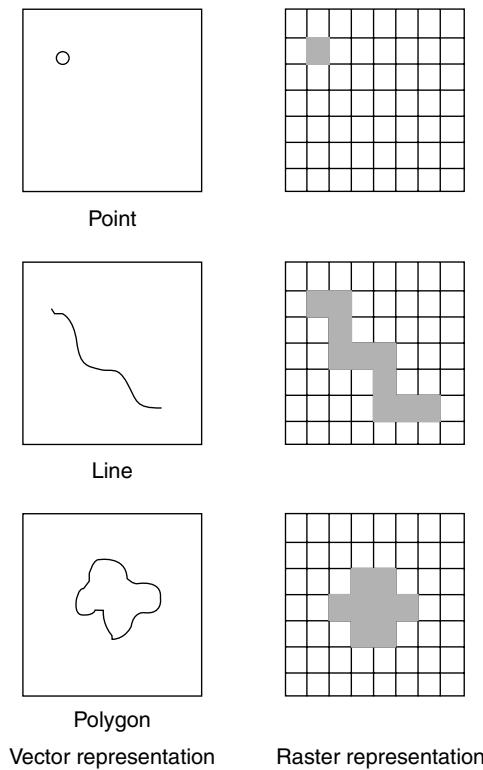


FIGURE 30.1 Vector and corresponding raster representation.

(4) neighborhood and statistical measures, and (5) connectivity. Each of these is discussed briefly in the remainder of this section.

Data transformation functions allow mathematical manipulation of raster or vector data to rotate, translate, scale, or mosaic different map layers. Such efforts are often necessary when acquiring GIS data sets. Transformation of data may be required to convert the data being developed by digitizing or scanning into useful forms. Restructuring functions include data conversion from raster to vector, vector to raster, and converting data between different formats.

Retrieval functions allow the query and retrieval of spatial and tabular data stored in the GIS database. They also facilitate spatial queries specific to a region on single or multiple map layers. Classification allows the user to rename attributes, such as to indicate that areas within 0.5 miles of a well are sensitive areas and special precautions should be taken with potential contaminants in this area. Generalization allows the combining of classes such as considering all water table depths less than 15 ft as “shallow” for a particular application. Measurement functions allow the user to determine distance from a feature, or length, area, volume, and perimeter of a feature or features.

Overlay functions constitute the third set of GIS analysis tools. Overlay allows the user to combine and overlay multiple thematic data (map layers) at different scales and in different formats. Overlay functions can also perform arithmetic operations such as addition, division, or multiplication on map layers and logical operations between map layers such as AND, OR, and NOT. A typical logical operation might be finding areas with soils having high permeabilities and areas with shallow depths to water tables and defining areas having this combination as being susceptible to groundwater contamination.

The fourth set of analysis procedures includes neighborhood functions that involve calculations based on the value of neighboring cells. For example, a search technique can be used to find the maximum contamination level in an area. The Thiessen polygon technique can be used to construct polygons

around a set of points such as rain gages or groundwater monitoring wells. Interpolation procedures, such as inverse distance weighting, kriging, and splines, can be used for interpolating a value based on the values of neighboring locations. Such procedures are useful to create continuous surfaces from point data, such as data observed in groundwater monitoring wells.

The fifth group of analysis procedures contains connectivity functions such as contiguity measures that can be used to determine the size of a catchment. Proximity can be used to determine distance from a feature and generate buffer zones. Network functions can be used to predict water flow in a stream network, and other connectivity functions can be used to define drainage basins, determine sinks in watersheds, and generate perspective views. Statistical measures allow calculation of spatial auto-correlation, correlation between maps, and confidence intervals.

30.5 Web GIS Services

With the introduction of the Internet in the early 1990s, data and information sharing has become easier and faster than ever before. With advances in GIS and information technologies, including the World Wide Web (WWW), spatial data sharing over the Internet has grown rapidly. Many Web GIS software tools have been developed to provide spatial information through the Web. Commercially available and public domain Web GIS software includes ArcIMS, AspMap, GeoTools, JShape, MapServer, MapGuide, and GeoServ.

Beyond just providing spatial information using Web GIS software, many Web GIS-based decision support systems have been developed to overcome data availability, expertise, ease-of-use, and access difficulties in solving real world problems, especially in natural resources and environmental areas. With the rapidly growing numbers of Internet users in the United States and beyond, the Web presents a means to supply large numbers of tools, databases, and other information to decision makers and other potential users. In addition, the Web-based approach provides opportunities to increase involvement of stakeholders in the decision making process by providing them with the required knowledge and data through these accessible, easy to use Web GIS-based systems.

The following are examples of Web GIS-based decision support systems to help solve natural resource and environmental concerns [3–8].

- NAPRA WWW (<http://danpatch.ecn.purdue.edu/~napra>): estimates impacts of agricultural management systems on surface and subsurface hydrology and water quality, and to identify location specific environmentally friendly agricultural management practices
- L-THIA WWW (<http://www.ecn.purdue.edu/runoff/lthianew>): estimates impacts of land use changes on hydrology and water quality
- ROMIN WWW (<http://pasture.ecn.purdue.edu/~romin>): provides assistance with environment-friendly land use planning
- SEDSPEC (<http://pasture.ecn.purdue.edu/~sedspec>): assists in analyzing runoff and erosion problem by determining the peak rate of runoff from the area and providing information about different types of runoff and erosion control structures
- WATERGEN (<http://pasture.ecn.purdue.edu/~watergen>): provides online watershed delineation tool and serves as an interface for other spatial decision support systems
- WHAT (<http://pasture.ecn.purdue.edu/~what>): provides automated baseflow separation and a hydrograph analysis tool to complement the USGS daily stream flow web site with a Web GIS interface

The Web GIS-based decision support systems overviewed above provide numerous advantages over traditional systems. In general, making such models, decision support tools, and databases accessible on the Web gives users access to vast and diverse options. Users of Web GIS-based systems do not need to learn how to use expensive and complex GIS software, and large databases including GIS data required to

run the model are also stored on a central server. This has a twofold advantage of significant financial and time savings for users of the Web GIS-based system.

30.6 Application of GIS to Groundwater Engineering

Typical examples of GIS applications in groundwater studies include site suitability analyses; mapping information such as water table depths, aquifer type and material, aquifer recharge, and groundwater remediation areas; managing site inventory data; computing statistics to estimate spatial correlation of a process occurring over a region such as non-point source pollution; estimating vulnerability of groundwater to pollution potential from non-point sources of pollution; integrating groundwater quality assessment models with spatial data to create spatial decision support systems; modeling groundwater and contaminant movement: advection and dispersion modeling, developing flow field or vector fields representing groundwater seepage velocities, tracking particle movement, and particle retardation; modeling solute transport and leaching; and evaluating soil salinity and salt loading into groundwater.

30.6.1 Mapping the Occurrence of Groundwater and Groundwater Features

Maps play an important role in portraying information about groundwater systems. GIS can be used to store these maps, allowing the user to create the desired map products. GIS further provide the ability to manipulate the mapped data.

Mapping the material of underlying aquifer systems requires information of geologic strata and other data including soil properties and depth to bedrock. Soil databases such as SSURGO and NASIS contain information that can be used in conjunction with glacial geology to interpret the material of the aquifers. Figure 30.2 shows the aquifer media data layer developed from the glacial geology map of Indiana at a 1:100,000 scale.

Another application is estimating the recharge flux to the water table. GIS can be utilized to evaluate and delineate patterns of recharge within a region by incorporating available information about observed temporal fluctuations of recharge at specified locations. Continuous remote acquisition of water table depths combined with GIS capabilities for distributing the flux spatially across the study region can result in near-real-time monitoring of spatial variability in recharge flux [9]. Recharge can also be estimated using information including current climate, soil, and vegetation/land use patterns and estimates of recharge for various soil-vegetation combinations [10].

30.6.2 Managing Spatial Data Aspects of Groundwater Projects

Siting a waste disposal facility is an example for which GIS can help manage the spatial data considerations including the potential impacts on groundwater. State and federal regulations for a site location can be implemented using GIS functions such as overlays, buffers, and arithmetic computations on data layers. For example, regulations for a hazardous waste landfill may indicate that they should not be located: (1) within 100 ft of any lake reservoir or continuously flowing stream, (2) within 100 ft from any of the site's property boundaries, (3) within 600 ft of potable domestic water supplies, (4) within 600 ft of any dwelling, and (5) within 1200 ft of a public water supply. Such exclusionary areas are easily computed as buffers within GIS. Further, GIS maps of soils, geology, and climate, hydrological and cultural data sets can be integrated in a GIS to develop maps showing areas suitable for hazardous waste landfill siting that meet the above and other regulations. Figure 30.3 shows the areas suitable for hazardous waste landfill siting in Tippecanoe County, Indiana that were identified using such an approach [11].

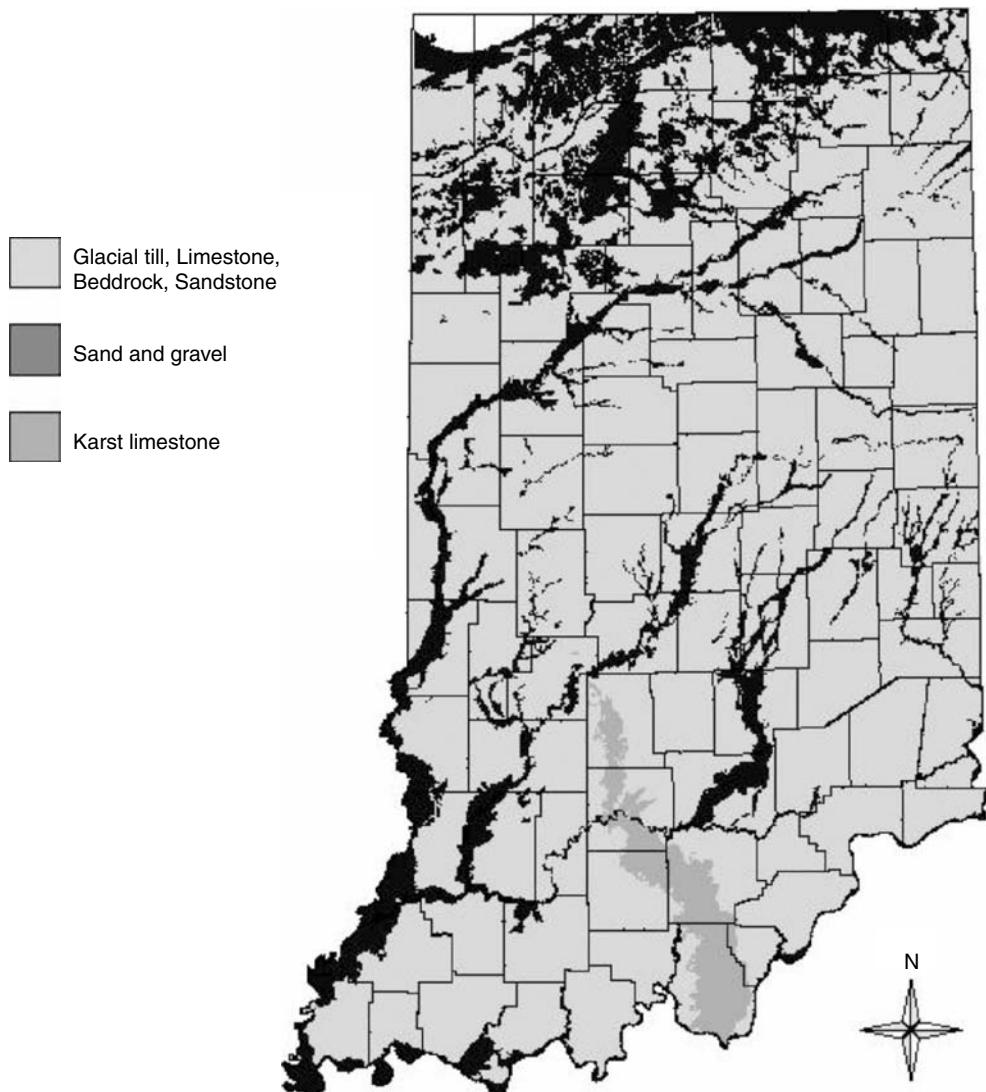


FIGURE 30.2 Aquifer material map of Indiana.

30.6.3 Spatial Data Analyses Using Spatial Statistics

Spatial statistics are a rapidly emerging field being applied to many diverse applications [12]. Spatial auto-correlation studies are often used in geo-statistics for determining the spatial dependencies of the factors being studied. In groundwater studies, spatial statistics can be computed to study the distributions of non-point source (NPS) contamination of groundwater on a regional scale, and spatial variation of hydrogeologic factors affecting solute transport including hydraulic conductivity of the aquifer, slope, and groundwater properties.

Spatial auto-correlation refers to the spatial ordering of a single variable and to the relationship between pairs of observations of this variable. The ordering of n observed values of some variable X is usually described with the aid of a connectivity matrix, C . Non-zero c_{ij} entries in the $n \times n$ matrix indicate that the corresponding polygons are juxtaposed. For data measured on an interval/ratio scale, Moran coefficient statistics can be used [13] for assessing spatial correlation. These spatial statistics can be used to determine

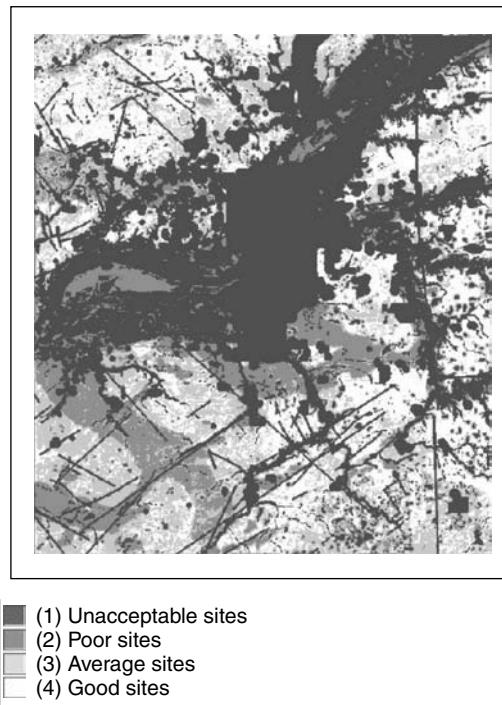


FIGURE 30.3 Sites suitable for hazardous waste disposal in Tippecanoe County, Indiana.

the spatial correlation of NPS pollution and to eliminate detections from point sources of pollution in a water quality database.

30.6.4 Surface Fitting and Interpolation

Surface fitting and interpolation techniques are commonly employed for creating continuous data (raster data) from a distributed set of data points over a geographical region (e.g., creating a depth to groundwater map from observations of water depth in wells). The underlying assumption in these techniques is that the data points are random, well-sampled, and are representative of the spatial distributions of the attribute. Commonly employed techniques for interpolation include inverse distance weighting interpolation (IDW), kriging using a semi-variogram, polynomial regression, and spline fitting [14,15].

Inverse distance weighting interpolation computes a cell value using a linear weighting combination of the set of sample points surrounding the cell. The weight is a function of inverse distance to the observed values. Usually a power of two is adopted for inverse distance and values greater than two will increase the influence of nearby data. Also the interpolated value can be controlled by limiting the search radius or by limiting the number of sample points to be used for interpolation.

Kriging is based on the regional variable theory that assumes the spatial variation of an attribute is statistically homogeneous throughout the surface. The variation of the attribute is measured using a semi-variogram, a graphical plot of semi-variance against the distance between pairs of sample points. Using the surface estimators, ordinary kriging, and universal kriging, the value at a cell can be interpolated. Polynomial regression interpolation depends on the order of polynomial used for fitting the surface. A first-order polynomial regression simply performs a least-squares fit of a plane to the set of data points. As the order of the polynomial increases, the interpolation becomes more complex.

Polynomial regression generally creates smooth surfaces, and the surface created seldom passes through the original data points. Spline interpolation is a two-dimensional (2D) minimum curve interpolation that creates a smooth surface, and unlike polynomial regression, the surface passes through the input data

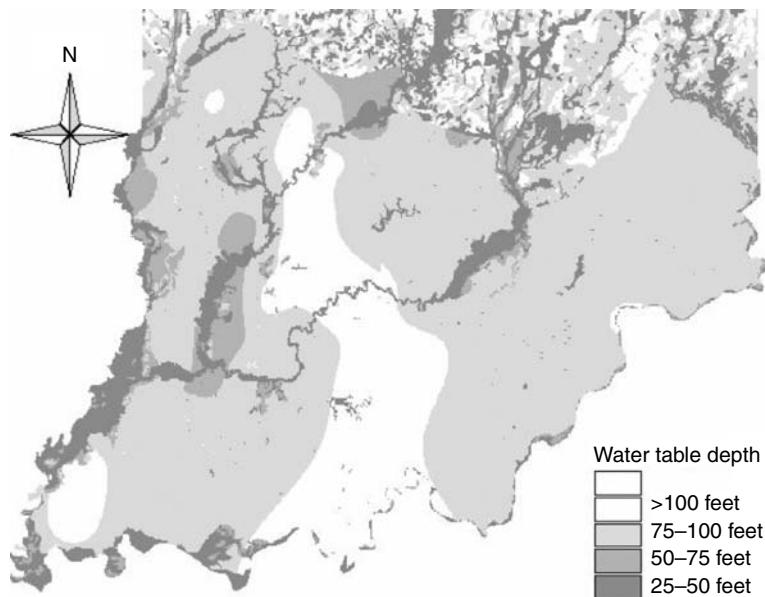


FIGURE 30.4 Interpolated water table depths based on water depths in wells.

points. The commonly used minimum-curvature technique is also called thin plate interpolation. Spline interpolation ensures a smooth surface together with continuous first-derivative surfaces.

Functions for conducting surface fitting and interpolation are often included in commercial GIS software [14,15]. These techniques are commonly used in geo-statistics and can be applied in groundwater studies. The distribution of sample points will determine which method is appropriate. An example application of these techniques within GIS includes developing groundwater table depths for a region in unconfined aquifers using observed water table depths in drinking water wells. Figure 30.4 shows the map of water table depths in southern Indiana developed using the inverse distance weighting method. Another example is the study in northeastern Colorado where inverse distance weighting interpolation was used for mapping spatial distributions of nitrate contamination in groundwater for the region [16].

30.6.5 Visualization

The saying “A picture is worth a thousand words” is certainly relevant to GIS. One of the advantages of using a GIS is its visualization capability. Visualization is a way of interpreting an image in digital format and improving human perception. When visualizing 3D phenomena in a 2D environment, a significant amount of information is lost. With advances in GIS technology, visualization of the phenomena in a 3D environment has become easier and more accurate, which provides opportunities for diverse analysis and exploitation of the data. Visualization of contaminant transport, plume movement, and delineation of remediation areas are typical 3D visualization examples in the groundwater engineering discipline.

Many efforts are underway and have been completed to develop 3D visualization for groundwater modeling systems, such as MODFLOW, MODPATH, MT3DMS, MOC3D, and SUTRA. For example, Model Viewer, developed by the US Geological Survey, is a public domain program for 3D visualization of groundwater model results. The Visual MODFLOW and GMS software are commercially available visualization programs.

30.6.6 Modeling Water Vulnerability Using Spatial Data

Non-point sources are often responsible for nitrates and pesticides in groundwater and surface water. However, it is difficult to predict NPS pollution in groundwater and surface water due to agricultural

activities because numerous factors including soil, climate, crop, rate, and timing of pesticide and nutrient applications and tillage affect the movement of NPS pollutants. GIS are increasingly being integrated with groundwater and surface water quality assessment models [17–21] to facilitate model use. The data required for these models can be stored and managed in a GIS to facilitate the model simulations.

GIS maps showing the areas vulnerable to groundwater contamination for a region would have many potential uses, such as implementation of groundwater management strategies to prevent degradation of groundwater quality and locating groundwater monitoring systems. These maps would also be helpful in evaluating existing and potential policies for groundwater protection. Groundwater vulnerability mapping can be classified into two categories: sensitivity mapping and vulnerability mapping. Sensitivity of groundwater to pollution is determined based on hydrogeologic setting, such as depth to groundwater, presence of confining layer, groundwater recharge, and hydraulic conductivity [20,21]. Groundwater vulnerability estimates how groundwater sensitivity is modified with human activities such as land use practices, pesticide and nutrient uses, and water use [20,21].

To develop regional maps of groundwater vulnerability to agricultural NPS pollution in Indiana, simple techniques like DRASTIC and SEEPAGE have been applied. DRASTIC is a regional scale groundwater quality assessment model that estimates groundwater vulnerability of aquifer systems based on the hydrogeologic factors of a region [22]. It is an empirical model and was developed by the National Well Water Association in conjunction with USEPA in the 1980s for evaluating pollution potential of groundwater systems on a regional scale. The factors considered in DRASTIC are: Depth to water table, Recharge (aquifer), Aquifer media, Soil media, Topography, Impact of vadose zone media, and Conductivity (hydraulic) of the aquifer.

A region to be rated by DRASTIC is subdivided into smaller areas in which the factors considered are nearly homogeneous. A grid cell subdivision is often used and thus a raster GIS approach is well suited for conducting DRASTIC analyses. Each of the subdivisions of the factors considered in the model is weighted based on the factor values. The DRASTIC Index [21,22], a measure of pollution potential, is computed by summing the weighted factors of each subdivision. The higher the DRASTIC index, the greater the relative pollution potential. The DRASTIC Index can be converted into qualitative risk categories of low, moderate, high, and very high.

Existing spatial databases like SSURGO contain inputs required for these groundwater models. GIS can be used to develop and store data representing hydrogeologic settings of a region required for DRASTIC and SEEPAGE. GIS classification, overlay, and arithmetic operations can be used for integrating the data layers to develop maps showing areas vulnerable to groundwater contamination from non-point source pollution. Figure 30.5 shows the vulnerability of groundwater systems in Indiana as predicted by DRASTIC analyses. Using a GIS, a comparison of the predictions of DRASTIC analyses with a GIS-based water quality database of nitrate detections in drinking water wells from USGS showed that 80% of nitrate detections >2 ppm were within the DRASTIC highly and very highly vulnerable areas [23]. The DRASTIC map is useful as a preliminary screening tool for policy and decision making in groundwater management strategies on a regional scale.

The DRASTIC approach considers only soil structure but not other soil properties such as texture. However, studies have shown that soil structure is important in estimating water and pollutant transport through preferential flow [24–26]. Thus, soil structure information with land use and agricultural chemical use data were used to improve the DRASTIC methodology for groundwater vulnerability mapping using fuzzy rules [23]. The fuzzy rule based DRASTIC results considering soil structure were better correlated with measured water quality data [23].

However, the DRASTIC and SEEPAGE approaches do not consider the effects of temporal changes in variable agronomic management, crop growth, land use, climate, and properties and amounts of nutrients and pesticides applied to local areas. The Web GIS-based National Agricultural Pesticide Risk Analysis (NAPRA) system [3], available at <http://danpatch.ecn.purdue.edu/~napra>, was developed to estimate the site-specific effects of land use and management on groundwater vulnerability. The NAPRA system builds the input parameter files for the Groundwater Loading Effects of Agricultural Management Systems

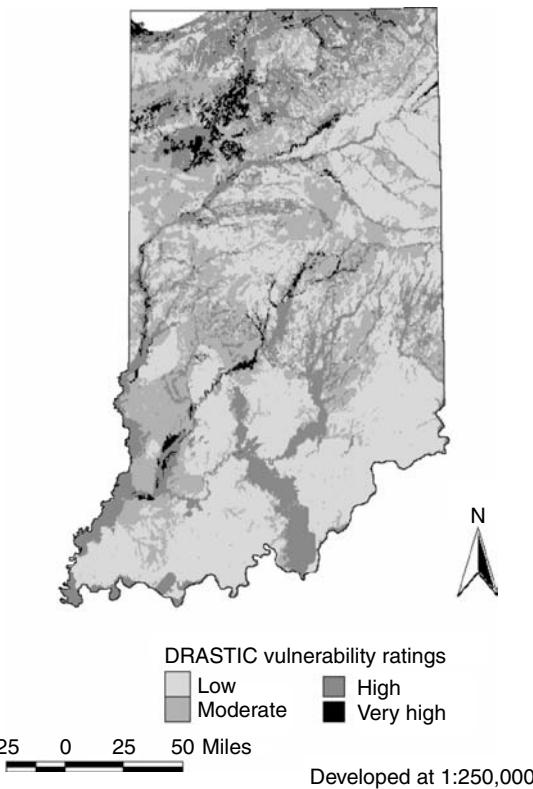


FIGURE 30.5 DRASTIC groundwater vulnerability map of Indiana.

(GLEAMS) [27] model from user provided crop management, pesticide, and nutrient data in the input interface by querying databases and by running weather generator models (Figure 30.6). The NAPRA postprocessor generates pesticide and nutrient loss probability of exceedance and historic loss curves and GIS maps.

The NAPRA system was used to simulate state average nutrient and atrazine application rates for continuous corn in Indiana to develop a regional scale shallow groundwater vulnerability map. The NAPRA system simulates the movement of nitrate and atrazine below the root zone, and it can be assumed that much of the atrazine and nitrate leached below the root zone is lost to shallow groundwater. Figure 30.7 shows the NAPRA predicted nitrate losses to shallow groundwater. The natural break classification method was used to classify the NAPRA predicted nitrate values in Figure 30.7. The NAPRA predicted nitrate loss map could be used to identify areas that are vulnerable to NPS nitrate reaching groundwater.

Statistical techniques used in conjunction with GIS show promise for various groundwater applications, including identifying and assessing nitrate leached to groundwater on a regional scale [28,29]. The Bayesian approach uses prior and posterior probabilities of occurrence of an event for combining data sets. The posterior probability can be computed and updated based on the prior probability and addition of new evidence. The Bayesian model, in a log-linear form, known as the weights of evidence model, is a data-driven model that estimates the relative importance of evidence of data by statistical means. The Bayesian approach can be used to build a map showing areas vulnerable to groundwater contamination from nitrates and other pollutants using factors considered in DRASTIC or SEEPAGE and a water quality database of observed contamination levels in drinking water wells in a region. Figure 30.8 shows the Bayesian risk map for Indiana developed using the USGS water quality database and DRASTIC factors as evidence [23]. The data layers required for the analyses were integrated within the ARC/Info GIS.

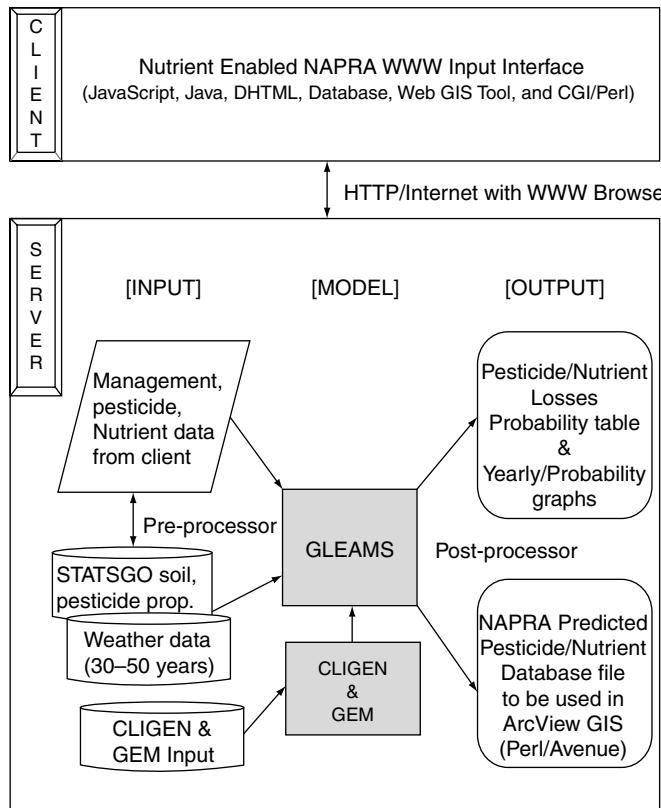


FIGURE 30.6 Overview of the NAPRA WWW System [3].

30.6.7 Modeling Groundwater Movement

Groundwater flow modeling is usually completed using a 2D or 3D finite-difference or finite-element approach. In 2D, the finite-difference method employs a rectangular discretization in which input parameters are assigned to each pixel and hydraulic head is computed for each cell. A 3D finite-difference can be visualized as a stack of grids with interaction between the grids as specified. Both common GIS data models, vector and raster, can be used to model groundwater movement using a finite-difference approach, but the grid/raster approach is more analogous to a finite-difference approach as each parameter/input into the model can be represented as a separate GIS data layer [30].

In the finite-element approach, a polygon network, typically a triangular network, is used to represent 2D flow in the aquifer [31]. GIS are often equipped with functions for creating triangular irregular networks (TIN), and this approach is similar to the finite-element approach. Hence, a vector data model with TIN capabilities may be more suitable for finite-element mesh generation for modeling groundwater movement.

GIS can facilitate the development, calibration, and verification of models as well as the display of model parameters and results in modeling groundwater movement. Spatial statistics and grid design capabilities of GIS can improve the modeling effort and aid in reliability assessment. A GIS-MODFLOW interface is an example of a groundwater finite-difference model linked with a GIS. Most parameters to the MODFLOW need to be represented in a raster format. There may be the case that there are no data in a cell after conversion to the raster data. Spatial interpolation algorithms in the GIS, such as Inverse Distance Weighting (IDW), can be used to interpolate point data [32]. Usually GIS and the model are autonomous and are linked by an interface or a set of programs. Typically, these programs simply convert the data from



FIGURE 30.7 NAPRA WWW predicted nitrate losses to shallow groundwater [28].

the GIS to a format that can be read by the model and the model output is converted for display in a GIS. GIS is primarily used for data management and spatial analysis capabilities in such interfaces.

Various functions are also generally available in a GIS that can be used for modeling groundwater advection and dispersion. The first step in groundwater flow modeling is to determine the flow velocity and direction at each point in the flow field. A flow field can be generated using Darcy's law for a regular grid. A volume balance can be computed for each cell. A zero volume balance residual might indicate a balance between inflow and outflow in a cell. This represents the steady state condition of groundwater flow.

30.6.8 Implementation of Policies Involving Spatial Dimensions

Groundwater is clearly one of the most important natural resources. The government has mandated protection of groundwater systems through programs that identify and protect municipal wellfields by designating wellhead protection areas. A wellhead protection area (WHPA) is defined as the surface and subsurface areas surrounding a public well or wellfield through which contaminants are likely to move toward and reach the well or wellfield. Five general criteria in wellhead protection as identified by EPA are: (1) distance, (2) drawdown, (3) time of travel, (4) flow boundaries, and (5) assimilative capacity. A GIS can be used to integrate the above concepts and synthesize the inputs and outputs required for the analysis [33]. The analytical capabilities of GIS software can be used for site characterization. Both the logical and visual overlay functions, as well as distance measurement capabilities in a GIS, can be used to define conditions and locations meeting specified protection criteria.

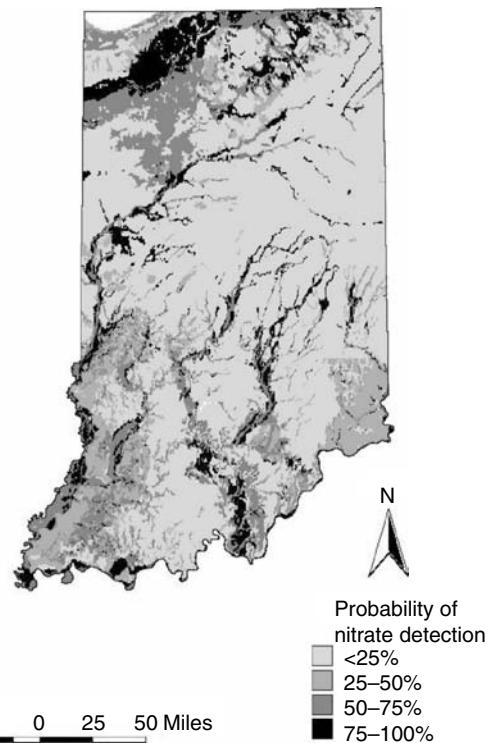


FIGURE 30.8 Bayesian risk map of nitrate detections in groundwater using DRASTIC factors.

GIS technology can be used for supporting remedial investigation of groundwater contamination. GIS can be used to integrate several software packages required for assessing environmental contamination during site characterization. Interfacing GIS with a groundwater movement model, such as MODFLOW, can facilitate investigation of 3D plumes [34].

GIS technology helps organize data about problems occurring over a geographic region and understand their spatial association, and provides a powerful means of analyzing and synthesizing information of such problems. GIS can also provide information to decision makers about the limitations of the accuracy of the data and the final products produced by GIS analyses.

Glossary

Buffer A zone of a specified distance around coverage features. Buffers are useful for proximity analysis (e.g., find all wetlands within 150 m of streams).

Cell Basic element of spatial information in a grid or raster data set.

Coordinate system A system used to measure horizontal and vertical distances on a planimetric map. Its units and characteristics are defined by a map projection. A common coordinate system is used to spatially register geographic data for the same area.

Coverage A set of thematically associated data considered as a unit. A coverage usually represents a single theme or layer such as roads, streams, land use, or aquifer depth.

Digitize To encode map features as x, y coordinates in a digital form. A digitizer (a table device with a cursor and keys for tracing map features) is used to create GIS coverages.

Geographic data The locations and descriptions of geographic features. The composite of spatial data and descriptive data.

Geographic database A collection of spatial data and related descriptive data organized for efficient storage and retrieval.

Georeference To establish the relationship between coordinates on a planar map and known real-world coordinates.

GIS Geographic information system. An organized collection of computer hardware, software, geographic data, and personnel to efficiently capture, store, update, manipulate, analyze, and display geographically referenced information.

Intersect Topological integration of spatial data sets. Common analysis function within GIS software.

Layer A set of thematic data stored for use within GIS. Layers organize data by subject matter (e.g., aquifer depths, slope, roads).

Line A set of ordered coordinates that represent the shape of linear features within a GIS coverage or layer. Data such as streams and roads are commonly represented as lines.

Map projection Conversion of locations on the Earth's surface from spherical to planar coordinates.

A projection is required to define a map's coordinate system.

Point An x, y coordinate that represents a geographic feature. Data representation technique used within GIS to represent small features such as well locations.

Polygon A multisided figure used to represent an area on a map. A polygon is defined by series of arcs. Data such as the locations of soil series are commonly represented as polygons within a GIS coverage or layer.

Raster A cellular or grid data structure composed of rows and columns. Groups of cells represent features such as land use or elevation. The value of each cell represents the value of the feature. Image and remotely sensed data is commonly stored in this format.

Resolution Resolution is the accuracy at which a map scale can depict the location and shape of map features. As map scale decreases, resolution diminishes and feature boundaries must be smoothed, simplified, or not shown at all.

Scanning Data input in raster format with a device called a scanner. Used to convert paper-based maps into GIS coverages or layers.

Spatial analysis The process of modeling, examining, and interpreting model results. Spatial analysis is the process of extracting or creating new information about a set of geographic features. Spatial analysis is useful for evaluating suitability and capability, for estimating and predicting, and for interpreting and understanding.

Spatial data Information about the location, shape of, and relationships among geographic features usually stored as coordinates and topology.

Surface Representation of geographic information as a set of continuous data in which the map features are not spatially discrete. Surfaces are used to represent spatial features such as depth to groundwater and are generally created by fitting a surface to point data.

Topology The spatial relationships between connecting or adjacent coverage features. Topology is useful in GIS because many spatial analysis functions do not require coordinates, only topological information.

References

- [1] Alterkawi, M.M. Measures towards a comprehensive municipal GIS-the case of Ar-Riyadh Municipality. *Habitat International* 29:689–698. 2005.
- [2] Griffin, C.B. Introduction to GIS for Public Agencies. Technical Report, Conservation Technology Information Center, State University of New York, College of Environmental Science and Forestry, Syracuse, NY, 1995.
- [3] Lim, K.J. and Engel, B.A. Extension and enhancement of national agricultural pesticide risk analysis (NAPRA) WWW decision support system to include nutrients. *Computers and Electronics in Agriculture* 38: 227–236. 2003.

- [4] Pandey, S., Gunn, R., Lim, K.J., Engel, B.A., and Harbor, J. Developing web-based tool to assess long-term hydrologic impacts of land use change: information technology issues and a case study. *Journal of Urban and Regional Information System Association (URISA)* 12: 5–17. 2000.
- [5] Lim, K.J. and Engel, B.A. RunOff MINimization (ROMIN) Web GIS for environment-friendly land use planning. 2004. <http://pasture.ecn.purdue.edu/~romin>. Accessed September 2005.
- [6] Tang, Z., Engel, B.A., Choi, J., Sullivan, K., Sharif, M., and Lim, K.J. A Web-based DSS for erosion control structure planning. *Applied Engineering in Agriculture* 20:707–714. 2004.
- [7] Engel, B.A., Choi, J., Harbor, J., and Pandey, S. Web-based DSS for hydrologic impact evaluation of small watershed land use changes. *Computers and Electronics in Agriculture* 39: 241–249. 2003.
- [8] Lim, K.J., Engel, B.A., Tang, Z., Choi, J., Kim, K., Muthukrishnan, S., and Tripathy, D. Automated Web GIS-based hydrograph analysis tool, WHAT. *Journal of American Water Resource Association* 41(6):1407–1416. 2005.
- [9] Keen, K.L. and Queen, L.P. An approach for continuously monitoring the recharge flux to the water table: Integration of field monitoring, laboratory characterization, and theory within GIS framework. In D.L. Corwin and R.J. Wagnet (ed), *Applications of GIS to the Modeling of Non-Point Source Pollutants in the Vadose Zone*. 1996 Bouyaucos Conference, USDA-ARS. 1996.
- [10] Fayer, M.J., Gee, G.W., Rockhold, M.L., Freshley, M.D., and Walters, T.B., Estimating recharge rates for a groundwater model using GIS. In D.L. Corwin and R.J. Wagnet (ed), *Applications of GIS to the Modeling of Non-Point Source Pollutants in the Vadose Zone*. 1996 Bouyanous Conference, USDA-ARS. 1996.
- [11] Levine, N.S. The Development and Use of a GIS-Based Expert System for Hazardous Waste Landfill Siting. Ph.D. thesis, Purdue University, December, 1995.
- [12] Schubert, M., Pena, P., Balcazar, M., Meissner, R., Lopez, A., and Flores, J.H. Determination of radon distribution patterns in the upper soil as a tool for the localization of subsurface NAPL contamination. *Radiation Measurement* 40:633–637. 2005.
- [13] Winiwarter, W., Dore, C., Hayman, G., Vlachogiannis, D., Gounaris, N., Bartzis, J., Ekstrand, S., Tamponi, M., and Maffei, G. Methods for comparing gridded inventories of atmospheric emissions — application for Milan province, Italy and Greater Athens Area, Greece. *The Science of the Total Environment* 303:231–243. 2003.
- [14] ESRI. *Using ArcGIS Geostatistical Analyst*. Redlands, CA. 2004.
- [15] Ruhaak, W. A Java application for quality weighted 3-d interpolation. *Computers & Geosciences* 32:43–51. 2006.
- [16] Follet, R.F., Shaffer, M.J., Brodahl, M.K., and Reichman, G.A. NLEAP simulation of residual soil nitrate for irrigated and non-irrigated corn. *Soil and Water Conservation* 49:375–382. 1994.
- [17] Tait, N.G., Davison, R.M., Whittaker, J.J., Lehrane, S.A., and Lerner, D.N. Borehole Optimisation System (BOS) — A GIS-based risk analysis tool for optimising the use of urban groundwater. *Environmental Modeling & Software* 19:1111–1124. 2004.
- [18] Wylie, B.K., Shaffer, M.J., and Hall, M.D. Regional assessment of NLEAP NO₃-N leaching indices. *Water Resources Bulletin* 31:399–408. 1995.
- [19] Luzio, M.D., Srinivasan, R., and Arnold, J.G. Integration of Watershed Tools and SWAT model into BASINS. *Journal of the American Water Resources Association* 38:1127–1141. 2002.
- [20] Dixon, B. Groundwater vulnerability mapping: A GIS and fuzzy rule based integration tool. *Applied Geography* 25:327–347. 2005.
- [21] Aller, L., Bennett, T., Lehr, J.H., and Petty, R.J. DRASTIC: A standardized system for evaluating groundwater using hydrogeologic settings. Technical report, USEPA EPA/600/2-85/0108, Robert S. Kerr Environmental Research Laboratory, Ada, OK. 1985.
- [22] Aller, L., Bennet, T., Lehr, J.H., Petty, R.J., and Hackett, G. DRASTIC: A standardized system for evaluating groundwater pollution using hydrogeologic settings. Technical report, USEPA EPA-600/2-87-035. 1987.
- [23] Navulur, K.C. and Engel, B.A. Groundwater vulnerability assessment to nonpoint source nitrate pollution on a regional scale using GIS. *Transaction of ASAE* 41:1671–1678, 1998.

- [24] Quisenberry, V.L., Smith, B.R., Phillips, R.E., Scott, H.D., and Norcliff, S. A soil classification system for describing water and chemical transport. *Soil Science Society of America Journal* 156:306–315. 1993.
- [25] Lin, H.S., McInnes, K.J., Wilding, L.P., and Hallmark, C.T. Effects of soil morphology on hydraulic properties: I. Quantification of soil morphology. *Soil Science Society of America Journal* 63:948–953. 1999.
- [26] Scott, H.D. *Soil Physics: Agricultural and Environmental Applications*. Ames, IW: Iowa State University Press. 2000.
- [27] Leonard, R.A., Knisel, W.G., and Still, D.A. 1987. GLEAMS: Groundwater loading effects of agricultural management systems. *Transactions of ASAE* 30:1403–1418.
- [28] Lim, K.J., Engel, B.A., and Tang, Z. Identifying regional groundwater risk areas using a WWW GIS model system. Accepted for the Special issue of *International Journal of Risk Assessment and Management*. 2006.
- [29] Bates, L.E., Barber, C., and Otto, C.J. Aquifer vulnerability mapping using GIS and Bayesian weights of evidence: Review of application. In *Hydro GIS*, Vienna, Austria, 1996.
- [30] Bonham-Carter, G.F. Geographic Information Systems for Geoscientists: Modeling with GIS. *Computer Methods in Geosciences* 13:302–303. 1994.
- [31] Watkins, D.W., McKinney, D.C., Maidment, D.R., and Lin, M., Use of Geographic Information Systems in Ground-Water Flow Modeling. *Water Resources Planning and Management* 122:88–96. 1996.
- [32] Fetter, C.W. *Applied Hydrogeology*. Third Edition. Prentice Hall, Upper Saddle River, NJ, USA. 1994.
- [33] Vieux, B.E., Mubaraki, M.A., and Brown, D. Wellhead protection area delineation using a coupled GIS and groundwater model. *Journal of Environmental Management* 54:205–214. 1998.
- [34] Juan, C.S. and Kolm, K.E. Conceptualization, characterization and numerical modeling of the Jackson Hole alluvial aquifer using Arc/INFO and MODFLOW. *Engineering Geology* 42:119–137. 1996.

Further Information

- Aronoff, S. 1989. *Geographic Information Systems: A Management Perspective*, WDL Publications, Ottawa, Canada.
- Bolstad, P. 2002. *GIS Fundamentals: A First Textbook on Geographic Information Systems*. Eider Press, OH.
- Clarke, K.C. 1997. *Getting Started with Geographic Information Systems*. Prentice Hall, Upper Saddle River, NJ.
- Corwin, D.L., Loague, K., Bigham, J.M., Kral, D.M., and Viney, M.K. 1996. *Applications of GIS to the Modeling of Non-Point Source Pollutants in the Vadose Zone*, SSSA No. 48. Soil Science Society of America, Madison, WI.
- Goodchild, M.F., Parks, B.O., and Steyart, L.T. (ed). 1993. *Environmental Modeling with GIS*. Oxford University Press, New York NY.
- Goodchild, M.F., Steyaert, L.T., Parks, B.O., Johnston, C., Maidment, D., Crane, M., and Glendinning, S. (eds). 1996. *GIS and Environmental Modeling: Progress and Research Issues*. GIS World, Fort Collins, CO.
- Heit, M. and Shortreed, A. 1991. *GIS Applications in Natural Resources*, GIS World, Fort Collins, CO.
- Longley, P.A., Goodchild, M., Maguire, D.J., and Rhind, D.W. 2002. *Geographic Information Systems and Science*. John Wiley & Sons, New York, NY.
- Maguire, D., Goodchild, M., and Rhind, D. (eds). 1991. *Geographical Information Systems*. John Wiley & Sons, New York, NY.
- Pinder, G.F. 2002. *Groundwater Modeling Using Geographic Information Systems*. John Wiley & Sons, New York, NY.

31

Biodegradation

31.1	Introduction.....	31-1
	Subsurface Microbiology and Geochemistry •	
	Environmental Conditions That Influence Biodegradation •	
	Biotransformations • Limits to Biodegradation	
31.2	Quantitative Description of Reactive Transport and Biodegradation.....	31-12
	Microbial Reaction Kinetics • Sorption as a Reactive Kinetic Process • Microbial Reactions and Sorption in a Batch System • Microbial Reaction, Sorption, and Transport	
31.3	Field Applications of Bioremediation	31-20
	Intrinsic Bioremediation • Engineered Bioremediation • Case Histories • Emerging Technologies	
	Glossary.....	31-22
	References	31-23
	Further Information	31-27

Loring F. Nies and
Chad T. Jafvert
Purdue University

31.1 Introduction

The objective of this chapter is to present the basic principles and concepts of groundwater microbiology and transport processes as they relate to biodegradation. Understanding of these processes is necessary to model, forecast, and manipulate subsurface contaminant transport and biodegradation. The basic principles presented may be applied to the biorestoration of contaminated groundwater, a field now commonly called “bioremediation” (Lee et al., 1988; National Research Council 1993; Thomas and Ward 1989).

31.1.1 Subsurface Microbiology and Geochemistry

31.1.1.1 Basic Microbiology

Microorganisms mediate the biotransformation of many different chemicals in the subsurface. Historically there has been much scientific debate over the relative contribution of abiotic and biotic chemical processes in the subsurface. Until the last 50 years, very little was known about microorganisms in groundwater, and thus abiotic transformations were thought to dominate. However, more recent research has contributed new knowledge about microbial life in the subsurface and this view has changed. While our understanding of subsurface microbiology and biogeochemistry is far from complete, biologically mediated reactions of both inorganic and organic compounds in groundwater are known to be significant, and in many cases control groundwater chemistry (Chapelle, 1993). With respect to the biotransformation of organic

pollutants, microbially mediated reactions are by far the most important. The abiotic transformation of pollutants does occur (e.g., dehalogenation, polymerization, and hydrolysis reactions), however, the environmental significance of these abiotic reactions relative to microbially mediated reactions is difficult to assess. This is partly due to the difficulty in distinguishing between abiotic and biotic processes *in situ*. In addition, certain abiotic reactions are dependent on two environmental parameters that are often controlled by microbial processes, redox, and pH. Thus these latter types of abiotic reactions are indirectly mediated by microorganisms as well.

Microorganisms can be classified in a number of different ways (e.g., phylogenetically). However, for the purpose of understanding biodegradation, it is most useful to begin the characterization of microorganisms according to their source of energy and their source of carbon for cell growth. Potential energy sources are organic carbon (including pollutants), inorganic compounds, and sunlight. Microbes that oxidize organic compounds to obtain energy are organotrophs and those that oxidize inorganic compounds are lithotrophs. Potential sources of carbon for cell growth are either organic carbon or inorganic carbon (HCO_3^- , CO_2). Microbes that degrade organic compounds to obtain carbon for the synthesis of cellular constituents are heterotrophs and those that utilize inorganic carbon are autotrophs. Logically, most lithotrophs are also autotrophs and most organotrophs are also heterotrophs. Photoautotrophs and photoheterotrophs may exist in groundwater, but their utilization of photometabolism is obviously limited by access to sunlight (Brock et al., 1997; Gottschalk, 1986).

The types of microorganisms found in groundwater include protozoa, fungi, bacteria (this chapter includes both bacteria and archaea), and viruses. For purposes of examining microscale phenomena (biodegradation) relative to macroscopic subsurface particle and pore sizes, it is useful to consider the size of the microorganism. Protozoa have a size range of approximately 2 to 200 μm , the smallest single-celled fungi are approximately 2 μm , and bacteria range in size from 0.1 to 15 μm , although a bacterium which is 500 μm long has recently been discovered. Viruses are much smaller, ranging from approximately 0.01 to 0.1 μm (Brock et al., 1997).

Protozoa are relatively large heterotrophic microorganisms whose distribution in groundwater has not been thoroughly studied. Protozoa are not known to be significant biodegraders of organic pollutants, although many protozoa do feed on bacteria. This predation is thought to reduce bacterial numbers in some groundwater (Chapelle, 1993). Thus, it is possible that protozoa could indirectly and adversely influence biodegradation rates in groundwater by reducing populations of pollutant biodegrading bacteria. Some protozoa are also human pathogens; therefore their distribution in groundwater is a public health interest (e.g., *Entamoeba histolytica*, the causative agent of dysentery). Drinking water is the primary route of human exposure to pathogenic protozoa. Fortunately, the cysts of the problematic pathogenic protozoa *Cryptosporidium* and *Giardia* are primarily detected only in surface waters (LeChevallier and Norton, 1995). However, groundwater under direct influence of surface water also very frequently tests positive for *Cryptosporidium* and *Giardia* (Hancock et al., 1998).

Fungi are heterotrophic biodegraders of immense capability. There are many types of single-celled microscopic fungi as well as multicelled macroscopic organisms, such as mushrooms. The ability of fungi to biodegrade organic pollutants is well known; however in groundwater the bacteria dominate in terms of both biomass and as catalysts of biodegradation. Therefore, all microbial biodegradation discussed in this chapter can be assumed to be mediated by bacteria, unless stated otherwise.

Viruses are obligate parasites whose hosts include bacteria as well as macroscopic organisms. The transport and survival of viruses in groundwater is of interest with respect to the control of infectious diseases (Bitton et al., 1983). Viruses have no catabolic capability and therefore do not contribute directly to biodegradation. The influence of the viral infection of bacteria on biodegradation, and on microbial ecology in general, has not been systematically examined. As viruses are essentially containers of genetic material that is inserted into the host, they can be used as gene vectors to insert desirable genetic information (e.g., biodegradation genes) into bacteria. While this type of technology is available in the laboratory, it is not currently used in the field.

Bacteria are ubiquitous in groundwater, and in addition, they have very diverse catabolic abilities, morphology, physiology, and biochemical constituents. This diversity allows bacteria to survive in some

of the most extreme environments on earth. Bacteria have been found up to 2.8 km below the surface (Fredrickson and Onstott, 1996). Although they are microscopic, bacteria are the dominant biodegraders and drivers of biogeochemical cycling in groundwater. Bacteria live attached to subsurface particles as individuals or in colonies and also as unattached motile organisms. Attached microbes conserve energy while removing nutrients from the surrounding water, as well as deriving some protection from predation. The growth and survival of attached organisms is dependent on obtaining nutrients by transport through groundwater. Motile organisms are known to move in response to chemical gradients (chemotaxis), moving toward higher concentrations of nutrients and away from toxics. Bacterial transport in groundwater and environmental factors that induce bacteria to exhibit attached or motile phenotypes have been studied (Dawson et al., 1981; Fletcher and Marshall, 1982; Corapcioglu and Haridas, 1984, 1985; Harvey et al., 1989). However, our ability to control and influence these processes in the subsurface to promote biodegradation is still under development (Gentry et al., 2004).

31.1.1.2 Basic Metabolism

All living organisms obtain energy by mediating oxidation/reduction reactions. Reduced organic and inorganic compounds serve as the reductant (electron donor) in the reaction. Oxidized, usually inorganic, species serve as the oxidant (electron acceptor). Aerobic respiring organisms utilize oxygen, which is the most ubiquitous electron acceptor in the environment, as an oxidant. However, the solubility of oxygen in water is relatively low ($\sim 8 \text{ mg/l}$ at 20°C) and the oxygen diffusion rate is slow, therefore, the availability of oxygen in groundwater is significantly limited by mass transfer. A variation of typical oxidation/reduction reactions occurs when a fraction of a compound is oxidized and the remaining fraction is reduced. This type of catabolism is called fermentation, in contrast to respiration. Almost all fermenting bacteria cannot tolerate oxygen and are found in anaerobic environments living in close association with sulfate reducing or methanogenic bacteria (Brock et al., 1997; Chapelle, 1993; Gottschalk, 1986).

Organisms have evolved elegant mechanisms for converting the electron transfer that occurs during oxidation/reduction reactions into energy. Respiring organisms oxidize their energy source, stripping off the high energy electrons. The electrons are shuttled on electron carriers (NADH) to the cell membranes where electron transport phosphorylation occurs (Figure 31.1). Respiring organisms use electron transport to pump protons across the membrane to create a pH differential (charge and proton gradient) between the inside and outside of the cell. In this energized state, protons are driven by the gradient through specialized enzymes which convert the “proton motive force” into stored chemical energy, adenosine triphosphate (ATP). Driving the proton pumps consumes the electron’s energy. At the terminal end of the electron transport chain the now low energy electrons are transferred to a terminal electron acceptor.

Fermenting bacteria make ATP through the direct conversion of chemical bond energy, in a process called “substrate level phosphorylation (SLP).” The energy-producing steps primarily occur during oxidation of carbonyl groups (Figure 31.2). As no external electron acceptor is utilized, and biological electron carriers are in limited supply and must be recycled, the electrons generated from the energy-producing oxidations of SLP must be discarded. This is accomplished by reducing some of the substrate, producing alcohols, or by reducing protons, producing hydrogen. Thus fermenters produce a mixture of both oxidized and reduced products such as CO_2 , carboxylic acids, alcohols, and H_2 .

Anaerobic respiration occurs when oxygen is depleted and other suitable oxidized species are available to act as an electron acceptor. Common anaerobic electron acceptors are oxidized nitrogen compounds (nitrate, nitrite), iron(III), oxidized sulfur compounds (sulfate, sulfite, S°), and CO_2 . Other oxidized inorganic species are utilized in anaerobic respiration as well, and in addition, some organic compounds, such as fumarate, may also serve as electron acceptors (Zehnder, 1988). The widespread existence of anaerobic respiration can be exploited for groundwater remediation when the contaminants of interest can function as microbial electron acceptors. Examples include nitrate, nitrite, perchlorate, chlorate, and oxidized species of chromate and arsenate. It is critically important to have a thorough understanding of the chemodynamics of inorganic species, particularly chromate and arsenate for example, as redox transformations alter speciation, solubility, and toxicity. Reduction of chromate species reduces mobility

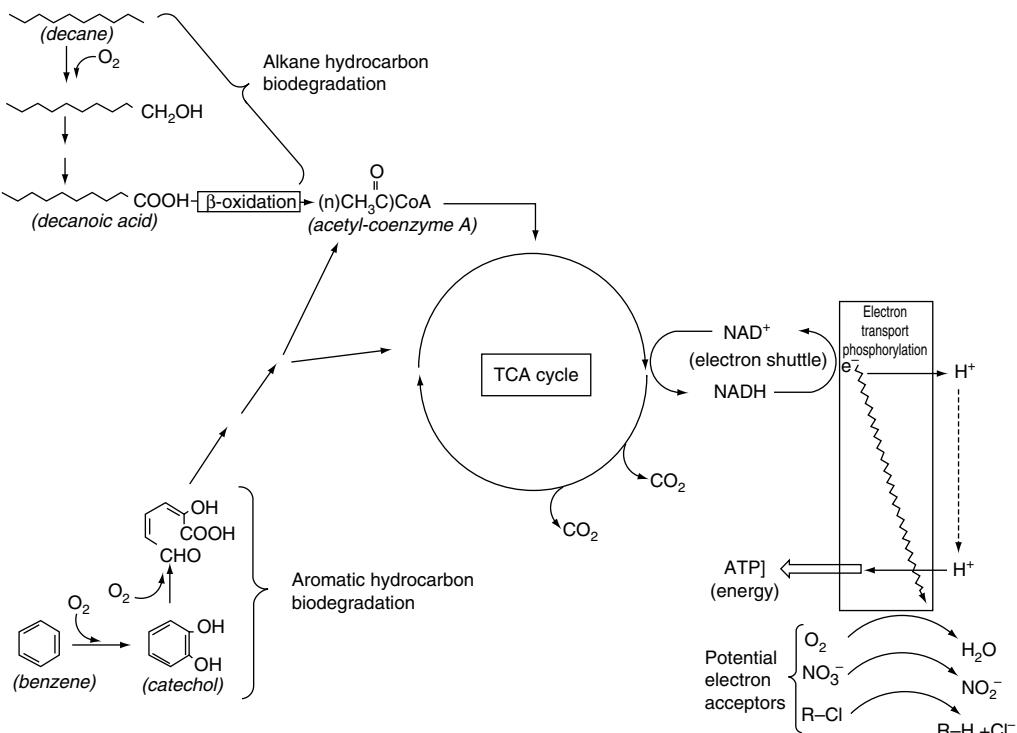


FIGURE 31.1 Idealized schematic of hydrocarbon biodegradation by a respiration organism.

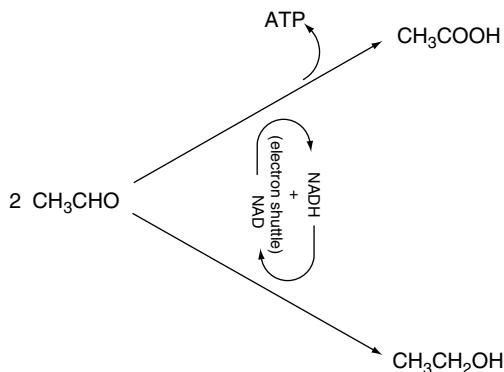


FIGURE 31.2 Schematic of fermentation.

and toxicity, while reduction of arsenate species may actually increase mobility in some instances. Therefore, redox transformations of inorganics may be desirable or undesirable and their long-term impacts must be assessed as well. Perchlorate is one inorganic groundwater contaminant with a high potential for successful biological transformation to the harmless end-product, chloride ion (Logan, 1998, 2001). Acquisition of fundamental knowledge about the microbiology and genetics of perchlorate reduction has been very significant (Bender et al., 2005).

Bacteria obtain less energy from using electron acceptors other than oxygen. The amount of energy obtained from the oxidation of a given substrate is proportional to the reduction potential (E°') of the electron acceptor (Table 31.1). Electron acceptors that yield the largest amount of energy (e.g., oxygen) tend to be utilized preferentially over others that are available, probably because bacteria with access

TABLE 31.1 Reduction Potentials of Important Biological Oxidants
(electron acceptors)

	$E^{\circ'}$ (Volts)
$\frac{1}{4}\text{O}_2(\text{g}) + \text{H}^+ + \text{e}^- = \frac{1}{2}\text{H}_2\text{O}$	0.820
$\frac{1}{2}\text{MnO}_2(\text{s}) + \frac{1}{2}\text{HCO}_3^-(10^{-3}) + \frac{3}{2}\text{H}^+ + \text{e}^- = \frac{1}{2}\text{MnCO}_3(\text{s}) + \text{H}_2\text{O}$	0.527
$\frac{1}{2}\text{NO}_3^- + \text{H}^+ + \text{e}^- = \frac{1}{2}\text{NO}_2^- + \frac{1}{2}\text{H}_2\text{O}$	0.423
$\frac{1}{6}\text{NO}_2^- + \frac{4}{3}\text{H}^+ + \text{e}^- = \frac{1}{6}\text{NH}_4^+ + \frac{1}{3}\text{H}_2\text{O}$	0.344
$\text{FeOOH}(\text{s}) + \text{HCO}_3^-(10^{-3}) + 2\text{H}^+ + \text{e}^- = \text{FeCO}_3(\text{s}) + 2\text{H}_2\text{O}$	−0.047
$\frac{1}{6}\text{SO}_4^{2-} + \frac{4}{3}\text{H}^+ + \text{e}^- = \frac{1}{6}\text{S}(\text{s}) + \frac{2}{3}\text{H}_2\text{O}$	−0.195
$\frac{1}{2}\text{S}(\text{s}) + \text{H}^+ + \text{e}^- = \frac{1}{2}\text{H}_2\text{S}(\text{g})$	−0.243
$\frac{1}{8}\text{CO}_2(\text{g}) + \text{H}^+ + \text{e}^- = \frac{1}{8}\text{CH}_4(\text{g}) + \frac{1}{4}\text{H}_2\text{O}$	−0.244
$\text{H}^+ + \text{e}^- = \frac{1}{2}\text{H}_2^-$	0.414

Values apply for unit activity (1 M or 1 atm) in water at pH 7.0 and 25°C, except $\text{HCO}_3^- = 10^{-3}$ M which more typically represents environmental conditions.

$\Delta G^{\circ'} = -nFE^{\circ'}$, where n is the number of electrons transferred and F is Faraday's constant. $F = 96.5 \text{ kJ/V-mol}$.

Source: Adapted from (Brock, T.D., M.T. Madigan, J.M. Martinko, and J. Parker. 1997. *Biology of Microorganisms*, Prentice Hall, Upper Saddle River, NJ.; Stumm, W. and J.J. Morgan. 1981. *Aquatic Chemistry*, John Wiley & Sons, Inc., New York.; Thauer R.K., K. Jungermann, and K. Decker. 1977. *Bacteriol. Rev.*, 41:100–180.)

to more energy can grow faster and compete for resources more successfully. In addition, biochemical mechanisms exist that maximize energy production. For example, facultative aerobes that can use both oxygen and nitrate as electron acceptors are common. The presence of oxygen inhibits synthesis of the enzyme at the terminus of the electron transport chain that transfers electrons to nitrate (nitrate reductase), thus ensuring more energetically favorable oxygen utilization whenever possible.

During biodegradation of organic compounds, large molecules are broken down into small molecules, which are oxidized, yielding electrons for energy production. For example, aromatic rings are cleaved open to form aliphatic chains, which are then cleaved into two-carbon pieces, which are subsequently oxidized to CO_2 (Figure 31.1). Likewise, long alkanes are cleaved into two-carbon pieces prior to oxidation to CO_2 . However, the initiation of biodegradation of many compounds requires “activation” by the insertion of molecular oxygen (O_2) into the hydrocarbon, forming mono- and dihydroxylated compounds. Oxygenase enzymes perform this function. Once activated by oxygenase enzymes, further biodegradation of hydrocarbons often does not require molecular oxygen. An exception is the biodegradation of aromatic rings, during which oxygen is used as a reactant for ring activation and ring cleavage. Thus, it should be observed that oxygen can be utilized by bacteria as both an electron acceptor and as a reacting co-substrate for initiating biodegradation reactions. A critical issue in biodegradation is that there are many alternatives to oxygen as an electron acceptor, while there are few alternatives to oxygen as a reactant for initiating hydrocarbon biodegradation. For this reason, when more rapid and complete biodegradation of certain hydrocarbons is desired, aerobic conditions are preferred. The bioremediation of groundwater aquifers that are contaminated with hydrocarbons is often limited by insufficient oxygen. Hydrogen peroxide, which decomposes to oxygen, is often used as a highly soluble source of oxygen for groundwater systems (Pardieck et al., 1992).

One good alternative to oxygen as an electron acceptor is nitrate (NO_3^-). Significant research also has been devoted to the feasibility of using nitrate as a supplemental electron acceptor (Dolfing et al., 1990; Evans et al., 1991; Hutchins, 1991a, 1991b). Nitrate is more soluble in water than oxygen, is a strong oxidant, and could potentially be more readily introduced into groundwater aquifers through injection wells. Interestingly, many monoaromatic compounds (BTEX) are biodegraded under anaerobic nitrate reducing conditions, although unlike oxygen, the nitrate is used only as an electron acceptor and not as a reactant in biotransformation reactions. Nitrate has relatively low toxicity, but does have known health

hazards (methemoglobinemia) and is a regulated drinking water contaminant, and therefore, is used as a supplemental electron acceptor only under carefully controlled situations.

31.1.2 Environmental Conditions That Influence Biodegradation

The activity and types of microorganisms present in groundwater are greatly influenced by subsurface physical and chemical properties (Alexander, 1994). Biodegradation kinetics are generally highly dependent on temperature because enzyme function is temperature sensitive. Bacteria can be classified by the range of temperature within which they can grow. At optimum temperatures biodegradation kinetics reach maximum rates, while slightly above optimum temperature, cell constituents usually begin to degrade. As temperatures decrease below the optimum, biodegradation rates decrease until enzymes function at rates too slow to support growth. Temperature extremes will select for climate-adapted organisms. For example, cold-loving bacteria (psychrophiles) can grow at temperatures below 0°C and generally die at temperatures greater than 20°C. Mesophilic bacteria thrive at temperatures that are comfortable to humans. Although the vast majority of biodegradation studies have been done with mesophilic bacteria at temperatures ranging between 15 and 45°C, some evidence exists that suggests that psychrophilic (or at least psychrotolerant) (Kellems et al., 1994) and thermophilic bacteria also possess pollutant biodegrading ability (Chen and Taylor, 1995).

Bacteria can also be classified according to the pH range within which they can grow. Our knowledge of biodegradation has been derived chiefly from studies conducted in the “neutral” pH range between approximately 6 and 8. Most groundwater falls within this pH range as well. Acidophilic bacteria can tolerate a pH as low as 2, although most are lithoautotrophs that do not degrade organics. Alkaline conditions exist (pH 9 to 11) in certain areas where carbonate rocks predominate. Sodium concentrations in these alkaline environments are often more than 10× greater than seawater, thus indigenous bacteria are halophiles in addition to their adaptation to extreme pH. Knowledge of pollutant biodegradation in extreme pH environments is scarce.

Bacteria are sensitive to salinity. Halophilic bacteria have evolved a mechanism to counteract the tendency for cells to become desiccated by high salt concentrations. Numerous studies have demonstrated that marine bacteria are capable of hydrocarbon biodegradation, however, knowledge about pollutant biodegradation by extreme halophiles is lacking. Because a primary concern is the bioremediation of contaminated aquifers that are sources of potable water, acquiring knowledge about organisms adapted to extreme environments has not been a high priority.

All living things, including bacteria require inorganic nutrients for growth, in addition to a source of energy and carbon. Nitrogen and phosphorus are the nutrients that are most likely to be a limiting factor for biodegradation and, similar to agricultural applications, are often added to the environment to remove potential nutrient limitations.

Oxidation/reduction potential is by definition the electrical potential (in volts) of the oxidation/reduction reaction occurring between the electron donor and electron acceptor. In practice, when considering conditions in the environment the term “redox” potential is commonly used. Redox potential typically refers to the reduction potential of the dominant electron acceptor in the environment. Probes are available to measure environmental redox potential; however, when assessing biodegradation, actual measurement of the electron acceptor of interest is more useful. For example, if aerobic biodegradation is desired oxygen concentrations should be monitored. Likewise, the simultaneous disappearance of nitrate and organics, with the concurrent appearance of nitrite is better circumstantial evidence of the existence of denitrifying bacteria than a redox measurement. More reducing, anaerobic conditions can be inferred from measurements of sulfate, sulfide, and the production of methane, as well as redox measurements.

31.1.3 Biotransformations

Bacteria biodegrade organic compounds by breaking large molecules apart with an array of biotransformation reactions such as hydrolysis, oxidation, reduction, dehalogenation, deamination,

TABLE 31.2 Biotransformation Reactions Mediated by Oxygenase Enzymes

Biotransformation	Compound	Reference
Deamination	Aniline	Bachofer et al. (1975)
Dehalogenation	Pentachlorophenol	Xun et al. (1992)
Ether cleavage	2,4,-dichlorophenoxyacetate	Tiedje and Alexander (1969)
Aromatic ring hydroxylation	Benzene	Gibson (1984)
Aromatic ring fission	Benzene	Gibson (1984)
Alkane hydroxylation	Octane	Watkinson and Morgan (1990)

decarboxylation, and rearrangement reactions. As mentioned previously, molecular oxygen participates directly in a number of biotransformation reactions (Table 31.2).

Anaerobic bacteria can perform many of the biotransformation reactions listed in Table 31.2 by different mechanisms without molecular oxygen (Schink, 1988). In addition, anaerobic bacteria can mediate reductive dehydroxylation, reductive deamination, and reductive dehalogenation reactions. Of these, reductive dehalogenations are of the greatest interest with respect to the biotransformation of hazardous pollutants. No anaerobic bacteria have been shown to be able to hydroxylate alkanes, thus saturated hydrocarbons persist under anaerobic conditions. While BTEX biodegradation by denitrifying and iron reducing bacteria has been well documented, and a few reports of aromatic hydrocarbon degradation under sulfate reducing and methanogenic conditions exist, detailed knowledge about the biochemistry of anaerobic aromatic hydrocarbon transformation is limited (Evans, 1988; Lovley et al., 1989; Lovley and Lonergan, 1990).

In most cases, bacteria biodegrade organic compounds to obtain energy for growth. Due to the relaxed specificity of some enzymes, bacteria perform certain biotransformation reactions on compounds that are not growth substrates. The term “cometabolism” can be broadly taken to mean the “gratuitous biotransformation of a compound from which the organism derives no benefit.” Often these gratuitous reactions produce products that can be more easily biodegraded by other organisms.

31.1.3.1 Hydrocarbons

Biodegradation of alkanes is usually initiated with terminal hydroxylation by a monooxygenase enzyme and subsequent oxidation of the alcohol to a carboxylic acid. The initial step requires oxygen, and thus far, no other anaerobic mechanism for the initiation of biodegradation of alkanes has been reported. The carboxylic acid can easily be further oxidized to CO₂ through two nearly universal biochemical pathways, β-oxidation and the Krebs (tricarboxylic acid) cycle (Figure 31.1). Branched alkanes can be more difficult to degrade depending on the degree of branching. Branching interferes with β-oxidation and significant branching can result in complete inhibition of biodegradation (Watkinson and Morgan, 1990). Alkenes can be aerobically biodegraded similarly to alkanes, however the double bond can also be hydrolyzed to initiate biodegradation under anaerobic conditions (Schink, 1988). As alkanes have relatively low water solubilities they are less of a hazard to migrate as soluble constituents of groundwater. Biodegradation of alkanes often occurs at the hydrocarbon/water interface and is mediated by bacteria that produce biosurfactants for hydrocarbon uptake.

As described previously, aerobic biodegradation of aromatic compounds is initiated by the hydroxylation of the ring by oxygenase enzymes. Ring cleavage also requires oxygen. The aromatic compounds of most common concern are benzene, ethylbenzene, toluene, o-xylene, m-xylene, and p-xylene (BTEX). Initiation of the biodegradation of the alkylbenzenes occurs either by dioxygenase attack on the aromatic ring or by monooxygenase attack on the alkyl group followed by a dioxygenase mediated ring cleavage (Smith, 1990). Biodegradation rates are generally observed in the following order which may vary from site to site: toluene, ethylbenzene > benzene > m-xylene, p-xylene > o-xylene. BTEX are constituents of gasoline and therefore often appear together as a mixture in contaminated groundwater. Interactions between bacteria and BTEX mixtures may be complex and site-specific (Alvarez and Vogel, 1991). Biodegradation of monoaromatic hydrocarbons (e.g., toluene, ethylbenzene) also occurs under anaerobic

nitrate reducing conditions. Benzene appears to be more recalcitrant under anaerobic conditions than other monoaromatics. The biodegradation of aromatic rings has also been observed under iron reducing, sulfate reducing, and methanogenic conditions (Grbic-Galic and Vogel, 1987; Lovley et al., 1989; Edwards et al., 1992).

Simple nonhalogenated hydrocarbon solvents such as ethanol, methanol, and acetone are easily biodegraded at dilute concentrations (e.g., 0.1% or less). The cyclical ethers, and ether structures in general, are typically recalcitrant; however the biodegradation of furans and 1, 4-dioxane has been reported, although biochemical mechanisms and the distribution of this ability in the environment remains unknown. Space limitations prohibit an exhaustive review of the biodegradation of all large production organics that might be found in groundwater, but several more complete references are available (Gibson, 1984; Howard, 1989; Leahy and Colwell, 1990; Young and Cerniglia, 1995).

31.1.3.2 Halogenated Compounds

Halogenated organic compounds are among the most problematic environmental pollutants encountered (Chaudhry and Chapalamadugu, 1991). Halogenation typically increases environmental stability and toxicity, and significantly alters the kinds of biochemical and chemical reactions compounds undergo (Vogel et al., 1987). Halogenation increases the oxidation state of a compound relative to analogous nonhalogenated compounds, and this significantly influences biodegradation as well. Several reviews of microbial transformation of halogenated compounds exist (Fetzner and Lingens, 1994; Mohn and Tiedje, 1992; Neilson, 1990). Biodegradation of halogenated compounds can be considered from two perspectives, bioenergetics and biochemical mechanisms. In general, extremely oxidized compounds are thermodynamically less favorable electron donors than reduced compounds, and therefore, as the degree of halogenation (and oxidation) increases compounds have fewer and fewer electrons to give up as electron donor and they potentially would yield correspondingly less energy when microorganisms oxidize them. Alternatively, polyhalogenated compounds are potentially good electron acceptors (Figure 31.3). Halogenated compounds acting as electron acceptors can undergo a reaction called reductive dehalogenation, in which two electrons are transferred to the compound, the halogen leaves as a halide ion, and is replaced by a hydrogen atom. Depending on the degree of halogenation, and the type of environment and microorganisms present where it is found, halogenated compounds may be used as either electron donors or electron acceptors, with the more halogenated compounds making better electron acceptors and the less halogenated compounds making better electron donors. Consideration of thermodynamics suggests that sequential anaerobic dechlorination followed by aerobic biodegradation would be successful.

Halogenated organics influence biodegradation mechanistically because of the large atomic size of halogens relative to hydrogen (which halogens usually replace), halogen electronegativity, and the strength of the carbon–halogen bond. Halogen size may prevent biochemical reactions simply due to steric hindrance. Halogen electronegativity causes charge separation in bonds and may result in dipole moments in molecules, profoundly affecting chemical reactivity. For example, halogen substitution may result in compounds more susceptible to nucleophilic substitution reactions (e.g., hydrolysis) whereas many oxygenases enzymes are strong electrophiles that are better suited to attack saturated non-halogenated hydrocarbons. Strong bonds require large activation energies for cleavage and may prevent reactions from occurring. For example, the carbon–fluorine bond is exceptionally strong and is rarely broken during biological processes.

31.1.3.3 Halogenated Aliphatic Solvents

The halogenated aliphatic solvents that are most commonly found in groundwater are the chlorinated methanes (e.g., carbon tetrachloride (CCl_4)) and chlorinated ethenes (tetrachloroethene or “perc” (PCE) and trichloroethene (TCE)) (Vogel et al., 1987; Schwarzenbach et al., 1993). Under anaerobic conditions, PCE can undergo stepwise reductive dechlorination to TCE, then dichloroethene (DCE), chloroethene (i.e., vinyl chloride (VC)), and finally ethene. The rates of dechlorination tend to be proportional to the number of chlorine, thus PCE dechlorinates faster than dichloroethene. In groundwater systems where PCE and TCE undergo reductive dechlorination, vinyl chloride often accumulates. As vinyl chloride is a

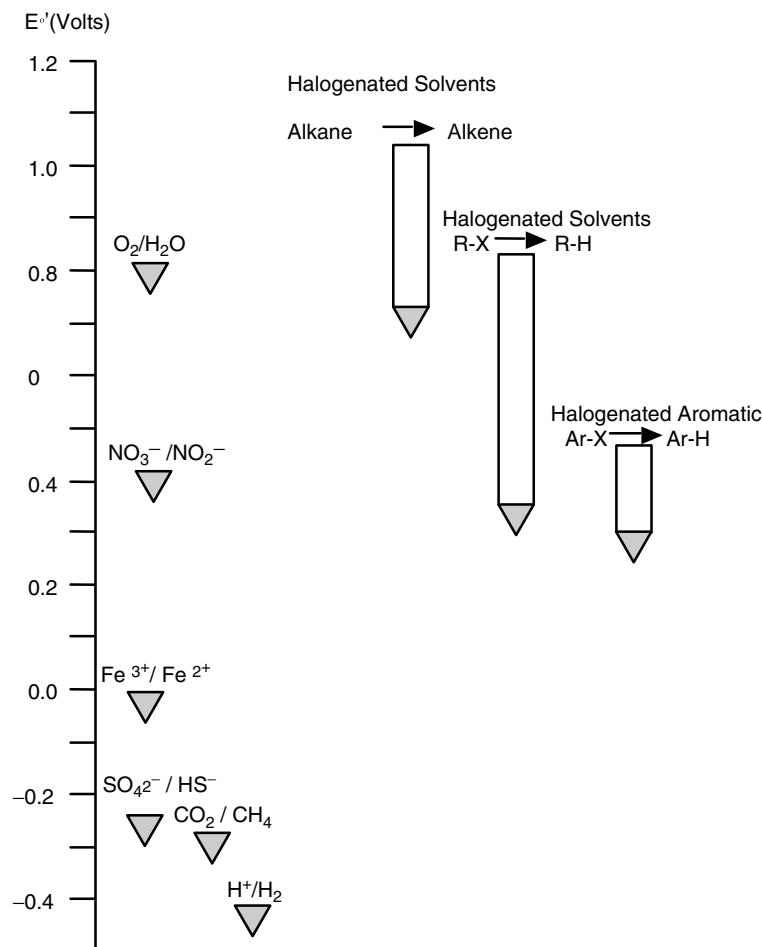


FIGURE 31.3 Redox potentials of biological electron acceptors and chlorinated compounds. (Brock, T.D., M.T. Madigan, J.M. Martinko, and J. Parker. 1997. *Biology of Microorganisms*, Prentice Hall, Upper Saddle River, NJ.; Dolfling, J. and B.K. Harrison. 1992. *Environ. Sci. Technol.*, 26:2213–2218.; Thauer R.K., K. Jungermann, and K. Decker. 1977. *Bacteriol. Rev.*, 41:100–180.; Holmes, D.A., Harrison, B.K., and Dolfling, J. 1993. *Environ. Sci. Technol.*, 27:725–731.; Schwarzenbach, R.P., P.M. Gbchwend, and D.M. Impoden. 1993. *Environmental Organic Chemistry*, John Wiley & Sons, Inc., New York.; Vogel, T.M., Criddle, C.S., and McCarty P.L. 1987. *Environ. Sci. Technol.*, 21: 2213–2218.)

known human carcinogen, the presence of this metabolite is extremely undesirable. It has recently been elucidated that there is a genetic basis for this phenomena and detection methods for the microbial ability to completely dechlorinate chloroethenes to ethene have been developed (Löffler et al., 2000). These novel microorganisms carry out “halorespiration” during which the chlorinated solvent is utilized as an energy producing electron acceptor. Molecular techniques have been used to characterize microbial communities with halorespiring capabilities at chlorinated solvent contaminated sites throughout North America and Europe (Hendrickson et al., 2002). Despite a generally poor history of success with bioaugmentation attempts, more recently a large number of successful field demonstrations have been completed with halorespiring bacteria (Ellis et al., 2000; Major et al., 2002). Advances and applications of halorespiration bioaugmentation in chlorinated solvent contaminated aquifers have been so substantial that the US Environmental Security Technology Certification Program commissioned a study to document them (Parsons Engineering Science, Inc., 2002).

Aerobically, TCE, DCE, and VC, but typically not PCE, can be co-metabolized by certain bacteria with monooxygenase enzymes. This phenomenon can potentially be exploited for use in the restoration of contaminated aquifers. These bacteria require specific growth substrates (e.g., methane or phenol) which induce synthesis of the monooxygenase enzymes that act on the chlorinated compounds. The metabolites of (at least TCE) co-metabolism are potentially toxic, and kill the cells mediating the reaction. As the growth substrate and the co-metabolized pollutant both compete for the same reactive enzyme site, a balance must be achieved between maintaining sufficient growth rates and acceptable degradation rates. Optimizing growth with excessive substrate could inhibit degradation of the pollutants, while low substrate concentrations may not induce the monooxygenase enzymes and biomass could be lost due to the production of suicide metabolites. For effective biodegradation, cell growth and pollutant biotransformation could potentially be separated in either space or time. This could be accomplished in groundwater by alternating periods of growth enhancement (adding growth substrate) with periods of starvation (Alvarez-Cohen and McCarty, 1991; Nelson et al., 1987).

31.1.3.4 Halogenated Aromatic Compounds

The biodegradation of halogenated aromatic compounds is highly dependent on the position and number of halogen substituents (Reineke and Knackmuss, 1988). It is useful to subdivide halogenated aromatics into two groups, non-ionizable compounds and ionizable compounds. PCBs and chlorobenzenes are examples of non-ionizable halogenated aromatic compounds found in the environment. In general, increased halogen substitution results in greater hydrophobicity. Therefore, the more soluble, less chlorinated PCB and chlorobenzene congeners are a greater threat to migration in groundwater relative to more highly chlorinated congeners. The biodegradation of PCBs and chlorobenzenes is highly dependent on the degree of chlorination as well (Furukawa et al., 1978; Bedard and Haberl, 1990). PCBs and chlorobenzenes can be aerobically degraded similar to non-halogenated aromatics; however, chlorine substitution often inhibits one or more enzymatic reactions. Thus, less chlorinated congeners are significantly more easily biodegraded by aerobic microorganisms than highly chlorinated compounds. Anaerobically, highly chlorinated PCBs and chlorobenzenes undergo microbially mediated reductive dechlorination to less chlorinated congeners that could potentially be biodegraded aerobically, but tend to persist under anaerobic conditions (Abramowicz, 1990). A well-documented field demonstration of PCB bioremediation was described by Harkness et al. (1993) during which it was found that bioavailability of PCBs severely limited biodegradation.

The behavior of ionizable aromatic compounds, such as phenols, anilines, and benzoates, is dependent on groundwater pH. For example, at a typical pH of 7, greater than 99% of the dissolved pentachlorophenol (PCP) ($pK_a = 4.75$) will exist in the more soluble, less hydrophobic, ionized form. Unlike most PCB and chlorobenzene congeners, the complete biodegradation of PCP to CO_2 occurs under both aerobic and anaerobic conditions. In the better understood aerobic process, a single microorganism can catabolize PCP for energy and carbon. Anaerobically, PCP is sequentially dechlorinated to phenol, which can be further degraded to methane and CO_2 . Relatively little is known about the microbiology and biochemistry of anaerobic PCP biodegradation (Haggblom and Valo, 1995; McAllister et al., 1996).

31.1.4 Limits to Biodegradation

It should be mentioned that the environmental conditions that influence biodegradation and the limits to biodegradation are intrinsically related, and therefore, should not necessarily be viewed as separate topics. The most extreme limit to biodegradation is the absence of any known biochemical mechanism for the transformation of a specific compound (McCarty and Semprini, 1993). Discoveries of new transformations and microbial evolution of new enzymes will continue to challenge this limitation (Shannon and Unterman, 1993). Biodegradation potential can often be predicted from structure-activity models and a review of the biodegradation literature (Scow, 1990; Huesemann, 1995). However, as microbial distribution and environmental conditions are extremely heterogeneous, actual site-specific assays, such as laboratory treatability studies, provide the most reliable evidence that *in situ* biodegradation is possible.

31.1.4.1 Toxic Environmental Conditions

As discussed previously, microorganisms have adapted to life in many naturally occurring extreme environments; however, most of these organisms are autotrophs that will not significantly biodegrade organic pollutants. Extreme toxic conditions resulting from human activities and chemical releases to the environment are more difficult to ameliorate. The addition of oxygen and nutrients, as well as a moderate ability to influence pH, comprise the options currently available to influence environmental conditions. High concentrations of pollutants may cause toxic conditions and prevent biodegradation. Under these circumstances removal of source material might lower groundwater concentrations to non-toxic levels at which biodegradation could occur. Moderate environmental conditions at near neutral pH, with adequate moisture, nutrients, and electron acceptors are the most likely to promote biodegradation.

31.1.4.2 Bioavailability and Mass Transfer Limitations

Most bacteria take up dissolved nutrients and substrate from the surrounding water. Phenomena that lower the dissolved concentrations or dissolution rate of compounds will limit biodegradation. Sorption of hydrophobic compounds to soils results in significant mass fractions of these compounds being unavailable to microorganisms. Biodegradation of the soluble phase results in desorption to reestablish the phase distribution equilibrium. In this situation, desorption rates may control biodegradation kinetics (Bosma et al., 1997). Long-term exposure of hydrophobic compounds to soils often results in a fraction of compound that remains unavailable for biodegradation due to mechanisms that are not yet fully understood (Hatzinger and Alexander, 1995). Therefore, it should be clear that *in situ* biodegradation kinetics may reflect processes other than microbial metabolism such as desorption of the pollutants, pollutant transport, and availability of electron acceptors.

31.1.4.3 Absence of Organisms

In some cases, novel pollutant degrading microorganisms have been isolated and cultured in laboratories, while the widespread existence of these microorganisms in the environment has not been observed. The introduction of novel non-indigenous pollutant degrading microorganisms to resolve this situation has several potential problems. The survival and effectiveness of non-indigenous microorganisms *in situ* has rarely been carefully documented. Current research is attempting to assess the transport and survival of introduced organisms in contaminated zones. Recent success with the bioaugmentation of halo respiring microorganisms into chlorinated solvent contaminated sites has firmly established this method as a viable remediation technique. An outstanding overall review of bioaugmentation applications and principles is now available (Gentry et al., 2004). However, the well-known ecological problems caused by the introduction of invasive non-indigenous macroscopic organisms (e.g., zebra mussels, invasive weeds, rabbits) should provoke careful evaluation of the practice of introducing non-native organisms, even at the microscopic scale.

31.1.4.4 Co-occurring Contaminants

Mixtures of different chemicals may influence biodegradation in several ways. Microorganisms have biochemical mechanisms for optimizing energy production by specifically utilizing preferred substrates while repressing catabolism of other substrates. Thus, compounds that are readily biodegraded when present individually, may persist when present in a mixture. Studies of BTEX biodegradation have revealed that substrate interactions are important, and are likely complex and diverse (Alvarez and Vogel, 1991). Compounds that require different redox conditions for biotransformation may further complicate the biodegradation of mixtures. For example, aerobic conditions are desirable for petroleum hydrocarbon biodegradation, but anaerobic conditions are necessary for reductive dechlorination of chlorinated solvents. Thus it may not be possible to have optimum biodegradation conditions for all compounds present in a mixture. Furthermore, readily biodegradable constituents of mixtures, such as ethanol in gasoline, may exert an oxygen demand sufficient to cause anaerobic conditions that result in a spatial expansion of BTEX plumes in groundwater (Ruiz-Aguilar et al., 2002).

31.2 Quantitative Description of Reactive Transport and Biodegradation

A quantitative understanding of coupled chemical transformation, sorption, and transport processes is required to predict, control, and possibly optimize biodegradation in natural and engineered environments. In this section, the basic mass conservation statements and idealized constitutive models that provide a framework for reactive transport modeling are presented. First, basic microbial reaction kinetics are discussed followed by a discussion of the most basic ways in which sorption can be idealized in reactive transport systems, and finally, the interactions that can be expected among these “reactive” mechanisms and the transport mechanisms are presented.

31.2.1 Microbial Reaction Kinetics

Probably the most familiar expression for microbial growth kinetics is attributed to Jacques Monod, after whom the general expression is named, although this equation by no means was Monod's greatest contribution to science. Later, he, along with F. Jacob and A. Lwoff conducted experiments on lactose metabolism by *E. coli* that led to the discovery of the lac-operon and the overall science of regulatory genes for which they were awarded the Nobel Prize in Medicine in 1965. A major milestone in Monod's contributions to the area of microbial growth kinetics came in 1942 with the publication of the first of his two books, *Recherches Sur la Croissance des Cultures Bactériennes*. Surprisingly and unfortunately, an English translation of this book was never published; consequently the writers of many current English textbooks, not only ignore the development of the equation, but tend to refer to many different equations as the Monod equation, with these alternate equations being variations on the same theme. An excellent discussion of Monod's equation and its development is provided by Panikov (1995) and is summarized here.

For *E. coli* and *B. subtilis* cultures, Monod observed that after complete substrate utilization, the biomass yield per unit mass of substrate, Y , was nearly constant, independent of culture age or initial substrate concentration,

$$[X]_m = [X]_0 + [S]_0 \cdot Y \quad (31.1)$$

where $[X]_m$ is the final biomass concentration after complete consumption (ML^{-3}), $[X]_0$ is the initial biomass concentration (ML^{-3}), $[S]_0$ is the initial substrate concentration (ML^{-3}), and Y is the yield coefficient for a specific substrate by a specific organism (MM^{-1}). For batch short-term-duration growing cultures, this relationship provided a material balance between substrate and biomass concentration that was employed to calculate the substrate concentration at any point along the growth curve,

$$[S] = [S]_0 - \frac{([X] - [X]_0)}{Y} \quad (31.2)$$

where X is the biomass concentration (ML^{-3}).

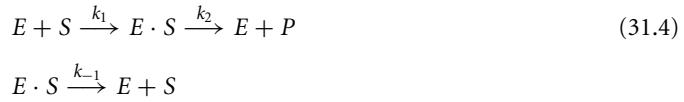
Plotting the calculated values of S against the instantaneous growth rate ($\mu = d(\ln(X)/dt)$, $\text{ML}^{-3}\text{T}^{-1}$) for various points along the growth curve resulted in a relationship that was best fit with an empirical equation with two fitting parameters, μ_m and K_s , that is generally considered in the strictest sense the Monod equation,

$$\mu = \mu_m \frac{[S]}{K_s + [S]} \quad (31.3)$$

where μ_m is the maximum rate ($\text{ML}^{-3}\text{T}^{-1}$) and K_s is the saturation constant (ML^{-3}).

The simple lesson of this equation is that at large substrate concentrations, where $\mu \rightarrow \mu_m$, microbial growth is limited by factors other than substrate concentration, and at small substrate concentrations, utilization is first-order in substrate concentration. Apparently, at the time of publishing his book, Monod was unaware of Michaelis and Menten's much earlier paper (1913) where an identical mathematic equation

was developed to describe the kinetics of substrate utilization by enzymes. The Michaelis–Menten equation is easily derived from the following reaction sequence,



where E is the enzyme, S is the substrate, $E \cdot S$ is the enzyme-substrate complex, and P is the product of the enzymatic reaction. In the derivation, it is assumed that the change in concentration of the enzyme–substrate complex over time is negligible (i.e., at pseudo-steady-state),

$$\frac{d[E \cdot S]}{dt} = k_1[E][S] - k_{-1}[E \cdot S] - k_2[E \cdot S] = 0 \quad (31.5)$$

leading to an approximation of the complex concentration at any point in time equal to,

$$[E \cdot S] = \frac{[S][E]_{\text{tot}}}{[S] + K_s} \quad (31.6)$$

where $[E]_{\text{tot}}$ is the total concentration of enzyme ($= [E] + [E \cdot S]$), (ML^{-3}), and K_s now is defined by the ratio $(k_{-1} + k_2)/k_1$, (M).

By measuring the velocity of product appearance, and realizing the maximum rate of product appearance defined by $v_{\max} = k_2[E]_{\text{tot}}$ (i.e., $[E]_{\text{tot}} = v_{\max}/k_2$), the rate of product appearance is defined by,

$$v = \frac{dP}{dt} = k_2[E \cdot S] = v_{\max} \frac{[S]}{K_s + [S]} \quad (31.7)$$

analogous mathematically to Equation 31.3. The right-hand-side of Equation 31.7 is obtained by substituting v_{\max}/k_2 for $[E]_{\text{tot}}$ in Equation 31.6, and substituting the resulting term for $[E \cdot S]$ into Equation 31.7.

As discussed by Panikov (1995), Monod cleverly invoked material balance constraints in developing Equation 31.3, yet the utility of this equation by itself in describing microbial kinetics is limited. Monod (1942) however proceeded further, defining the rate of microbial growth as a first-order process in biomass concentration, with the rate constant obviously being a function (e.g., $f(x)$) equal to the instantaneous growth rate, $\mu(S)$ (where the substrate concentration, S , is the function variable),

$$\frac{d[X]}{dt} = \mu(S) \cdot [X] = \mu_m \cdot \frac{[S]}{K_s + [S]} \cdot [X] \quad (31.8)$$

After substituting Equation 31.2 into Equation 31.8, separating variables and integrating, the following relationship was obtained,

$$(1 + P) \ln([X]/[X]_0) - P \ln(Q - [X]/[X]_0) + P \ln(Q - 1) = \mu_m \cdot t \quad (31.9)$$

where $P = K_s \cdot Y/(Y \cdot [S]_0 + X_0)$, and $Q = (Y \cdot [S]_0 + X_0)/[X]_0$.

Equation 31.9 has three parameters, μ_m , K_s , and Y , that upon determining their values experimentally for any substrate-organism combination and set of environmental conditions (i.e., pH, temperature, mineral content), can be applied in a predictive manner to estimate the growth dynamics from the initial (boundary value) concentrations of substrate and biomass. Depending upon the magnitude of the three “fitting” parameters, plotting X versus t often results in the characteristic S-shaped growth curve generally associated with the “Monod kinetics.” For example, Figure 31.4 is a plot of Equation 31.9 and a plot of the corresponding substrate concentration for the case where $\mu_m = 1.2 \text{ h}^{-1}$, $Y = 0.3$, and $K_s = 75 \text{ mg/l}$.

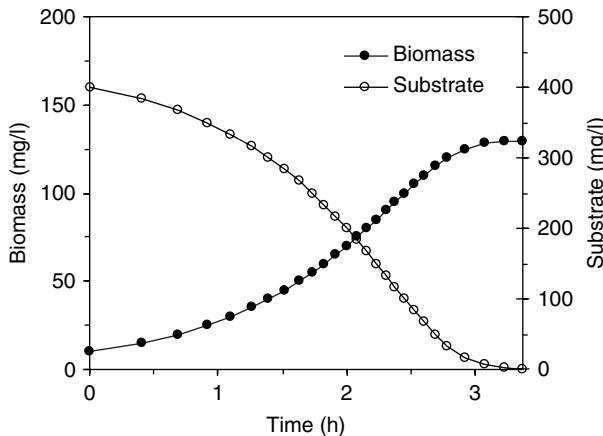


FIGURE 31.4 Substrate utilization curve with initial concentrations of substrate and biomass equal to 400 and 10 mg/l, respectively, and $\mu_m = 1.2 \text{ h}^{-1}$, $Y = 0.3$, and $K_s = 75 \text{ mg/l}$.

The ensuing literature on microbial growth kinetics over the past ~ 7 decades is replete with examples where either Monod's original model has been invoked to interpret experimental data or where modified versions have been used (see e.g., Simkins and Alexander, 1984; Brezonik, 1994). One such form is the generalized kinetic model that considers simultaneous utilization of numerous substrates by numerous microbial species, and includes reaction terms to account for organism death (Schnoor, 1996),

$$\frac{dX_i}{dt} = \sum_{j=1}^m \left(\frac{\mu_{i,j} S_j X_i}{K_{s_{i,j}} + S_j} \right) - b_i X_i \quad (31.10)$$

$$\frac{dS_j}{dt} = \sum_{i=1}^n \left(\frac{-\mu_{i,j} S_j X_i}{K_{s_{i,j}} + S_j} \right) \cdot \frac{1}{Y_{i,j}} \quad (31.11)$$

where all terms are as previously defined, with n denoting the number of chemical substrates, and m denoting the number of microbial species, and b_i is the death rate constant of species i (T^{-1}).

While Equation 31.10 and Equation 31.11 are mathematically straightforward, they are very impractical for the purpose of evaluating chemical transformation in groundwater, or essentially any complex system for that matter. Even for the simple case of one organism utilizing four substrates, a 13 parameter model results. For the case of BTEX (including all three xylene isomers) being metabolized by four organisms a model with 76 coefficients results, and we have yet to consider sorption and transport! Clearly, the model must be simplified for practical purposes, and with reasonable insight into underlying metabolic processes, this can be accomplished for many specific cases.

In many cases, for example, metabolism of the chemical substrate of interest is very slow, as it may be an anthropogenic chemical at low concentration that may not act as the main carbon source. In a non-engineered system where the biomass concentration and supply of electron acceptors (i.e., redox condition) are relatively constant, the transformation kinetics of such a chemical may easily be characterized as pseudo-first order, as simplification of Equation 31.11 for such conditions suggests. Hence, environmental half-lives of many chemicals often vary little from site to site at similar temperatures, with this being particularly true for agricultural pesticides in groundwater.

Alternatively, in many highly contaminated systems and engineered systems, the overall rate of degradation is established (i.e., controlled) by one or two "limiting factors." A large input of reduced carbon, for example, may drive the system to a low redox potential, such that all organisms capable of specific transformations do not continue to grow, effectively stopping transformation of certain

chemicals. High redox potentials have similar effects on the transformation of other chemicals. In such cases, the kinetics of degradation often are directly controlled by the rate of input of a suitable electron acceptor, leading to zero-order kinetics with respect to substrate concentration. Additional factors limiting degradation include, but are not limited to, toxicity to the chemical or to other chemicals present in the system including reaction products; extreme system conditions such as very low or very high pH; and insufficient concentrations of essential elements in absorbable forms, including phosphorus and nitrogen.

31.2.2 Sorption as a Reactive Kinetic Process

Sorption of chemicals to soils, sediments, and aquifer solids has been an area of extensive research over the past several decades. As with biological transformation and growth kinetics, a large portion of this effort has been devoted to the discovery of generalized models that can be used in predictive ways to characterize chemical behavior in natural and contaminated environments, and in engineered environments. Obviously, similar to biological transformations, sorption affects chemical transport and fate, often in a big way. Hence, well-trained quantitative algorithms for such processes become essential components in evaluating the overall risks associated with existing contaminated sites, in evaluating the risk of new chemicals before they are approved for use for various purposes within our society, and in evaluating treatment efficacy and potential costs associated with remediation technologies at existing contaminated sites.

31.2.2.1 Sorption Thermodynamics

While modern site remediation efforts and associated costs generally have been concerned with several hundred chemicals, with a few dozen chemicals marshalling major attention (PCE, MTBE, BTEX components, etc.), the development of quantitative sorption algorithms has been largely through successful elucidation of coefficients, consistent with physical chemistry theory that can be easily measured or estimated for thousands of chemicals. As the first studies examined non-polar organic chemicals, including PAHs, PCBs, pesticides, and chlorinated solvents, the dominant influence of soil organic matter as the sorptive phase was noted with the linearity of sorption isotherms indicating that partitioning (i.e., solubilization) was occurring, rather than adsorption to a surface, leading to the common expressions for phase distribution,

$$K_p = \frac{[C]_s}{[C]_{aq}} = K_{oc} \cdot f_{oc} = \frac{f_{oc} \cdot [C]_{oc}}{[C]_{aq}} \quad (31.12)$$

where $[C]_s$ is the chemical concentration in the bulk solid (MM^{-1}), $[C]_{aq}$ is the chemical concentration in the aqueous phase (ML^{-3}), $[C]_{oc}$ is the chemical concentration in the solid normalized to organic carbon mass in the solid phase (MM^{-1}), K_p is the partition coefficient (L^3M^{-1}), K_{oc} is the carbon-normalized partition coefficient (L^3M^{-1}), and f_{oc} is the fraction of organic carbon in the solid phase (MM^{-1}).

Correlations of K_{oc} with the octanol–water partition coefficient, K_{ow} , (Karickhoff et al., 1979) and water solubility, S_w (Chiou et al., 1979; Karickhoff et al., 1979) were reported within the same year, with Karickhoff et al. correctly noting that a much better correlation existed between K_{oc} and K_{ow} than with S_w because of the fact that in these early correlations, the crystal energy contributions of those chemicals that are solids at room temperature were not quantitatively taken into consideration. This has led to the concept of adjusting the solubility of room-temperature solids to the hypothetical solubility value that would occur if the chemical were a liquid at room temperature (see e.g., Schwarzenbach et al., 1993), referred to as the hypothetical sub-cooled (also, super-cooled) liquid solubility. The log of this hypothetical solubility is inversely proportional to the energy change in transferring the chemical's molecules from the hypothetical liquid phase to water, as the energy required to “break” the crystal is eliminated from consideration (i.e., not involved). If S^* is defined as the water solubility of ambient-temperature liquids or the hypothetical sub-cooled solubility of ambient temperature solids, correlations between $-\log S^*$ and $\log K_{ow}$ have a sound theoretical basis, and correlations between $-\log S^*$ and $\log K_{oc}$ are generally robust and can be considered semi-empirical.

In the case of partitioning between water and either sediment organic carbon (oc) or octanol, the distributions involve only solvent–solute interactions, and crystal energy need not be considered. Karickhoff et al. (1979) determined the relationship between the associated partition coefficients as,

$$K_{oc} = 0.63 \cdot K_{ow} \quad (31.13)$$

To this day, this relationship remains a reasonable approximation for thousands of non-polar chemicals over numerous soil, sediment, and aquifer media classifications. In a paper two years later, Karickhoff (1981) reported on a slightly different proportionality constant of 0.411 (rather than 0.63) based on the sorption of benzene and four polycyclic aromatic hydrocarbons to soils. This new, trained model was then used to estimate the K_{oc} values of 38 other chemicals for which K_{oc} (i.e., K_p and f_{oc}) values were reported in the literature with agreement between estimated and measured values generally agreeing within a factor of three. While this error may seem large, it should be noted that Karichkoff's data set included chemicals that vary by over six orders in magnitude in K_{ow} ($=1^{1.0}-10^{6.5}$)!

One may ask why this rather simple correlation works so well. For example, soil organic carbon is derived from different sources and exists at different stages along its diagenetic path. The simple answer is because it is the interactions between the solutes with water, rather than with the organic or solid matrix, that influence the magnitude of these phase distribution processes most significantly. Basic physical chemistry demonstrates that by invoking the same standard (i.e., reference) state (i.e., pure liquid or hypothetical subcooled liquid phase) the thermodynamic (i.e., equilibrium) chemical potentials, hence the activities, of a chemical solubilized within different phases are equal, where activity is defined as the product of the chemical concentration and the activity coefficient,

$$\gamma_1 \cdot [C]_1 = \gamma_2 \cdot [C]_2 \quad (31.14)$$

An important system for which this relationship applies is the distribution of a chemical between its pure liquid phase (i.e., liquid benzene) and water, defining the chemical's "water solubility." With the pure liquid as the reference state, it can be shown that the activity of chemical in the pure liquid phase equals 1 ($1 = \gamma \cdot [C]$), in which case the solubility of the chemical in the water phase, S^* equals the reciprocal of the associated activity coefficient ($=1/\gamma$). Equation 31.14 also applies to partitioning of a chemical, C , between octanol and water,

$$K_{ow} = \frac{[C]_o}{[C]_{aq}} = \frac{\gamma_{aq}}{\gamma_o} \quad (31.15)$$

As K_{ow} of nonpolar organic compounds increases through a series of chemicals, the respective water phase activity coefficients increase almost proportionally, as the respective octanol phase activity coefficients remain nearly constant, and actually increase very slightly. This can be shown by applying Equation 31.15 to chemical distributions at the solubility limit (or hypothetical sub-cooled liquid solubility limit). For example, from the K_{oc} and K_{ow} values reported by Karickhoff (1981) for benzene, naphthalene, phenanthrene, anthracene, and pyrene and the corresponding solubilities, S^* , reported by Schwarzenbach et al. (1993), the solubility of each chemical in water-saturated octanol and soil organic carbon can be calculated with Equation 31.15 or Equation 31.12, respectively. Figure 31.5 shows these solubility values (S^*) regressed against K_{ow} values on logarithmic scales. Clearly, the aqueous phase solubility decreases as K_{ow} increases, however solubility in octanol and soil organic carbon also slightly decreases as K_{ow} increases! Hence, partitioning to soil organic matter by these solutes is not because of a greater affinity of the low solubility compounds for this natural organic carbon phase; rather it is because of their much less favorable affinity for water: The molecules truly are *hydro-phobic*. The affinity, or attractive forces, for soil organic carbon actually decreases slightly with increasing K_{ow} ; however, the much more significant decrease in affinity for water is what results in larger partition coefficients. The parallel slopes of the octanol and organic carbon solubility lines in Figure 31.5 indicate that octanol has about twice the affinity

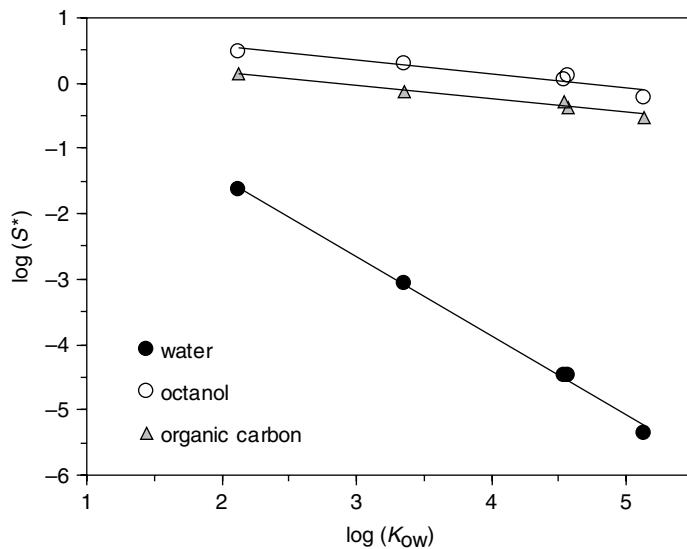


FIGURE 31.5 Solubility or sub-cooled liquid solubility (S^*) of benzene, naphthalene, phenanthrene, anthracene, and pyrene in regressed against.

for these compounds as soil organic carbon, independent of the specific chemical's K_{ow} value. At the macro-scale, this same phenomenon of very weak attractive forces between nonpolar molecules accounts for the low surface tension of organic liquids such as gasoline that spread easily across a surface, as the cohesive forces holding the molecules together are very weak.

The significance of these predictive algorithms for *a priori* estimation of contaminant transport in porous media was quickly recognized, resulting in the ASCE Hydraulics Division Committee on Hydrologic Transport and Dispersion to solicit from S. W. Karickhoff a review paper on the subject in 1984 (Karickhoff, 1984). Importantly, in this paper, situations where Equation 31.12 and Equation 31.13 are poor estimators of chemical distribution are reviewed. These situations include (1) high clay to organic carbon content soils; (2) low (<0.05%) organic carbon soils; and (3) systems where the specific chemical interacts with the solid matrix through polar or ionic (i.e., non-hydrophobic) forces. In the case of surface soils and sediments, there is much recent interest on nonlinear and enhanced sorption caused by char and soot content in some soils (Bucheli and Gustafsson, 2000; Zhu and Pignatello, 2005), although characterization of char and soot in soils is difficult at best. Clearly, for the idealized case, where all partition processes are assumed linear, the partition equation can easily be generalized to consider sorption to numerous sorptive components (i.e., organic carbon, amorphous oxides, and aged clay minerals), each with unique affinities,

$$[C]_{\text{aq}} = \sum_{i=1}^n (K_{p_i} \cdot [C]_{s_i}) \quad (31.16)$$

Whether sorption to phases other than organic carbon ought to be considered generally can be gleaned from the polarity of the chemical of interest and the organic carbon content of the matrix. For some extremely hydrophobic chemicals, partitioning to aqueous phase humic acids or colloidal particles is important and can result in "facilitated" transport through porous media.

31.2.2.2 Time-Dependent Sorption

An additional consideration is the rate at which thermodynamic equilibrium is achieved. Sorption requires diffusion of chemical to the solid matrix, often through the matrix, and often through tortuous pores within the solid matrix. As a result, whether or not local sorption equilibrium can be presumed depends

on the relative characteristic times of other processes to that of the sorption process. In the case of sandy aquifer material, any time-dependency can be approximated quite reasonably with a model that accounts for chemical mass transfer, via diffusion, through a film-layer around the sand grains, with local sorption equilibrium occurring at the surface of the particles, as sorption may occur within a thin organic coating, or for very low organic carbon sands, may occur directly on the silica surface. For cohesive soils and sediments, approach to equilibrium can be modeled either as a kinetic process or a diffusive process, with sorption occurring throughout the soil pods. One of the simplest empirical kinetic models for this type of matrix is a “2-box model,” in which sorption occurs rapidly to a portion of the matrix (i.e., always at equilibrium with the surrounding solution), and sorption is treated as a kinetic process to the remaining portion (Karickhoff, 1980),



where K_p is the equilibrium partition coefficient ($\text{L}^3 \text{ M}^{-1}$), C_{S1} is chemical concentration in the fraction of sites (X) accessed rapidly (M M^{-1}), C_{S2} is the chemical concentration in sites accessed slowly, ($1 - X$), (M M^{-1}), and k is the first-order exchange constant for chemical passing from the equilibrium to kinetically accessed sites (T^{-1}).

The advantages of this model are that only two additional model parameters exist (k and X), and the temporal dependency is treated as a first-order process and can be easily encoded within larger flow models. If approach to equilibrium is treated as a diffusive process, then definition of particle geometry is required, along with definition of the active matrix through which diffusion is occurring (i.e., water pores versus organic matrix). Excellent descriptions of these types of models exist in the literature (Wu and Gschwend, 1986).

31.2.3 Microbial Reactions and Sorption in a Batch System

Before considering flowing systems, much can be gleaned from simple systems where only first-order microbial transformation of a chemical and equilibrium sorption are considered. In most cases, microbial transformation does not occur if the chemical is partitioned to a solid matrix, especially if this matrix is the organic fraction of a soil, in which case the chemical may be considered “dissolved” by the organic matrix and unavailable to the microorganisms. Hence, only chemical dissolved in the water phase is readily accessible. A sufficient set of equations that may define this system include partition, mass balance, and biodegradation kinetic equations, respectively,

$$K_p = \frac{[C]_s}{[C]_{\text{aq}}} = K_{\text{oc}} \cdot f_{\text{oc}} = 0.411 \cdot K_{\text{ow}} \cdot f_{\text{oc}} \quad (31.18)$$

$$V_{\text{tot}}[C]_{\text{tot}} = V_{\text{aq}}[C]_{\text{aq}} + M_s[C]_s \quad (31.19)$$

$$\frac{d[C]_{\text{tot}}}{dt} = -k_{\text{bio}}[C]_{\text{aq}} \quad (31.20)$$

where V_{tot} is total (or unit) volume (L^3), $[C]_{\text{tot}}$ is the total chemical mass divided by the total volume (ML^{-3}), M_s is the mass of the soil (M), C_s is the chemical concentration in the soil (M M^{-1}), and k_{bio} is the pseudo-first order biodegradation rate constant of aqueous phase chemical (T^{-1}).

Substitution of Equation 31.18 and Equation 31.19 into Equation 31.20 yields Equation 31.21 that is easily solved by separation of variables,

$$\frac{d[C]_{\text{tot}}}{dt} = -k_{\text{eff}}[C]_{\text{tot}} \quad (31.21)$$

where

$$k_{\text{eff}} = k_{\text{bio}} \cdot \left(\frac{V_{\text{aq}}}{V_{\text{tot}}} + \frac{0.411 \cdot K_{\text{ow}} \cdot f_{\text{oc}} \cdot M_s}{V_{\text{tot}}} \right)^{-1} \quad (31.22)$$

For any first-order process such as this, the observed or “effective” half-life of any chemical in the system is defined by $t_{1/2,\text{eff}} = 0.693/k_{\text{eff}}$. To evaluate the influence that sorption has on the effective half-life of several chemicals in this system, the pseudo-first order biodegradation rate constant, k_{bio} , can be held constant for the purpose of demonstration (say, $k_{\text{bio}} = 0.173 \text{ d}^{-1}$ for a half-life in the water phase of 4 days) while varying K_{ow} . For this comparison, assume constant soil mass (10 g), organic carbon content (2%) soil density (2 g/ml), and water volume (100 ml). Under these conditions, while the rate of decay in the aqueous phase is set to be the same with a half-life of 4 days, the effective half-lives of total chemical mass in the system for chemicals with hydrophobicities equal to: chlorobenzene ($\log K_{\text{ow}} = 2.84$), 1, 2, 3, 4-tetrachlorobenzene ($\log K_{\text{ow}} = 4.60$), and p,p'-DDT ($\log K_{\text{ow}} = 6.36$) are 6 days, 128 days, and 19.7 years, respectively. Note that these effective half-lives are in general agreement with observed loss rates of these compounds in the environment (i.e., days, months, years) under reducing conditions. Interestingly, some chemicals like p,p'-DDT are known to undergo facile reduction in organic solvents like hexane or acetonitrile as long as a suitable reducing agent is present, again indicating that the environmental recalcitrance of chemicals like this may result more from their limited availability rather than from limited reactivity.

31.2.4 Microbial Reaction, Sorption, and Transport

These same idealized processes of first-order microbial transformation and sorption equilibrium can be considered in a 1-dimensional (1-D) saturated porous media flowing system. Because concentration, and hence total chemical mass, may vary with distance, x , along the 1-D path, sorption equilibrium occurs only locally at each discrete value of x . The equation that describes transport in 1-D for this simplest of cases is,

$$\frac{\partial [C]_{\text{aq}}}{\partial t} = -\mu \cdot \left(\frac{\partial [C]_{\text{aq}}}{\partial x} \right) + D \cdot \left(\frac{\partial^2 [C]_{\text{aq}}}{\partial x^2} \right) \pm \sum (rxn) \quad (31.23)$$

where μ is the water velocity (LT^{-1}), D is the dispersion coefficient (L^2T^{-1}) and the reaction terms (rxn) that are considered include those for pseudo first-order biodegradation and mass transfer of chemical from the aqueous phase to the solid phase, respectively,

$$\frac{d[C]_{\text{aq}}}{dt} = -k_{\text{bio}}[C]_{\text{aq}} \quad (31.24)$$

$$\frac{d[C]_{\text{aq}}}{dt} = \frac{-(1-n)\rho}{n} \left(\frac{d[C]_s}{dt} \right) \quad (31.25)$$

where n is the bulk porosity of the media (water volume/total volume), ρ is the density of the solid phase (ML^{-3}), and again $[C]_s$ is the concentration of chemical in the solid phase, MM^{-1} .

Equation 31.25 is derived strictly via mass balance, as any chemical that leaves the aqueous phase due to sorption must accumulate in the solid phase and vice versa. Because, locally $[C]_{\text{aq}}$ and $[C]_s$ are in equilibrium, Equation 31.12 can be substituted into Equation 31.25 eliminating $[C]_s$ as a parameter, and the resulting two reaction equations can be substituted into Equation 31.23, which upon rearrangement yields,

$$\frac{\partial [C]_{\text{aq}}}{\partial t} = \frac{-\mu \cdot (\partial [C]_{\text{aq}} / \partial x) + D \cdot (\partial^2 [C]_{\text{aq}} / \partial x^2) - k_{\text{bio}}[C]_{\text{aq}}}{t_r} \quad (31.26)$$

where

$$t_r = \left(1 + \frac{K_p(1-n)\rho}{n} \right)$$

The sorption term, t_r is commonly referred to as the retardation factor and may be thought of as the ratio of the water flow rate to the chemical transport rate. A value of 5, for example, implies that the chemical on average is moving through the porous media at one-fifth the rate of the water. A large retardation factor therefore attenuates the chemicals advective transport (the first term on the right-hand side of Equation 31.26); however, because t_r is the denominator of all three terms on the right-hand-side, it also attenuates dispersive transport, and as before, biological degradation. Many analytical solutions exist for this equation as the solution depends upon the boundary conditions of interest. Very often, numerical solutions are used to solve for $[Cl]_{aq}$ as a function of time and distance, as additional factors are considered that prohibit finding a suitable analytical solution, such as media heterogeneity with distance (variation in K_p , n , etc.), variable flow rates, etc. Additional variations to this simplest construct include accounting for (1) advective and dispersive flow in three dimensions, (2) sorption nonlinearity, (3) sorption dynamics, (4) the simultaneous transport of substrate and electron acceptors (i.e., higher order biodegradation kinetics), (5) biomass growth, and the list goes on. Very importantly, even for this simplest of cases, the solution to Equation 31.26 is only half the answer. The other half is in calculating $[C]_s$ with time and distance; however, because local equilibrium is assumed in the derivation, this becomes a trivial task. For any chemical that has a retardation coefficient greater than two, however, more than half the total chemical mass will be found sorbed on the media.

31.3 Field Applications of Bioremediation

Bioremediation is the utilization of naturally occurring microbial biodegradation processes to restore a site to a non-hazardous condition. A detailed discussion of bioremediation application and design is presented elsewhere (King et al., 1992; Riser-Roberts, 1992; Baker and Herson, 1994; Flathman et al., 1994; Cookson, 1995). There are many advantages to using bioremediation. As pollutant destruction occurs *in situ*, the potential liability and environmental risk associated with the removal, handling, transport, and storage of hazardous contaminated materials is eliminated. An additional important advantage of bioremediation is that it is often the most economical solution available. Successful bioremediation requires a thorough site investigation and evaluation of treatment options. A site history is compiled which should include property uses, chemicals stored, and location of utilities and buildings, as well as the location of nearby wells. Available information about local hydrogeology, geology, and topography should also be included. Soil and groundwater samples are obtained to identify the contaminants present, and estimate their concentration and distribution. Soil borings are made to determine the local hydraulic gradient and sometimes in conjunction with pumping tests, the hydraulic conductivity is estimated. Whenever possible, an assessment of the local groundwater geochemistry should be made by measuring alternative electron acceptor concentrations (nitrate, sulfate), pH and aquifer buffering capacity. Data from samples taken outside the zone of contamination are useful for assessing background microbiological and geochemical conditions. Laboratory studies are currently the most reliable method for assessing the biodegradation potential of the indigenous organisms. From the assembled information, predictions of contaminant migration and biodegradation can be used to evaluate the current hazard posed and potential remediation options.

31.3.1 Intrinsic Bioremediation

Intrinsic bioremediation is the natural *in situ* biodegradation of pollutants without the engineered manipulation of environmental conditions. Intrinsic bioremediation is appropriate when a site investigation shows that natural biodegradation processes are sufficiently attenuating the migration of contaminants and there is little imminent risk or liability associated with the site. An ongoing monitoring program is essential to regularly evaluate the progress of intrinsic bioremediation until site closure.

31.3.2 Engineered Bioremediation

Engineered systems are installed when it is necessary to overcome some limitation to biodegradation, or health and liability concerns make it desirable to accelerate naturally occurring processes. In current practice, the addition of some form of oxygen to the aquifer is usually the primary objective. Nutrients (N&P) are often added as well, usually without site-specific evidence of a nutrient deficiency, but because it is an easy and inexpensive option. Pumping wells, injection wells, infiltration galleries, vacuum pumps, and compressors may be installed in various combinations and configurations depending on the design objective. In addition to adding an electron acceptor or nutrients to an aquifer, it may be desirable to attempt to contain the contaminant plume, install a groundwater recirculation system, or combine bioremediation with other remediation processes.

31.3.3 Case Histories

The objective of listing a few case histories is to provide the reader with references to further study and well-documented examples of applied bioremediation. Therefore, only brief summaries of several types of bioremediation field studies are presented and the interested reader is encouraged to obtain the cited references for detailed descriptions.

31.3.3.1 Aerobic Petroleum Hydrocarbon Bioremediation

The aerobic bioremediation of motor oil, diesel fuel, and gasoline that had leaked from a used oil sump into groundwater was described by Nelson et al. (1994). Hydrogen peroxide and nutrients were infiltrated into the contaminated zone through an installed groundwater recirculation system. A vapor extraction system was combined with bioremediation to enhance hydrocarbon removal in the unsaturated zone.

31.3.3.2 BTEX Bioremediation under Denitrifying Conditions

The use of nitrate as an electron acceptor for the bioremediation of JP-4 jet fuel was described by Hutchins et al. (1991). The system consisted of an infiltration gallery, interdiction wells, pumping wells, and monitoring wells. The bioremediation was effective in removing BTEX constituents; however, a significant amount of relatively insoluble alkanes remained as NAPL.

31.3.3.3 Aerobic Chlorinated Aliphatic Hydrocarbon Bioremediation

The feasibility of *in situ* methanotrophic co-oxidation of chlorinated ethenes was tested at Moffett Naval Air Station, CA (Roberts et al., 1990; Semprini et al., 1990). With sufficient oxygen and methane the methanotrophic bacteria cometabolized TCE, DCE, and VC in a 2-m long biostimulated zone. Approximately 90% of the VC and 80 to 90% of the trans-DCE was biodegraded in this short distance. TCE and cis-DCE were more recalcitrant; however, longer residence times in biostimulated zones would improve the extent of removal.

31.3.3.4 Anaerobic Chlorinated Aliphatic Hydrocarbon Bioremediation

Several excellent case studies involving bioaugmentation of chlorinated solvent contaminated sites with halorespiring bacteria are available (Ellis et al., 2000; Major et al., 2002). In addition, an extensive compilation of the specific techniques used at different sites and the cost of application and other information is available as well (Parsons Engineering Science, Inc., 2002).

31.3.3.5 Intrinsic Bioremediation

Intrinsic biodegradation processes have been responsible for the containment of petroleum hydrocarbons from a pipeline leak in Minnesota (Baecker et al., 1993). Careful monitoring has confirmed that the soluble BTEX constituents are not migrating with the groundwater flow. This is due to the biodegradation rate being equal to the release rate of BTEX from the residual hydrocarbon mixture.

31.3.4 Emerging Technologies

Increased utilization of bioremediation will depend on technological advances that will remove limitations to when bioremediation can be reliably and effectively applied (Atlas, 1995). Advances in understanding novel bioremediation technologies have been made in several areas, for example, bioavailability, molecular biology, microbiology, and anaerobic biodegradation.

Microbially produced biosurfactants and synthetic surfactants increase the apparent solubility of hydrophobic compounds and could potentially reduce bioavailability limitations. However, much remains to be learned about the biodegradation of surfactant solubilized compounds and potential surfactant toxicity, as well as cost and effectiveness. A combined treatment process consisting of surfactant/solvent soil washing followed by bioremediation to remove residual contamination has potential for the remediation of NAPL (Non-Aqueous Phase Liquids) contaminated sites. However, the microecological effect of the soil washing is unknown.

Advances in molecular biology have contributed much to our understanding of the biochemistry and microbiology of biodegradation, and will continue to do so. Use of molecular techniques is now commonplace in biodegradation research. The importance of understanding the biochemistry, genetics, and microbiology of biodegradation cannot be overestimated. In the coming years, molecular techniques for detecting and enumerating bacteria *in situ* will likely become routine. These techniques have been used to characterize anaerobic chlorinated solvent sites (Hendrikson et al., 2002) and aerobic petroleum contaminated sites (Baldwin et al., 2003). Genetic techniques are being used to construct novel biodegradation pathways, essentially creating organisms with new capabilities.

Additional examples of emerging technologies include water recirculation systems, oxygen releasing peroxides, the use of zero valent metals to enhance reductive dechlorination, as well as advances in *in situ* measurement of physical, chemical, and biological parameters. Ongoing development of pollutant degrading thermophiles, alkaline tolerant bacteria, and bacteria with membranes resistant to high solvent concentrations will ensure that biodegradation will continue to grow as a primary groundwater remediation option.

Glossary

Aerobic Microorganisms that use molecular oxygen as a terminal electron acceptor or an environment that contains molecular oxygen.

Anaerobic Microorganisms that do not use molecular oxygen as a terminal electron acceptor or an environment that contains no molecular oxygen.

Autotroph An organism that uses inorganic carbon as a source for biosynthesis of cellular constituents.

Bacteria In this chapter, bacteria includes all prokaryotic organisms. In fact, Achaea and Bacteria are separate phylogenetic domains of prokaryotic organisms.

Bioavailability Refers to the availability of a substance for uptake and transformation by living organisms.

Bioremediation The use of naturally occurring biodegradation processes for the remediation of sites contaminated with pollutants.

Cell An assembly of highly structured macromolecules which is the smallest living unit. Bacteria are single-celled organisms.

Citric Acid Cycle A nearly universal biochemical cycle that oxidizes organic carbon to CO₂ and NADH.

Cometabolism A gratuitous biochemical reaction from which the microorganism derives no benefit.

Electron acceptor The terminal oxidant in respiratory processes. Oxygen is the electron acceptor of aerobic organisms.

Electron donor The compound that is oxidized as a source of energy by organisms.

Eukaryote Organisms that have a membrane enclosed compartment containing their DNA (a nucleus).

Fermentation Substrate transformation without the use of an external electron acceptor, therefore the substrate serves as both electron donor and electron acceptor.

Halogens Elements in group VII of the Periodic Table. F, Cl, Br, I, At are halogens.

Heterotroph An organism that uses organic carbon as a source for biosynthesis of cellular constituents.

Lithotroph An organism that oxidizes inorganic compounds for energy.

Organotroph An organism that oxidizes organic carbon for energy.

pH A logarithmic unit of H⁺ ion concentration, pH = $-\log[H^+]$.

Prokaryote Single-celled organisms without a membrane enclosed compartment containing DNA.

Redox potential The electrical potential of a given oxidation-reduction reaction.

Reductive Dechlorination Removal of halogens as halide ions by the reduction of the organic compound with the replacement of the halogen by a hydrogen atom.

Respiration Substrate transformation in which an external oxidant is used as a terminal electron acceptor.

References

- Abramowicz, D.A. 1990. Aerobic and anaerobic biodegradation of PCBs: A review. *CRC Crit. Rev. Biotechnol.*, 10:241–251.
- Alexander, M. 1994. *Biodegradation and Bioremediation*, Academic Press, San Diego, CA.
- Alvarez, P.J.J. and Vogel, T.M. 1991. Substrate interactions of benzene, toluene, and para-xylene during microbial degradation by pure cultures and mixed culture aquifer slurries. *Appl. Environ. Microbiol.*, 57:2981–2985.
- Alvarez-Cohen, L. and McCarty, P.L. 1991. Product toxicity and cometabolic competitive inhibition modeling of chloroform and trichloroethylene transformation by methanotrophic resting cells. *Appl. Environ. Microbiol.*, 57:1031–1037.
- Alvarez, P.J.J., Anid, P.J., and Vogel, T.M. 1991. Kinetics of aerobic biodegradation of benzene and toluene in sandy aquifer material, *Biodegradation*, 2:43–51.
- Atlas, R.M. 1995. Bioremediation. *Chem. Eng. News* 73:32–42.
- Bachofer, R., Lingens, F., and Schafer, W. 1975. Conversion of aniline into pyrocatechol by a Nocardia sp.: Incorporation of oxygen-18. *FEBS Lett.*, 50:288–290.
- Baedecker, M.J., Cozzarelli, I.M., Eganhouse, R.P., Siegel, D.I., and Bennett, P.C. 1993. Crude oil in a shallow sand gravel aquifer III. Biogeochemical reactions and mass balance modeling in anoxic groundwater. *Appl. Geochim.* 8:569–586.
- Baker, K.H. and Herson, D.S. 1994. *Bioremediation*, McGraw-Hill, Inc., New York.
- Baldwin, B.R., Nakatsu, C.H., and Nies, L. 2003. Detection and enumeration of aromatic oxygenase genes by multiplex and real-time PCR. *Appl. Environ. Microbiol.*, 69:3350–3358.
- Bedard, D.L. and Haberl, M.L. 1990. Influence of chlorine substitution pattern on the degradation of polychlorinated biphenyls by eight bacterial strains. *Microb. Ecol.*, 20:87–102.
- Bender, K.S., Shang, C., and Chakraborty, R. 2005. Identification, characterization, and classification of genes encoding perchlorate reductase. *J. Bacteriol.*, 187:5090–5096.
- Bitton, G., Farrah, S.R., Ruskin, R.H., Butner, J., and Chou, Y.J. 1983. Survival of pathogenic and indicator organisms in groundwater. *Ground Water*, 21:405–410.
- Bosma, T.N.P., Middeldorp, P.J.M., Schraa, G., and Zehnder, A.J.B. 1997. Mass transfer limitation of biotransformation: quantifying bioavailability. *Environ. Sci. Technol.*, 31:248–252.
- Brezonik, R.L. 1994. *Chemical Kinetics and Process Dynamics in Aqueous Systems*, Lewis Publishers, Boca Raton, FL.
- Brock, T.D., Madigan, M.T., Martinko, J.M., and Parker, J. 1997. *Biology of Microorganisms*, Prentice Hall, Upper Saddle River, NJ.

- Bucheli, T.D. and Gustafsson, O. 2000. Quantification of the soot-water distribution coefficient of PAHs provides mechanistic basis for enhanced sorption observations. *Environ. Sci. Technol.*, 34:5144–5151.
- Chapelle, F.H. 1993. *Groundwater Microbiology & Geochemistry*, John Wiley & Sons, Inc., New York.
- Chaudhry, G.R. and Chapalamadugu, S. 1991. Biodegradation of halogenated organic compounds. *Microbiol. Rev.*, 55:59–79.
- Chen, C.I. and Taylor, R.T. 1995. Thermophilic biodegradation of BTEX by two *Thermus* species. *Biotech. Bioeng.*, 48:614–624.
- Chiou, C.T., Peters, L.J., and Freed, V.H. 1979. A physical concept of soil–water equilibria for nonionic organic compounds. *Science*, 206:831–832.
- Chiou, C.T., Schmedding, D.W., and Manes, M. 1982. Partitioning of organic compounds in octanol–water systems. *Environ. Sci. Technol.*, 16:4–10.
- Cookson, J.T. Jr. 1995. *Bioremediation Engineering: Design and Application*, McGraw-Hill, New York.
- Corapcioglu, M.Y. and Haridas, A. 1984. Transport and fate of microorganisms in porous media: a theoretical investigation. *J. Hydrol.*, 72:149–169.
- Corapcioglu, M.Y. and Haridas, A. 1985. Microbial transport in soils and groundwater: A numerical model. *Adv. Water Resour.*, 8:188–200.
- Dawson, M.P., Humphrey, B.A., and Marshall, K.C. 1981. Adhesion: A tactic in the survival strategy of a marine vibrio during starvation. *Curr. Microbiol.*, 6:195–199.
- Dolfing, J. and Harrison, B.K. 1992. Gibbs free energy of formation of halogenated aromatic compounds and their potential role as electron acceptors in anaerobic environments. *Environ. Sci. Technol.*, 26:2213–2218.
- Dolfing, J., Zeyer, J., Binder-Eicher, P., and Schwarzenbach, R.P. 1990. Isolation and characterization of a bacterium that mineralizes toluene in the absence of molecular oxygen. *Arch. Microbiol.*, 154:336–341.
- Edwards, E.A., Wills, L.E., Reinhard, M., and Grbic-Galic, D. 1992. Anaerobic degradation of toluene and xylene by aquifer microorganisms under sulfate-reducing conditions. *Appl. Environ. Microbiol.*, 58:794–800.
- Ellis, D.E., Lutz, E.J., Odom, J.M., Buchanan, R.J.Jr., Bartlett, C.L., Lee, M.D., Harkness, M.R., and Deweerd, K.A.. 2000. Bioaugmentation for accelerated in situ anaerobic bioremediation. *Environ. Sci. Technol.*, 34:2254–2260.
- Evans, W.C. 1988. Anaerobic degradation of aromatic compounds. *Annu. Rev. Microbiol.*, 42:289–317.
- Evans, P.J., Mang, D.T., Kim, K.S., and Young, L.Y. 1991. Anaerobic degradation of toluene by a denitrifying bacterium. *Appl. Environ. Microbiol.*, 57:1139–1145.
- Fetzner, S. and Lingens, F. 1994. Bacterial dehalogenases: biochemistry, genetics, and biotechnical applications. *Microbiol. Rev.*, 58:641–685.
- Flathman, P.E., Jerger, D.E., and Exner, J.H. 1994. *Bioremediation — Field Experience*, Lewis Publishers, Boca Raton, FL.
- Fletcher, M. and Marshall, K.C. 1982. Are solid surfaces of ecological significance to aquatic bacteria? *Adv. Microbial. Ecol.*, 6:199–236.
- Fredrickson, J.K. and Onstott, T.C. 1996. Microbes deep inside the earth. *Sci. Am.*, 275:68–73.
- Furukawa, K., Tonomura, K., and Kamibayashi, A. 1978. Effect of chlorine substitution on the biodegradability of polychlorinated biphenyls. *Appl. Environ. Microbiol.*, 35:223–227.
- Gentry, T.J., Rensing, C., and Pepper, I.L. 2004. New approaches for bioaugmentation as a remediation technology. *Crit. Rev. Environ. Sci. Technol.*, 34:447–494.
- Gibson, D.T. 1984. *Microbial Degradation of Organic Compounds*, Marcel Dekker, New York.
- Grbic-Galic, D. and T.M. Vogel. 1987. Transformation of toluene and benzene by mixed methanogenic cultures. *Appl. Environ. Microbiol.*, 53:254–260.
- Gottschalk, G. 1986. *Bacterial Metabolism*, Springer-Verlag, New York.
- Hagblom, M.M. and Valo, R.J. 1995. Bioremediation of chlorophenol wastes, pp. 389–434. In L.Y. Young and C.E. Cerniglia (eds.), *Microbial Transformation and Degradation of Toxic Organic Chemicals*, Wiley-Liss, Inc., New York.

- Hancock, C.M., Bose, J.B., and Callahan, M. 1998. Crypto and Giardia in US groundwater. *J. Am. Water Works Assoc.*, 90:58–61.
- Harkness, M.R., McDermott, J.B., Abramowicz, D.A., Salvo, J.J., Flanagan, W.P., Stephens, M.L., Mondello, F.J., May, R.J., Lobos, J.H., Carroll, K.M., Brennan, M.J., Bracco, A.A., Fish, K.M., Warner, G.L., Wilson, P.R., Dietrich, D.K., Lin, D.T., Morgan, C.B., and Gately, W.L. 1993. In situ stimulation of aerobic PCB biodegradation in Hudson River sediments. *Science*, 259:503–507.
- Harvey, R.W., George, L.H., Smith, R.L., and LeBlanc, D.R. 1989. Transport of microspheres and indigenous bacteria through a sandy aquifer: Results of natural and forced gradient tracer experiments. *Environ. Sci. Technol.*, 23:51–56.
- Hatzinger, P.B. and Alexander, M. 1995. Effect of aging of chemicals in soil on their biodegradability and extractability. *Environ. Sci. Technol.*, 29:537–545.
- Hendrickson, E.R., Payne, J.A., Young, R.M., Starr, M.G., Perry, M.P., Fahnestock, S., Ellis, D.E., and Ebersole, R.C. 2002. Molecular analysis of *Dehalococcoides* 16S rRNA DNA from chloroethene-contaminated sites throughout North America and Europe. *Appl. Environ. Microbiol.*, 68:485–495.
- Holmes, D.A., Harrison, B.K., and Dolfing, J. 1993. Estimation of Gibbs free energies of formation for polychlorinated biphenyls. *Environ. Sci. Technol.*, 27:725–731.
- Howard, P.H. 1989. *Handbook of Environmental Fate and Exposure Data for Organic Chemicals*, Lewis Publishers, Inc., Chelsea, MI.
- Huesemann, M.H. 1995. Predictive model for estimating the extent of petroleum hydrocarbon biodegradation in contaminated soils. *Environ. Sci. Technol.*, 29:7–18.
- Hutchins, S.R. 1991a. Biodegradation of monoaromatic hydrocarbons by aquifer microorganisms using oxygen, nitrate, or nitrous oxide as the terminal electron acceptor. *Appl. Environ. Microbiol.*, 57:2403–2407.
- Hutchins, S.R. 1991b. Optimizing BTEX biodegradation under denitrifying conditions. *Environ. Toxicol. Chem.*, 10:1437–1448.
- Hutchins, S.R., Downs, W.C., Wilson, J.T., Smith, G.B., Kovacs, D.A., Fine, D.D., Douglas, R.H., and Hendrix, D.J. 1991. Effect of nitrate addition on bioremediation of fuel-contaminated aquifer: field demonstration. *Ground Water* 29:571–580.
- Karickhoff, S.W. 1980. Sorption kinetics of hydrophobic pollutants in natural sediments. In Robert A. Baker (ed.), *Contaminants and Sediments*, Vol. 2, Ann Arbor Science, Ann Arbor, MI, pp. 193–205.
- Karickhoff, S.W. 1981. Semi-empirical estimation of sorption of hydrophobic pollutants on natural sediments and soils. *Chemosphere*, 10:833–846.
- Karickhoff, S.W. 1984. Organic pollutant sorption in aqueous systems. *J. Hydraul. Eng.*, 110:707–733.
- Karickhoff, S.W., Brown, D.S., and Scott, T.A. 1979. Sorption of hydrophobic pollutants on natural sediments. *Water Res.*, 13:241–248.
- Kellems, B.L., Leeson, A., and Hinchee, R.E. 1994. Review of bioremediation experience in Alaska, p. 438–443. In R.E. Hinchee, B.C. Alleman, R.E. Hoeppl, and R.N. Miller (ed.), *Hydrocarbon Bioremediation*, Lewis Publisher, Boca Raton, FL.
- King, R.B., Long, G.M., and Sheldon, J.K. 1992. *Practical Environmental Bioremediation*, Lewis Publishers, Boca Raton, FL.
- Leahy, J.G. and Colwell, R.R. 1990. Microbial degradation of hydrocarbons in the environment. *Microbiol. Rev.*, 54:305–315.
- LeChevallier, M.W. and Norton, W.D. 1995. Giardia and Cryptosporidium in raw and finished water. *J. Am. Water Works Assoc.*, 87:54–68.
- Lee, M.D., Thomas, J.M., Borden, R.C., Bedient, P.B., Ward, C.H., and Wilson, J.T. 1988. Bioremediation of aquifers contaminated with organic compounds. *CRC Crit. Rev. Environ. Control*, 18: 29–88.
- Löffler, F.E., Sun, Q., Li, J., and Tiedje, J.M. 2000. 16S rRNA gene-based detection of tetrachloroethene-dechlorinating *Desulfuromonas* and *Dehalococcoides* species. *Appl. Environ. Microbiol.*, 66: 1369–1374.

- Logan, B.E. 1998. A review of chlorate- and perchlorate-respiring microorganisms. *Bioremediation J.* 2:69–79.
- Logan, B.E. 2001. Assessing the outlook for perchlorate remediation. *Environ. Sci. Technol.*, 482A–487A.
- Lovley, D.R. and Lonergan, D.J. 1990. Anaerobic oxidation of toluene, phenol and p-cresol by the dissimilatory iron-reducing organism, GS-15. *Appl. Environ. Microbiol.*, 56:1858–1864.
- Lovley, D.R., Baedecker, M.J., Lonergan, D.J., Cozzarelli, I.M., Phillips, E.J.P., and Siegel, D.I. 1989. Oxidation of aromatic contaminants coupled to microbial iron reduction. *Nature*, 339:297–299.
- Major, D.W., McMaster, M.L., Cox, E.E., Edwards, E.A., Dworatzek, S.M., Hendrickson, E.R., Starr, M.G., Payne, J.A., and Buonamica, L.W. 2002. Field demonstration of successful bioaugmentation to achieve dechlorination of tetrachloroethene to ethene. *Environ. Sci. Technol.*, 36:5106–5116.
- McAllister, K.A., Lee, H., and Trevors, J.T. 1996. Microbial degradation of pentachlorophenol. *Biodegradation*, 7:1–40.
- McCarty, P.L. and Semprini, L. 1993. Engineering and hydrogeological problems associated with in-situ treatment. *Hydrol. Sci. J.*, 38:261–272.
- Michaelis, L. and Menten, M.L. 1913. The kinetics of invertin action, *Biochemische Zeitschrift*, 49:333.
- Mohn, W.W. and Tiedje, J.M. 1992. Microbial reductive dehalogenation. *Microbiol. Rev.*, 56:482–507.
- Monod, J.K. 1942. *Recherches Sur la Croissance des Cultures Bactériennes*, Herman et Cie, Paris.
- National Research Council. 1993. *In Situ Bioremediation. When does it work?* National Academy Press, Washington, D.C.
- Neilson, A.H. 1990. A review—the biodegradation of halogenated organic compounds. *J. Appl. Bacteriol.*, 69:445–470.
- Nelson, M.J.K., Montgomery, S.O., Mahaffey, W.R., and Pritchard, P.H. 1987. Biodegradation of trichloroethylene and involvement of an aromatic biodegradative pathway. *Appl. Environ. Microbiol.*, 53:949–954.
- Nelson, C.H., Hicks, R.J., and Andrews, S.D. 1994. In situ bioremediation: an integrated systems approach, pp.125–132. In R.E. Hinchee, B.C. Alleman, R.E. Hoeppel, and R.N. Miller (ed.), *Hydrocarbon Bioremediation*, Lewis Publishers, Boca Raton, FL.
- Panikov, N.S. 1995. *Microbial Growth Kinetics*, Chapman and Hall, London.
- Pardieck, D.L., Bouwer, E.J., and Stone, A.T. 1992. Hydrogen peroxide use to increase oxidant capacity for in situ bioremediation of contaminated soils and aquifers: A review. *J. Contam. Hydrol.*, 9:221–242.
- Parsons Engineering Science, Inc. 2002. Evaluation of performance and costs associated with anaerobic dechlorination techniques. ESTCP. [http://www.estcp.org/documents/techdocs/Phase%20I%20Site%20Survey-Final%20Version%20\(3\).pdf](http://www.estcp.org/documents/techdocs/Phase%20I%20Site%20Survey-Final%20Version%20(3).pdf)
- Reineke, W. and Knackmuss, H.J. 1988. Microbial degradation of haloaromatics. *Annu. Rev. Microbiol.*, 42:263–287.
- Riser-Roberts, E. 1992. *Bioremediation of Petroleum Contaminated Sites*, C.K. Smoley, Boca Raton, FL.
- Roberts, P.V., Hopkins, G.D., Mackay, D.M., and Semprini, L. 1990. A field evaluation of in-situ biodegradation of chlorinated ethenes: part I, methodology and field site characterization. *Ground Water*, 28:591–604.
- Ruiz-Aguilar, G.M., O'Reilley, K., and Alvarez, P.J.J. 2002. A comparison of benzene and toluene plumes for sites contaminated with regular versus ethanol-amended gasoline. *Ground Water Monit. R.*, 23:48–53.
- Schink, B. 1988. Principles and limits of anaerobic degradation: Environmental implications and technological aspects, pp. 771–846. In A.J.B. Zehnder (ed.), *Biology of Anaerobic Microorganisms*, John Wiley & Sons, New York.
- Schnoor, J.L. 1996. *Environmental Modeling, Fate and Transport of Pollutants in Water, Air, and Soil*, John Wiley and Sons, Inc., New York.
- Schwarzenbach, R.P., Gschwend, P.M., and Imboden, D.M. 1993. *Environmental Organic Chemistry*, John Wiley & Sons, Inc., New York.
- Scow, K.M. 1990. Rate of Biodegradation, In W.J. Lyman, W.F. Reehl, and D.H. Rosenblatt (ed.), *Handbook of Chemical Property Estimation Methods*, American Chemical Society, Washington, D.C.

- Semprini, L., Roberts, P.V., Hopkins, G.D., and McCarty, P.L. 1990. A field evaluation of in-situ biodegradation of chlorinated ethenes: Part 2. Results of biostimulation and biotransformation experiments. *Ground Water*, 28:715–738.
- Shannon, M.J.R. and Unterman, R. 1993. Evaluating bioremediation: distinguishing fact from fiction. *Annu. Rev. Microbiol.*, 47:715–738.
- Simkins, S. and Alexander, M. 1984. Models for mineralization kinetics with the variables of substrate concentration and population density. *Appl. Environ. Microbiol.*, 47:1299–1306.
- Smith, M.R. 1990. The biodegradation of aromatic hydrocarbons by bacteria. *Biodegradation* 1:191–206.
- Stumm, W. and Morgan, J.J. 1981. *Aquatic Chemistry*, John Wiley & Sons, Inc., New York.
- Thauer, R.K., Jungermann, K., and Decker, K. 1977. Energy conservation in chemotrophic anaerobic bacteria. *Bacteriol. Rev.*, 41:100–180.
- Thomas, J.M. and Ward, C.H. 1989. In situ bioremediation of organic contaminants in the subsurface. *Environ. Sci. Technol.*, 23:760–766.
- Tiedje, J.M. and Alexander, M. 1969. Enzymatic cleavage of the ether bond of 2,4-dichlorophenoxyacetate. *J. Agric. Food Chem.*, 17:1080–1084.
- Vogel, T.M., Criddle, C.S., and McCarty, P.L. 1987. Transformations of halogenated aliphatic compounds. *Environ. Sci. Technol.*, 21:722–736.
- Watkinson, R.J. and Morgan, P. 1990. Physiology of aliphatic hydrocarbon degrading microorganisms. *Biodegradation*, 1:79–92.
- Wu, S. and Gschwend, P.M. 1986. Sorption kinetics of organic hydrophobic compounds to natural sediments and soils. *Environ. Sci. Technol.*, 20, 717–725.
- Xun, L., Topp, E., and Orser, C.S. 1992. Confirmation of oxidative dehalogenation of pentachlorophenol by a Flavobacterium pentachlorophenol hydroxylase. *J. Bacteriol.*, 174:5745–5747.
- Young, L.Y. and Cerniglia, C.E. 1995. *Microbial Transformation and Degradation of Toxic Organic Chemicals*, Wiley-Liss, New York.
- Zehnder, A.J.B. 1988. *Biology of Anaerobic Microorganisms*, John Wiley & Sons, New York.
- Zhu D.Q. and Pignatello, J.J. 2005. Characterization of aromatic compound sorptive interactions with black carbon (charcoal) assisted by graphite as a model. *Environ. Sci. Technol.*, 39:2033–2041.

Further Information

- Brock et al. (1997) provides an excellent comprehensive work on microbiology covering introductory cell chemistry and cell biology as well as much more advanced material.
- Environmental Security Technology Certification Program, (<http://www.estcp.org/index.cfm>).
- Gottschalk (1986) is an outstanding reference that describes basic bacterial biochemistry.
- National Research Council (1993) describes the basic fundamental principles and limitations of bioremediation.
- Office of Superfund Remediation and Technology Innovation (<http://clu-in.org>).

32

Legal Framework for Groundwater Protection in the United States

John A. Veil, Deborah
Elcock, Nancy L. Ranek,
and Markus G. Puder
Argonne National Laboratory

32.1	Introduction.....	32-2
	Relationship between Laws and Regulations • Relationship between Federal, State, and Local Laws and Regulations • Interaction between Surface Water and Groundwater • Organization of Chapter	
32.2	Groundwater Allocation and Management	32-3
	The Need for Regulation • State Groundwater Rights Systems • Federal and Indian Groundwater Rights • Takings Issues	
32.3	Clean Water Act (CWA)	32-7
	CWA Legislative Background • CWA Regulations • CWA Summary	
32.4	Safe Drinking Water Act (SDWA)	32-10
	SDWA Legislative Background • SDWA Regulations • SDWA Summary	
32.5	Resource Conservation and Recovery Act (RCRA)....	32-16
	RCRA Legislative Background • RCRA Regulations • RCRA Summary	
32.6	Comprehensive Environmental Response, Compensation, and Liability Act (CERCLA)	32-23
	CERCLA Legislative Background • CERCLA Regulations • CERCLA Summary	
32.7	Other Laws	32-28
	SMCRA • FIFRA	
	Glossary.....	32-29
	Further Information	32-32

32.1 Introduction

This chapter describes the four major Federal environmental laws that place controls on the management of groundwater. These are:

- The Clean Water Act (CWA), based on the Federal Water Pollution Control Act (FWPCA) Amendments of 1972, Public Law 92-500, 42 U.S.C. §1251 *et seq.*¹
- The Safe Drinking Water Act (SDWA), Public Law 93-523, 42 U.S.C. §300f *et seq.*
- The Resource Conservation and Recovery Act of 1976 (RCRA), Public Law 94-580, 42 U.S.C. 3251 *et seq.*, and
- The Comprehensive Environmental Response, Compensation, and Liability Act of 1980 (CERCLA or Superfund), Public Law No. 96-510, 42 U.S.C. §9601 *et seq.*, as amended.

The chapter also offers a brief discussion of groundwater protection themes in two other laws:

- The Surface Mining Control and Reclamation Act of 1977 (SMCRA), Public Law 95-87, 30 U.S.C. §1201 *et seq.*
- The Federal Insecticide, Fungicide, and Rodenticide Act (FIFRA) of 1947, Public Law 80-104, 7 U.S.C. §136 *et seq.*

The legal framework for groundwater protection has been summarized as concisely as possible while still presenting the important issues. We strongly recommend that readers consult the authentic source texts of the relevant laws and regulations before forming opinions and making decisions. Copies can be obtained from Federal and State agencies. Also, various online services and agency homepages offer electronic access.

The first edition of this handbook did not delve into water quantity issues. This second edition includes a new section (32.2) that describes groundwater allocation and management. Again, readers are encouraged to investigate the law in the light of the specific facts of their case.

32.1.1 Relationship between Laws and Regulations

A law, or statute, is a formal written enactment of a legislative body — the Congress or a State legislature. Normally, laws outline general requirements and programs without prescribing in great detail the specific methods to accomplish the statutory objectives. A regulation is a type of “delegated legislation” promulgated by a state, federal, or local administrative agency that has been given the regulatory authority by the appropriate legislature (e.g., the U.S. Environmental Protection Agency [EPA] or State and local environmental, health, or natural resources agencies). Regulations generally are very specific in nature. They are also referred to as “rules.” Regulations are promulgated following a more or less lengthy administrative procedure that includes opportunity for public comment. All Federal regulations are codified in the Code of Federal Regulations (CFR).

32.1.2 Relationship between Federal, State, and Local Laws and Regulations

Groundwater protection efforts are managed at all three levels of government. Federal laws and regulations establish basic frameworks and minimum national standards, and entrust Federal agencies with program administration. In many statutes Congress enabled the federal agencies entrusted with program management to give authority (primacy) to willing and able States. Other responsibilities, such as funding programs, conducting large research programs, and preparing baseline regulatory criteria, remain with the Federal government. When States assume Federal programs, their programs must be at least as strict

¹This is the reference for section 1251 forth following in title 42 of the United States Code. The United States Code (U.S.C.) is the codification by subject matter of the general and permanent laws of the United States. It is published by the Office of the Law Revision Counsel of the U.S. House of Representatives. The U.S.C. is divided by broad subjects into 50 titles.

or as effective as the Federal programs. States may choose to fashion stricter programs. In addition, States can adopt groundwater protection programs that are not considered under the Federal laws. For example, many States operate groundwater discharge permit programs that cover a wider range of discharges than the corresponding Federal program — the Underground Injection Control (UIC) program.

Local governments also play a role in groundwater protection. Local health departments frequently represent the first line of defense against unsound septic system installation and operation. Local and State governments are the driving force behind *source water protection programs* that are geared to prevent pollution from entering a drinking water aquifer. Under RCRA, for example, local governments must protect groundwater by complying with all applicable Federal and State regulations governing *municipal solid waste landfills (MSWLFs)*.

32.1.3 Interaction between Surface Water and Groundwater

Surface water and groundwater are intimately linked through the hydrologic cycle. Water tables move up and down seasonally, and can either contribute water to or take water from surface water bodies. This interchange of water can also carry contaminants. The technical aspects of the hydrologic cycle and surface water/ground water exchange are described in other sections of this handbook. For the most part, the major laws and regulations governing groundwater do not extensively address the overlap and exchange of surface water and groundwater. One of the few examples relates to EPA's efforts over the past decade to promote a *watershed management approach*.

32.1.4 Organization of Chapter

The remainder of this chapter discusses in detail the CWA, SDWA, RCRA, and CERCLA, including the legislative backgrounds, statutory authorities, implementing regulations, and broader perspectives. A final section highlights the SMCRA and FIFRA.

32.2 Groundwater Allocation and Management

32.2.1 The Need for Regulation

Groundwater in the United States faces increasing demands from multiple parties. In a scenario where several interests compete for the same groundwater, one person's extensive use can deny the resource to other parties. Groundwater overdraft and drawdown can create significant problems for sustaining groundwater resources over time. When groundwater stocks are exhausted, costs for drilling and pumping rise because the distance between the water stored in an aquifer and the surface increases. In addition, water quality suffers because groundwater pumped from greater depths typically contains more salts and minerals. More dramatically, the eventual depletion of the supply can create regional problems, including reductions in spring and stream flows, saltwater intrusions in coastal areas, and adverse consequences to wetlands and riparian ecosystems. Finally, in areas of severe groundwater loss, a sinking land surface can cause subsidence by creating cracks or fissures with the potential to damage roads, foundations of buildings, and underground structures. The expected population growth, especially in the southeastern and western parts of the United States, will only compound the pressures faced by a relatively limited groundwater resource.

Groundwater rights law governs the rights associated with the use of water in the ground portion of the hydrologic cycle, including the type and scope of the right, and the point in time when the right arises. Conceptually, the existence of the groundwater must be distinguished from the right to lift the groundwater. The right to lift, which involves the energy contained in the groundwater pool, allows users to withdraw a volume of groundwater by reasonable means from a water table at a fairly stable level without having to deepen their wells and use more powerful pumps. In practice, most disputes revolve around the right to lift.

Groundwater rights law in the United States is highly fragmented. Each State has developed its own unique system based on English common law, Spanish-Mexican law elements, and administrative law. Also, federal and tribal groundwater rights law exists. Sources of groundwater rights law include constitutions, statutes, administrative regulations and practice, and court decisions. In light of the complexities associated with groundwater rights law, this section only highlights key elements of the issues.

32.2.2 State Groundwater Rights Systems

The distinctions among State systems involve the foundation and scope of groundwater rights. In most States, groundwater rights spring from ownership of the overlying land surface. Several subcategories of systems have developed: absolute dominion (English rule or rule of capture), reasonable use (American rule), Restatement of Torts, and correlative rights (common resource rule). Other States separate water rights from the rights of land ownership. Groundwater belongs to the State, and any rights are based on specific administrative authorization. These jurisdictions follow the law of prior appropriation.

A survey conducted by the Water Systems Council identifies the States following the different subcategories of groundwater rights law (see Table 32.1). It is often difficult to determine a State's particular type of groundwater rights system. The courts and legislatures in the different jurisdictions have crafted specific exceptions and limitations to the various rules. Further, administrative permitting systems as well as local and regional groundwater management and conservation schemes can modify the traditional rule frameworks.

32.2.2.1 Absolute Dominion

Under the absolute dominion doctrine, landowners have the right to withdraw and use, as they please, unlimited amounts of groundwater lying beneath their surface land, unconstrained by liability to surrounding landowners who might claim that the pumping operations have depleted their wells. The absolute dominion doctrine is based on a right to pump groundwater. This includes capturing and selling the groundwater for offsite use, but does not cover the right to save it in place for subsequent use and protect it against the potential use by others.

Eight States — Connecticut, Indiana, Louisiana, Maine, Massachusetts, Mississippi, Rhode Island, and Texas — have adopted or indicate a preference for the absolute dominion rule. Among the arid western

TABLE 32.1 State Survey of Groundwater Rights Rules

Doctrine	Number of States	States in Alphabetical Order
Absolute dominion	8	Connecticut, Indiana, Louisiana, Maine, Massachusetts, Mississippi, Rhode Island, and Texas
Reasonable use	21	Alabama, Arizona, Arkansas*, Delaware*, Florida, Georgia, Illinois, Kentucky, Maryland, Missouri*, Nebraska**, New Hampshire, New York, North Carolina, Oklahoma, Pennsylvania, South Carolina, Tennessee, Virginia, West Virginia, and Wyoming*
Restatement of torts	3	Michigan, Ohio, and Wisconsin
Correlative use	6	California, Hawaii, Iowa, Minnesota, New Jersey, and Vermont
Prior appropriation	12	Alaska, Colorado, Idaho, Kansas, Montana, Nevada, New Mexico, North Dakota, Oregon, South Dakota, Utah, and Washington

Note: In conjunction with the prior appropriation rule ** In conjunction with the correlative rights rule.

Source: Water Systems Council.

states, Texas represents the lone holdout for its version of the absolute dominion doctrine — the rule of capture.

32.2.2.2 Reasonable Use

Like the absolute dominion doctrine, the reasonable use rule allows landowners to pump and use as much groundwater as they choose, as long as they make “reasonable use” of the withdrawn groundwater. This means that withdrawals must be reasonably related to the overlying land and do not result in waste. The reasonable use rule does not provide for proportional sharing or preference for prior users. Neighbors injured by reasonable use have no recourse. However, unreasonable groundwater uses that injure adjacent landowners can result in lawsuits. In traditional reasonable use jurisdictions, offsite uses, including selling and transportation, have been presumed to be unreasonable. The considerations for determining reasonableness generally include the purpose, suitability, economic and social value of the use, the overall supply quantities, the presence or absence of competition, the extent and equity of harm caused by the use, and the efficiency or wastefulness of competing uses. For example, persons are generally allowed to withdraw a fair share of the groundwater for their artificial wants (irrigation); however, natural needs (household use) enjoy priority over artificial wants.

Recognizing the need for efficient and effective groundwater management, especially in times of water scarcity, many reasonable use jurisdictions have implemented regulatory and administrative variants of the doctrine. For example, certain groundwater users who pump in quantities that exceed a regulatory threshold (e.g., 100,000 gallons per day) are held to comply with all applicable notification and permitting requirements.

Twenty-one states — Alabama, Arizona, Arkansas, Delaware, Florida, Georgia, Illinois, Kentucky, Maryland, Missouri, Nebraska, New Hampshire, New York, North Carolina, Oklahoma, Pennsylvania, South Carolina, Tennessee, Virginia, West Virginia, and Wyoming — have adopted or indicate a preference for the reasonable use rule. Arkansas, Delaware, Missouri, and Wyoming combine the reasonable use rule with the prior appropriation doctrine. Nebraska blends the reasonable use and correlative rights rules.

32.2.2.3 Restatement of Torts

Restatements are collections of law published by the American Law Institute. Although they do not have binding authority, Restatements are highly persuasive because they reflect long periods of effort and receive extensive input from law professors, practicing attorneys, and judges. Section 858 of the Restatement of Torts (second edition) “Liability for Use of Groundwater” provides specific criteria for comparing the reasonableness of competing uses of groundwater. A well owner is not liable for withdrawal of groundwater unless the withdrawal causes well interference by lowering the water table or by reducing water pressure, involves pumping more than the well owner’s reasonable share, or interferes with levels of streams and lakes that depend on groundwater. This language protects against over-pumping, and seems flexible enough to take into account uses that are more beneficial than others.

Three states — Michigan, Ohio, and Wisconsin — have adopted or indicate a preference for the Restatement of Torts approach. Also, most States with a reasonable use approach rely on some of the considerations offered in the restatement.

32.2.2.4 Correlative Rights

The correlative rights doctrine realizes that landowners using groundwater from a common pool share common rights and duties with respect to one another. Landowners may withdraw only their portion of water for reasonable use. Groundwater is prorated among the overlying landowners. When conflicts or shortages occur, each owner is entitled to a share of the available supplies that is proportionate to the amounts allocated to that landowner’s use. Unlike absolute dominion and reasonable use, a correlative rights system attempts to accommodate all overlying owners through proportional reductions when all reasonable needs cannot be met.

Six states — California, Hawaii, Iowa, Minnesota, New Jersey, and Vermont — have adopted or indicate a preference for the correlative rights rule. The doctrine originated in California. Interestingly, California

also recognizes the municipal pueblo water right possessed by a municipality as successor of a Spanish or Mexican pueblo. The pueblo right entitles the municipality to the beneficial use of all needed, naturally occurring groundwater of the original pueblo watershed. Water use under a pueblo right must occur within the modern city limits, and excess water may not be sold outside the city.

32.2.2.5 Prior Appropriation

Following the concept that water rights are separate from the rights of land ownership, Spanish and Mexican water law required specific authorization to use water. Under the appropriative rights doctrine, groundwater is allocated according to the temporal sequence of use (“first come first serve”). The first person to use water (“senior appropriator”) acquires the right (“priority”) to future use against all subsequent users (“junior appropriators”).

In almost all prior appropriation states, the water right is regulated and managed by a permit system that specifies pumping rates, well spacing, and construction requirements. Groundwater permits generally require actual and beneficial use. Permits reflect seniority, recognizing the higher priority legal right in the first user. Landowners whose usage predates the permitting system receive grandfathered rights. The appropriation date of a water right reflects the point in time when the applicant successfully demonstrates the initiation of the appropriation. Prior appropriation can apply to all groundwater, although in some states, the doctrine applies only to particular sources (underground streams).

Twelve states — Alaska, Colorado, Idaho, Kansas, Montana, Nevada, New Mexico, North Dakota, Oregon, South Dakota, Utah, and Washington — have adopted or indicate a preference for the prior appropriation rule.

32.2.2.6 Prescriptive Rights

Prescriptive rights, which exist outside the relationships among and between common law and appropriative users, are gained by trespass or unauthorized taking. The title is based on conduct allowed to continue beyond the applicable statute of limitations. Prescriptive rights can only be obtained against private entities. The claim of a prescriptive water right to non-surplus water requires that pumping was actual, open and notorious, hostile, adverse to the original rights holder, continuous and uninterrupted, and under a claim of right.

32.2.2.7 Conjunctive Management

A common practice among western states with respect to groundwater involves conjunctive management — the coordinated and integrated management of surface and groundwater resources. Groundwater is often connected to surface water. For example, seepage from a stream may recharge an underlying aquifer, or a particular stream may be fed by an aquifer. Several states are managing interconnected surface water and groundwater in a single system that recognizes the importance of the whole hydrological cycle, and facilitates better management of water resources.

32.2.3 Federal and Indian Groundwater Rights

Federal authority and jurisdiction over groundwater rights is based on several clauses of the U.S. Constitution. The Commerce Clause permits federal regulation of groundwater that may be involved in or affect interstate commerce. The Property Clause allows federal regulation of groundwater as necessary for the beneficial use of federal property. The Compact Clause bars state compacts without the consent of Congress; for example, the creation of interstate river basin commissions regulating groundwater withdrawals through permits and fee schedules. In many instances, federal law requires federal agencies managing groundwater resources to defer to state laws and cooperate with state regulators. Examples of such legislation include the Reclamation Act, the Water Supply Act, the McCarran Amendment, and the Endangered Species Act.

In interstate disputes over groundwater rights federal courts resolve the issue through equitable apportionment. Reasonable and efficient usage and volumes of prior appropriation constitute important factors for making groundwater allocations among states. In controversies among appropriative rights

systems, the groundwater supply is generally allocated by way of seniority. In disputes involving states that follow ownership doctrines, courts will generally follow a reasonableness standard to maintain groundwater availability to all states concerned. Finally, the US Supreme Court has held that states are barred from obstructing the exportation of groundwater to other states.

Indian groundwater rights law must be carefully evaluated and administered within the prevailing groundwater allocation systems. In light of the early establishment of Indian reservations often predating the onset of most other water uses, especially in the west, tribes generally retain the oldest and most valuable groundwater rights. Indian groundwater rights generally fall into aboriginal, pueblo, and *Winters* rights. Aboriginal groundwater rights attach to lands occupied since time immemorial. They are accorded a corresponding priority date. Pueblo groundwater rights enjoy early priority dates in the wake of Spanish land grants and the US Treaty of Guadalupe Hidalgo with Mexico. Most federal Indian groundwater rights are examined in light of the *Winters* decision by the US Supreme Court. In *Winters*, the high court held that when ceding to the federal government vast territories of land in return for permanent reservations for Indian occupation and use, the tribes reserved, along with the land, the amount of water required to meet the needs and purposes of the reservation. This means that Indian water rights are created according to the date of the reservation. Actual quantification and sizing of vested Indian water rights occurs through litigation and congressional action. The most widely used method is the practicable irrigation acreage approach, which quantifies the amount of water needed to irrigate the reservation's arable lands. In general, the amounts are measured in terms of (future) needs, as opposed to the quantities in actual use. On reservations where state, federal, and tribal jurisdictional interests and concerns compete, courts will regularly fashion the type of applicable scheme of groundwater rights administration. In general, courts have recognized important tribal sovereignty interests to regulate Indian and non-Indian water use on a reservation.

32.2.4 Takings Issues

Questions of property takings through governmental action can be associated with groundwater use and regulation. Groundwater pumping operations by governmental entity (municipality) pumps can physically invade the landowner's property by diverting water from beneath the landowner's property. And high-volume groundwater pumping can deny the landowner all economically viable uses of the property. Other issues of regulatory property takings are raised when a state converts groundwater rights under ownership doctrines to an appropriative rights system. In all scenarios, courts will need to determine whether a governmental action rises to the level of a taking. The outcome will depend on the specific facts of each case.

32.3 Clean Water Act (CWA)

32.3.1 CWA Legislative Background

32.3.1.1 Historical Perspective

The CWA was passed in 1972 as a comprehensive package of amendments to the Federal Water Pollution Control Act of 1948. The name "Clean Water Act" was not applied to the law until 1977, when another round of amendments was enacted. The 1972 amendments, which gave the law its current shape, serve as the Federal foundation for water quality protection and water pollution control. Through the CWA Congress responded to the country's growing public concerns over water quality degradation in its rivers, lakes, and oceans. Only a few years earlier the Cuyahoga River in Cleveland had actually caught fire due to floating oils and other flammable contaminants. In the light of the Nation's greater awareness of surface water issues, the CWA is mainly geared to the surface water component of the hydrological cycle. The amendments of 1977, 1981, and 1987 did little to change this focus of the CWA. This also holds true for the minor revisions to the CWA made by Congress since 1987.

32.3.1.2 Goals and Objectives

The 1972 law declared as its objective “to restore and maintain the chemical, physical, and biological integrity of the Nation’s waters.” The phrase “Nation’s waters” is never explained, but the term “navigable waters,” defined as “the waters of the United States,” is used throughout the CWA. The definition, which determines the reach of the CWA, has been litigated on numerous occasions over the recent years. Navigable waters include surface water bodies but not groundwater.

The goals listed in §101 of the CWA do not directly speak to groundwater. However, §102(a) states that EPA “shall, after careful investigation . . . prepare or develop comprehensive programs for preventing, reducing, or eliminating the pollution of the navigable waters and ground waters and improving the sanitary condition of surface and underground waters.” Through summer 2005, no regulatory programs have been promulgated under the CWA to implement this section. Several other references to groundwater are sprinkled throughout the CWA. §104 directs EPA to maintain a groundwater monitoring network. Under §208, states are to identify areas that have substantial water quality problems and develop areawide waste treatment management plans. The plans must include, among others, processes that identify contaminated underground mine drainage, recognize salt water intrusion due to excessive groundwater extraction, and control disposal of pollutants on the land or in subsurface excavations within such area to protect groundwater quality.

§304 requires EPA to develop and publish water quality criteria that reflect the latest scientific knowledge on effects on health and welfare, which result from the presence of pollutants in groundwater. Under §304 EPA must develop and publish information on the factors necessary to restore and maintain the chemical, physical, and biological integrity of all navigable waters and groundwater. §304 also directs EPA to issue to appropriate Federal agencies, the states, water pollution control agencies, and certain agencies processes, procedures, and methods to control pollution resulting from various sources, including: underground mining activities; the disposal of pollutants in wells or in subsurface excavations; salt water intrusion resulting from reductions of fresh water flow (e.g., in the wake of extraction of ground water); and changes in the movement, flow, or circulation of any navigable waters or ground waters. EPA has fulfilled some of these obligations through other programs, particularly those under the SDWA, but has not developed a comprehensive program under the CWA. §319 — discussed below — offers the only program specifically geared toward protecting groundwater. States have accessed grant funding under §106 to develop their State groundwater protection and wellhead regulations.

32.3.1.3 General Provisions

The primary regulatory control mechanism of the CWA — the National Pollutant Discharge Elimination System (NPDES) program (§402) — covers point source discharges (mainly from industries and sewage treatment plants, including publicly owned treatment works — POTWs). Congress also established a comparable pretreatment program for industries that discharge wastewater to POTWs rather than directly to navigable waters.

Nonpoint sources of pollution are not subject to a formal Federal regulatory program such as the NPDES program. Under §319, states are required to submit nonpoint source management programs to EPA for approval. The State programs must outline the best management practices and measures that will be used to reduce nonpoint source pollution, and the regulatory or non-regulatory programs (including enforcement, technical and financial assistance, education, training, technology transfer, and demonstration projects) designed to achieve implementation. §319 represents one of the few sections of the CWA that explicitly references groundwater protection by name. State management programs are to consider impacts of the recommended practices on groundwater quality (§319(b)(2)(A)). When making Federal nonpoint source grants, EPA must give priority to projects that will protect groundwater from nonpoint sources (§319(h)(5)(D)). Further, EPA may make grants that assist states in carrying out groundwater protection activities designed to implement a comprehensive nonpoint source program (§319(i)). Unfortunately, nonpoint source control programs have not been widely successful. This may be attributable to the lack of a strong Federal program, insufficient levels of Federal financial support, and

the absence of responsible and financially viable parties (such as industries or POTWs) willing and able to carry out the control program.

Although the Federal NPDES program is limited to discharges to surface water, many State water pollution control laws are broader and mandate regulatory programs to protect both the surface water and groundwater of their state. For example, some states require permits for discharges of wastewater not only to surface water but also to groundwater or the land surface. Septic systems for disposal of sewage or household wastes are not regulated at the Federal level through the CWA, but are controlled by State and local regulations. As the details of these State programs vary widely, readers are advised to consult with the appropriate State regulatory agencies.

32.3.1.4 Issues and Outlook

As of mid-2005, Congress is not working on any significant revisions to the CWA in the area of groundwater quality and protection. Over the past decade, EPA has emphasized a watershed management approach that involves consideration of all activities occurring within a watershed rather than focusing on a single regulatory program. A well-designed and operated watershed management program ideally considers impacts of various activities on groundwater quantity and quality. EPA has begun to encourage watershed management programs; however, these programs might have greater success and clout if backed up by explicit statutory authority.

32.3.2 CWA Regulations

The CWA regulations do not exercise direct regulatory controls over groundwater. However, the point source stormwater regulations (40 CFR 122.46²), which require that stormwater discharges associated with industrial activity, stormwater discharges from construction sites, and discharges from large and medium municipal separate storm sewer systems must be covered by NPDES permits, do indirectly affect groundwater. In most instances, stormwater permits do not use numerical limits as compliance tools. Instead, they require that permittees adopt best management practices (BMPs), which can be grouped into three categories. A first type of BMPs prevents clean stormwater from coming into contact with contaminant sources. This can for example be accomplished through berms, curbing, and roofing. Another category of BMPs treats the contaminated runoff to lower pollutant levels. Treatment technologies in this category include oil/water separators, straw bale dikes, and sediment ponds. A third type of BMPs involves reducing the total volume or rate of runoff to surface waters — for example, through detention and retention ponds, or stormwater drainage wells. The third category of BMPs represents a very effective mechanism for controlling surface water quality, but can conceivably contribute to groundwater contamination. For example, dissolved pollutants contained in the stormwater can infiltrate through the bottom of detention or retention ponds and enter the groundwater. Disposal of stormwater into drainage wells directly introduces pollutants into the groundwater.

Nonpoint source and point source stormwater discharges contain the same contaminants. EPA has not adopted nonpoint source stormwater regulations; rather control programs in this area rely almost entirely on BMPs, and groundwater contamination scenarios comparable to the point source category described above exists. EPA's sewage sludge regulations (40 CFR 503) represent another example of CWA regulations with indirect effects on groundwater quality and protection. Two of the approved methods for sewage sludge management involve land application and surface disposal. The land application and surface disposal regulations do not contain direct groundwater protection controls. Rather, the prescribed application rates and management practices are designed to protect surface water from excessive land application of sludge contaminants.

²40 CFR 122.46 refers to Section 122.46 in Part 122 of Title 40 of the Code of Federal Regulations.

32.3.3 CWA Summary

Only a few sections in the CWA speak directly to groundwater. Yet, §319 — the nonpoint source control program — offers a statutory lever for imposing groundwater protection requirements on landowners, businesses, farmers, or other dischargers.

32.4 Safe Drinking Water Act (SDWA)

32.4.1 SDWA Legislative Background

32.4.1.1 Historical Perspective

The SDWA, Title XIV of the Public Health Service Act, was passed in 1974 to establish a national framework for ensuring a clean drinking water supply. The SDWA establishes two main programmatic themes. The public water system program specifies how clean drinking water must be when it is delivered to customers. A variety of other programs protect underground sources of drinking water (USDWs).

Minor amendments to the SDWA were made in 1976, 1977, and 1980, and the first significant changes were adopted in 1986. The 1986 amendments greatly expanded the scope of contaminants subject to drinking water standards, and increased the level of monitoring for nonregulated contaminants. States were required to establish wellhead protection areas. In 1996, Congress passed another round of comprehensive amendments to the SDWA. Many of those have affected the public water supply system requirements by changing how drinking water standards are established and implemented. Consideration was given for small public systems with limited economic resources. Much greater emphasis was placed on protecting source water, and the new source water requirements have indeed helped to improve ground water quality.

32.4.1.2 Goals and Objectives

The public water system program strives to ensure that drinking water at the tap is clean and safe to drink. This is accomplished through a combination of treatment requirements (like filtration and disinfection) and numerical quality standards for individual contaminants. The SDWA focuses solely on public water systems, which are defined as those that have at least 15 service connections or that serve at least 25 persons. The SDWA does not require testing, controls, or standards for those drinking water systems that supply fewer than the specified number of connections or persons (e.g., private wells).

The remaining SDWA programs are designed to protect USDWs from contamination by restricting activities in sensitive areas and controlling discharges of contaminants to groundwater or to the subsurface soils. The SDWA does not contain program components specifically requiring protection of surface water supplies. Under the SDWA, not all groundwater must be protected; protection is required only for USDWs.

32.4.1.3 General Provisions

All major SDWA programs affect groundwater to varying degrees. The national primary drinking water regulations form the heart of the public water system program. EPA has published maximum contaminant levels (MCLs) and maximum contaminant level goals (MCLGs) for more than 75 contaminants (see Table 32.2). The MCLGs reflect the levels of contaminants in drinking water below which no known or expected risk to health exists. MCLGs allow for a margin of safety and represent non-enforceable public health goals. For many contaminants, MCLGs cannot be economically or technically achieved. MCLs are the highest concentrations of contaminants allowed in drinking water. MCLs are established as close to MCLGs as feasible using the best available treatment technology and taking cost into consideration. In some cases, EPA has the authority to establish treatment requirements that control a particular contaminant in lieu of promulgating a numerical MCL. Under the SDWA, variances and waivers from meeting MCLs or treatment requirements are available for public drinking water systems. The primary

TABLE 32.2 EPA's National Primary Drinking Water Standards

Contaminant	MCLG (mg/l)	MCL (mg/l)
<i>Inorganics</i>		
Antimony	0.006	0.006
Arsenic	0	0.010
Asbestos (>10 um)	7 MFL ^a	7 MFL
Barium	2	2
Beryllium	0.004	0.004
Cadmium	0.005	0.005
Chromium (total)	0.1	0.1
Copper	1.3	TT ^b
Cyanide	0.2	0.2
Fluoride	4	4
Lead	0	TT
Mercury (inorganic)	0.002	0.002
Nitrate	10	10
Nitrite	1	1
Selenium	0.05	0.05
Thallium	0.0005	0.002
<i>Coliform and surface water treatment</i>		
<i>Cryptosporidium</i>	0	TT
Giardia lamblia	0 detected	TT
Legionella	0 detected	TT
Heterotrophic plate count	N/A	TT
Total coliform	0 detected	c
Turbidity	N/A	TT
Viruses	0 detected	TT
<i>Organics</i>		
Acrylamide	0	TT
Alachlor	0	0.002
Atrazine	0.003	0.003
Benzene	0	0.005
Benzo-a-pyrene	0	0.0002
Carbofuran	0.04	0.04
Carbon tetrachloride	0	0.005
Chlordane	0	0.002
Chlorobenzene	0.1	0.1
2,4-D	0.07	0.07
Dalapon	0.2	0.2
1, 2-Dibromo-3-chloropropane	0	0.0002
o-Dichlorobenzene	0.6	0.6
p-Dichlorobenzene	0.075	0.075
1, 2-Dichloroethane	0	0.005
1,1-Dichloroethylene	0.007	0.007
cis-1, 2-Dichloroethylene	0.07	0.07
trans-1, 2-Dichloroethylene	0.1	0.1
Dichloromethane	0	0.005
1, 2-Dichloropropane	0	0.005
Di(2-ethylhexyl) adipate	0.4	0.4
Di(2-ethylhexyl) phthalate	0	0.006
Dinoseb	0.007	0.007
Diquat	0.02	0.02
Dioxin	0	0.00000003
Endothall	0.1	0.1
Endrin	0.002	0.002
Epichlorohydrin	0	TT
Ethylbenzene	0.7	0.7

Continued

TABLE 32.2 *Continued*

Contaminant	MCLG (mg/l)	MCL (mg/l)
Ethylene dibromide	0	0.00005
Glyphosate	0.7	0.7
Heptachlor	0	0.0004
Heptachlor epoxide	0	0.0002
Hexachlorobenzene	0	0.001
Hexachlorocyclopentadiene	0.05	0.05
Lindane	0.0002	0.0002
Methoxychlor	0.04	0.04
Oxyamyl	0.2	0.2
PCBs	0	0.0005
Pentachlorophenol	0	0.001
Phthalate (di-2-ethylhexyl)	0	0.006
Picloram	0.5	0.5
Simazine	0.004	0.004
Styrene	0.1	0.1
Tetrachloroethylene	0	0.005
Toluene	1	1
Toxaphene	0	0.003
2, 4, 5-TP	0.05	0.05
1, 2, 4-Trichlorobenzene	0.07	0.07
1,1,1-Trichloroethane	0.2	0.2
1,1, 2-Trichloroethane	0.003	0.005
Trichloroethylene	0	0.005
Total trihalomethanes	N/A	0.080
Vinyl chloride	0	0.002
Xylenes (total)	10	10
<i>Disinfectants and disinfectant byproducts</i>		
Bromate	0	0.010
Chloramines (as Cl ₂)	4.0 ^d	4.0 ^d
Chlorine (as Cl ₂)	4.0 ^d	4.0 ^d
Chlorine dioxide (as ClO ₂)	0.8 ^d	0.8 ^d
Chlorite	0.8	1.0
Haloacetic acids (HAA5)	N/A	0.06
Total trihalomethanes (TTHMs)	N/A	0.080
<i>Radionuclides</i>		
Radium 226 + radium 228	0	5 pCi/l
Alpha particles	0	15 pCi/l
Beta particle + photon radioactivity	0	4 mrem/yr
Uranium	0	0.030

^a MFL = million fibers per liter.^b TT = treatment techniques in lieu of a numerical standard.^c No more than 5% of samples can be total coliform-positive.^d Expressed as maximum residual disinfectant.

MCLs described above regulate the concentration of numerous naturally occurring and synthetic, organic, and inorganic contaminants in drinking water. Primary MCLs are enforceable. EPA also has adopted secondary MCLs for 16 drinking water contaminants or constituents. Secondary MCLs are suggested guidelines that are related to aesthetic qualities (taste and smell) of drinking water. They rarely present health threats to consumers because the drinking water in question would become unpalatable well before the constituents could possibly reach harmful levels. Secondary MCLs are not enforceable. The public water system requirements under the SDWA do not directly control groundwater quality. They are designed to protect drinking water consumers and not the groundwater itself. For example, a water supplier can withdraw groundwater that does not meet all MCLs, treat it, and then supply it to customers. However,

MCLs play important roles in several other regulatory programs. The Underground Injection Control (UIC) program frequently requires that MCLs must be met at the point where waste injection occurs, where the discharge plume intersects the USDW, or where the USDW leaves the property boundaries. Under CERCLA, the MCLs are often enlisted as applicable or relevant and appropriate cleanup standards. This is discussed in Section 32.6.2.1.

The SDWA contains several important USDW protection programs. The UIC program governs injection of fluids into wells (§§1421–1424). The SDWA envisions that willing and able states will have the opportunity to gain UIC program primacy (§§1422 and 1425). Many states have been granted UIC program authority over certain or all classes of *injection wells*. Otherwise EPA's regional offices administer the UIC program in direct implementation. Underground injection operations are subject to authorization by permit or rule, and must not endanger USDWs.

The second USDW protection program is the *sole source aquifer* program. Under §1424(e) and §1427, EPA has the authority to designate certain aquifers as the sole or principal drinking water source for a given area. When an aquifer is designated as the sole source aquifer, EPA is barred from committing Federal financial assistance (through a grant, contract, loan guarantee, or otherwise) for any project deemed to contaminate such aquifer through a recharge zone and create a significant public health hazard. As of 2005, EPA has designated more than 70 aquifers as “sole source.” State and local governments and regional planning entities with jurisdiction over certain critical aquifer protection areas may apply to EPA for demonstration-program funding.

When selecting critical aquifer protection areas EPA must consider the vulnerability of the aquifer to contamination, the population using the aquifer as their drinking water source, the economic, social, and environmental benefits resulting from protecting the aquifer, and the economic, social, and environmental costs resulting from aquifer degradation. The application for a demonstration program must include a comprehensive management plan for the proposed protection area. Such plans would, among other items identify existing and potential point and nonpoint sources of ground water degradation; assess the relationship between activities on the land surface and ground water quality; specify implementation and management activities and practices for preventing adverse impacts on ground water quality in the critical protection area; determine the quality of the existing groundwater recharged through the special protection area, and the natural recharge capabilities of the special protection area watershed; provide requirements designed to maintain existing underground drinking water quality or improve underground drinking water quality if prevailing conditions fail to meet drinking water standards; and describe actions in the special protection area that would avoid adverse impacts on water quality, recharge capabilities, or both. The third USDW protection program relates to protection of wellheads. Under this program, states are directed to protect the areas around water supply wellheads from contaminants (§1428). Wellhead protection programs must delineate the size of the protection area for each wellhead and identify the activities that will be undertaken for protecting the water supply within wellhead protection areas from contaminants. The Federal requirements for the wellhead protection program provide a framework and funding for the program. The actual regulatory requirements and controls are designed and administered at the State or local levels. With the advent of source water protection programs in the wake of the 1996 amendments, many wellhead protection programs have been made part of new source water protection programs.

Two new USDW protection programs were added to the SDWA in 1996. §1429 authorizes EPA to make grants to states for developing and implementing comprehensive programs for groundwater protection. Under §1453, states are required to develop and implement source water assessment programs protective of surface water and groundwater drinking water. State source water assessment programs must delineate the boundaries of assessment areas, identify contaminants deemed to be a public health threat, and trace the origins of those contaminants. Under §1454, states may establish programs that enable community water systems or local governments to submit source water quality protection partnership petitions to the state requesting assistance in developing voluntary, incentive-based partnerships to reduce the presence of contaminants in drinking water, receive financial or technical assistance, and develop a long-term source water protection strategy.

32.4.1.4 Issues and Outlook

The SDWA was reauthorized and amended in 1996. No significant revisions have been made since then.

32.4.2 SDWA Regulations

EPA's extensive regulations for the public drinking water program are codified in 40 CFR Parts 141–143. As discussed above, most of the drinking water requirements are not directly based on protecting groundwater. As this chapter concerns groundwater issues, only the MCLs and MCLGs portions, as opposed to the full range of drinking water programs, are discussed here. Many parts of the drinking water regulations were adopted at different times; therefore, not all MCLs or MCLGs are housed in one single consolidated section of the regulations. MCLs are found in 40 CFR §§141.11–16, 141.61–64, and 141.66; and MCLGs in 40 CFR §§141.50–53, and .55. All MCLs and MCLGs that have been promulgated by EPA are listed in Table 32.2. EPA's detailed UIC regulations are codified in 40 CFR Parts 144–148. Underground injection, unless authorized by permit or rule, is prohibited by §144.11. Moreover, §144.12 states: “no owner or operator shall construct, operate, maintain, convert, plug, abandon, or conduct any other injection activity in a manner that allows the movement of fluid containing any contaminant into underground sources of drinking water, if the presence of that contaminant may cause a violation of any primary drinking water regulation under 40 CFR Part 142 or may otherwise adversely affect the health of persons.”

§146.4 provides for aquifer exemptions if the aquifer does not currently serve and cannot in the future serve as a USDW by virtue of its depth, location, existing contamination, or its use for mineral or hydrocarbon extraction; the total dissolved solids content of the aquifer's ground water is more than 3,000 mg/l and less than 10,000 mg/l; and the aquifer is not reasonably expected to supply a public water system.

§146.5 establishes five classes of injection wells. Class I wells are used for the disposal of industrial hazardous, industrial non-hazardous wastes, and municipal (non-hazardous) waste below the lowermost USDW. §§146.11–14, and .61–.73 contain detailed requirements for siting, constructing, operating, monitoring, and closing Class I wells to ensure that contaminants are unable to migrate into a USDW. Parts 144.60–70 define the financial responsibilities for Class I well owners and operators. Nearly 500 Class I injection wells at more than 290 active facilities are scattered throughout the United States in 19 states. The greatest concentration of Class I wells occurs in the Gulf Coast, Great Lakes, and Florida regions. Class II wells are used for injection of oil and gas industry wastes, injection of fluids for enhanced oil recovery, or underground storage of liquid hydrocarbons. §§146.21–24 provide detailed requirements for construction, monitoring, operations, and closure. Approximately 170,000 Class II wells exist in 31 states.

Class III wells are used for mineral extraction. The two primary mining processes that use Class III wells are solution mining and *in situ* leaching. §§146.31–34 contain detailed requirements for construction, monitoring, operations, and closure. The Class III category includes approximately 20,000 wells at 200 facilities in 18 states, primarily in the South and Southwest. Under the Federal UIC regulations, all Class I, II, or III wells must be permitted by the EPA or an approved primacy state (§144.31).

Class IV wells inject hazardous or radioactive wastes into or above a USDW (§144.13). Class IV wells, which are considered to be a threat to USDWs, are prohibited, except when used for CERCLA or RCRA cleanups, or injections to exempted aquifers. Identified Class IV wells must be closed and considered for remediation.

The final class of injection wells is not a specific type of well but serves as a catch-all for injection wells that do not fall into one of the other classes. Class V wells include drywells, sumps, drain fields, drainage wells, and some septic systems, among others. Class V wells have been widely used in many parts of the country without having typically been subjected to extensive permitting requirements. Most Class V wells are drainage wells or wells involving sewage disposal — septic tanks, drain fields, and cesspools. EPA does not have an accurate count or even rough estimate of the number of Class V wells in the United States. The number very likely exceeds 1 million wells.

New Class V regulations were adopted in 2002 (144.79–.89). Class V wells are generally authorized by rule — a type of authorization given through blanket regulatory approval. However, large-capacity cesspools and motor vehicle waste disposal wells located in a groundwater protection area or sensitive groundwater area must be permitted or closed. Other types of Class V wells are subject to various regulatory requirements. Movement of fluid into USDWs that might cause endangerment is not permitted. Moreover, well owners must comply with other Federal or State UIC requirements. And finally, wells must be properly closed. Under certain circumstances EPA can require a permit for individual wells (§§144.24–.25). §§144.51–.55 contain various conditions that are applicable to all UIC permits. In 40 CFR Part 147 the applicable Underground Injection Control (UIC) programs for the states, territories, and possessions are codified. The applicable UIC program for a state is either a State-administered program approved by EPA, or a federally administered program promulgated by EPA. In some cases, the UIC program may consist of a State-administered program applicable to some classes of wells and a federally administered program applicable to other classes of wells. In all cases readers should consult 40 CFR Part 147 concerning the current program authorization status for UIC activities in a particular state.

40 CFR Part 148 identifies hazardous wastes that are restricted from disposal into Class I hazardous waste injection wells, and defines those circumstances under which a waste that is otherwise prohibited from injection may be injected. EPA's hazardous waste injection restrictions actually implement §3020 of RCRA, but are discussed here because EPA's UIC regulatory program covers both SDWA and RCRA requirements.

EPA's sole source aquifer regulations are codified in 40 CFR Part 149. Federal regulations are minimal as the language in the SDWA (§§1424(e) and 1427) is already rather specific. Subpart A (§§149.1–.3) provides two alternative sets of criteria for identifying critical aquifer protection areas (§149.3). A critical aquifer protection area is an area designated as a sole or principal source aquifer prior to June 19, 1986, and for which an areawide groundwater quality protection plan was approved prior to that date. A critical aquifer protection area is also a major recharge area of a sole or principal source aquifer, designated before June 19, 1988, for which the sole or principal source aquifer is particularly vulnerable, contamination is likely to occur without a control program, and significant cost would accrue. Subpart B (§§149.100–.110) describes Federal requirements for protecting the Edwards Underground Reservoir. This is the sole or principal drinking water source for the San Antonio area, and if contaminated, would create a significant hazard to public health. The Edwards Underground Reservoir is the only aquifer subject to specific regulations promulgated by EPA. EPA's Office of Ground Water and Drinking Water periodically publishes a list of the sole source aquifers in the United States. The types of requirements placed on these sole sources aquifers through State programs are similar to the Federal regulations governing the Edwards Underground Reservoir.

EPA has not adopted regulations governing wellhead protection programs, although EPA does require states to submit wellhead protection plans (§1428 of the SDWA). As in the case for the sole source aquifer program, the SDWA language for the wellhead protection program is rather specific. Most states have developed wellhead protection programs, and the actual implementation of most wellhead protection programs occurs at the state and local levels. Over 4000 communities nationwide conduct some form of wellhead protection.

This section has only presented the Federal SDWA regulations. As states take a lead role in groundwater protection activities, readers are encouraged to work with state regulatory agencies to determine specific groundwater protection requirements. Although not required by the SDWA, EPA has encouraged each state to develop a Comprehensive State Ground Water Protection Program (CSGWPP) that focuses on a hierarchy of priorities: prevention of contamination whenever possible; prevention of contamination based on the relative vulnerability of the resource, and where necessary, the groundwater's use, value, and vulnerability; and remediation commensurate with the relative use and value of groundwater. A CSGWPP incorporates goal and priority setting, assignment of responsibilities to different stakeholders, implementation, data collection, and public participation. States with CSGWPPs are eligible for greater flexibility in grant funding and program administration.

32.4.3 SDWA Summary

The SDWA places indirect and direct controls on groundwater quality. The MCLs serve as overriding targets for groundwater quality even though they are intended as drinking water standards that must be met at the tap. The UIC program regulates injection of fluids and wastes into underground formations. The sole source aquifer, wellhead protection, groundwater protection, and source water protection programs are designed to prevent drinking water aquifers from becoming contaminated. States play the largest role in program managing and receive extensive support from local governments.

32.5 Resource Conservation and Recovery Act (RCRA)

32.5.1 RCRA Legislative Background

32.5.1.1 Historical Perspective

RCRA, as enacted in 1976, gave EPA broad authority to develop a hazardous and solid waste management regulatory program to promote recycling and ensure protective waste management practices. However, little guidance was included on how to do it, and EPA first focused on the regulation of non-hazardous solid wastes. As a result, by 1984, the Agency had managed to issue permits (EPA's approach to ensuring protective waste management practices) to only five of the thousands of operating hazardous waste land disposal facilities. Meanwhile, Congress had become even more aware of the effects that land disposal of hazardous wastes was having on the environment (especially groundwater). Congress was intent on preventing contaminated sites that could require future cleanups under CERCLA, but lacked confidence in EPA's ability to develop an effective program. Hence, the Hazardous and Solid Waste Amendments of 1984 (HSWA) were passed, containing a level of detail unprecedented in environmental law and fundamentally altering the RCRA program. Provisions were included that would automatically take effect unless EPA issued within specified time frames regulations to significantly reduce, and comprehensively control, land disposal of hazardous wastes. As a result, between 1986 and 1998, treatment standards specifying the method or level of treatment for all hazardous wastes were identified. After 1998, development and updating of hazardous waste treatment standards became an ongoing process that continues as new hazardous wastes are identified. The HSWA also mandated that EPA create a Corrective Action Program under which cleanups have been initiated at RCRA-permitted hazardous waste treatment, storage, and disposal facilities where waste has been determined to be leaking into the environment from any source.

32.5.1.2 Goals and Objectives

§1003(b) of RCRA declares the national policy of the United States to be: "wherever feasible, the generation of hazardous waste is to be reduced or eliminated as expeditiously as possible. Waste that is nevertheless generated should be treated, stored, or disposed of so as to minimize the present and future threat to human health and the environment."

32.5.1.3 General Provisions

In general, RCRA regulates solid waste, which includes nonhazardous solid waste and hazardous waste produced by industrial processes. Subtitle D of the statute deals with nonhazardous solid waste by directing EPA to promulgate guidelines for use by states, which are directed to develop and implement solid waste management plans and to establish permitting programs for certain solid waste management facilities. Subtitle C, on the other hand, authorizes EPA to develop and implement all or portions of an extensive Federal program for regulating the management of hazardous wastes. Upon request, qualified states can be authorized by EPA to implement the Federal hazardous waste program. State regulations must be at least as stringent as EPA's. When EPA finalizes new Federal hazardous waste regulations, EPA implements them in all states until the states become authorized, unless the new Federal regulations are based on RCRA authority that pre-dated HSWA, or are less stringent than existing ones. In the latter two situations,

the new Federal regulations usually take effect in a state only after the state has adopted and received EPA approval for implementing its regulations. States also sometimes use their own laws as authority to regulate hazardous wastes, independent of RCRA. Readers should check with the state waste management agencies to determine which state and federal hazardous waste regulations are applicable to their facilities.

Also of interest in RCRA is Subtitle I, which regulates underground storage tanks (USTs) containing petroleum and hazardous substances as defined by the CERCLA. Subtitle I does not cover USTs containing hazardous wastes that are regulated under Subtitle C. Like Subtitle C, Subtitle I allows states to be authorized by EPA to implement the Federal UST program, provided that their regulations are no less stringent than the corresponding EPA requirements.

32.5.1.4 Issues and Outlook

During the mid- and late 1990s, leaders in both the Administration and Congress were interested in amending RCRA in ways that would speed the cleanup of hazardous waste sites. However, despite much discussion of draft legislation at the time, no bills were passed, and the momentum for such legislation subsided significantly after 2000. Notwithstanding, improvements in the coordination, speed, and effectiveness of cleanups continues to be an issue. EPA's One Cleanup Program (<http://www.epa.gov/oswer/onecleanupprogram/>) was established to address this issue by (1) promoting more consistent and effective cleanups, (2) providing clear and useful information to stakeholders about the cleanup of contaminated properties, and (3) developing meaningful measures to demonstrate the overall effectiveness and benefit of cleanup efforts.

Another primary focus of EPA's RCRA regulatory activities is the goal of reducing the amount of waste generated and changing the ways in which society identifies materials as wastes. In furtherance of this goal, EPA continues to pursue finalization of its proposed rule entitled "Revisions to the Definition of Solid Waste," which would revise the regulatory definition of solid waste by identifying certain recyclable hazardous secondary materials as not "discarded" and thus not subject to regulation as wastes under Subtitle C of RCRA (68 FR 61558; October 28, 2003). In addition, EPA is promoting legislation, regulations, and voluntary programs that encourage the reuse or recycle of electronic products. EPA also supports development and use of Environmental Management Systems by all types of organizations to provide a structured approach for managing environmental and regulatory responsibilities.

32.5.2 RCRA Regulations

Because of the large number of RCRA regulations that deal with groundwater protection, this section of the chapter relies primarily on summary tables.

32.5.2.1 Hazardous Waste

32.5.2.1.1 Introduction

Regulations implementing RCRA Subtitle C, "Hazardous Waste Management," are located in 40 CFR 260–279. Of these, Part 264, "Standards for Owners and Operators of Hazardous Waste Treatment, Storage and Disposal Facilities," Part 265, "Interim Status Standards for Owners and Operators of Hazardous Waste Treatment, Storage and Disposal Facilities," and Part 280, "Underground Storage Tanks," contain specific groundwater monitoring provisions. HSWA created two regulatory programs particularly designed to prevent the migration of hazardous constituents into groundwater as a result of hazardous waste management activities: the Land Disposal Restrictions (LDR) program and the Corrective Action program. Regulations addressing corrective action are located in Part 264, Subparts F, "Releases from Solid Waste Management Units," and S, "Special Provisions for Cleanup." The LDR program is governed by Part 268, "Land Disposal Restrictions," which identifies hazardous wastes that are restricted from land disposal and defines the limited circumstances under which an otherwise restricted waste may continue to be stored and land disposed.

Elsewhere in Title 40 of the CFR, Part 260 contains definitions and general provisions that apply to all facilities and wastes governed by RCRA Subtitle C. Part 261 defines the terms *solid waste* and *hazardous*

waste. For a waste to be regulated under RCRA Subtitle C, it must first fall within the definition of solid waste. Waste that does not meet this definition cannot be RCRA hazardous waste. Unfortunately, the definitions of solid waste and hazardous waste are very complicated, making it sometimes difficult to determine whether a particular waste is regulated.

32.5.2.1.2 Facility Standards

Once a waste meets EPA's hazardous waste definition, generators, transporters, and all facilities that treat, store or dispose of it, must comply with applicable RCRA regulatory requirements. Such requirements are intended to encourage minimization of hazardous waste generation, and ensure that hazardous waste management facilities and practices prevent migration of hazardous constituents into the environment, including groundwater. They are also intended to ensure that if hazardous constituents do migrate into the environment, corrective action will occur.

Of particular concern with respect to groundwater protection are hazardous waste facilities and practices involving placement of hazardous waste on or into the land. Therefore, RCRA regulations governing such facilities and practices contain specific groundwater protection provisions, which are summarized briefly in Table 32.3 and Table 32.4. Table 32.3 covers the regulations applicable to owners and operators of hazardous waste treatment, storage and disposal facilities that hold RCRA permits (Part 264). Table 32.4 covers the regulations applicable to interim status facilities (Part 265). To qualify for interim status, the owner or operator must: (a) have been treating, storing, or disposing of the hazardous waste, or commenced facility construction on or before October 31, 1976 (the date RCRA was enacted), or on or before the date on which the waste being managed became subject to RCRA permitting requirements; (b) notify EPA of hazardous waste activities, as required in RCRA §3010; and (c) apply for a permit under Part 270.

Generally, the Part 265 regulations specify a minimal groundwater monitoring program to detect hazardous releases to the uppermost aquifer at interim status surface impoundments, landfills, and land treatment facilities. In contrast, Part 264 imposes minimum technology designs on tanks, surface impoundments, waste piles, land treatment facilities, and landfills that are seeking RCRA permits, and requires all *solid waste management units (SWMUs)* at facilities seeking RCRA permits to comply with corrective action requirements designed to identify and remediate hazardous releases.

32.5.2.1.3 Land Disposal Restrictions

Under the LDR program, no hazardous waste can be land disposed after a specified date unless it has been: (a) treated prior to disposal in accordance with standards established by EPA; (b) treated prior to disposal in accordance with a case-by-case treatability variance granted by EPA; (c) disposed of in a facility that will prevent migration of hazardous constituents for as long as the waste remains hazardous; (d) approved for a national capacity variance granted by EPA; or (e) approved for a case-by-case capacity variance granted by EPA. In addition, storage of untreated hazardous waste beyond one year is prohibited, unless the storage occurs in tanks, containers, or containment buildings solely to facilitate proper recovery, treatment, or disposal and certain labeling and record keeping requirements are met. While this program does not directly address groundwater monitoring or remediation, its treatment requirements and limitations on storage of hazardous waste have significantly affected hazardous waste management practices that could result in groundwater contamination.

32.5.2.1.4 Corrective Action Program

Prior to HSWA, *corrective action* requirements under RCRA were applicable only to onsite releases of hazardous waste to groundwater from surface impoundments, waste piles, land treatment units, and landfills that received wastes after July 26, 1982 (referred to as regulated units). HSWA expanded corrective action requirements to cover both onsite and offsite releases to all environmental media from solid waste management units at all facilities seeking RCRA permits, including interim status facilities. EPA incorporated the HSWA language regarding corrective action very quickly into its rules (existing 40 CFR 264.90–101). Even so, the Agency did not propose comprehensive implementing regulations until 1990 (55 FR 30798; July 27, 1990). As proposed, these regulations (referred to as the “1990 Subpart S

TABLE 32.3 Groundwater Protection Provisions for Hazardous Waste Treatment, Storage and Disposal Facilities with RCRA Permits

Provision	Description
40 CFR 264.90	Requires solid waste management units to comply with corrective action requirements in 40 CFR 264.101, but allows surface impoundments, waste piles, land treatment units or landfills that received hazardous waste after July 26, 1982 (referred to as “regulated units”) to comply with 40 CFR 264.91–264.100 in lieu of 264.101 and, under certain circumstances, allows the EPA Regional Administrator to replace all or part of the requirements in 40 CFR 264.91–264.100 with alternative requirements for groundwater monitoring and corrective action for releases to groundwater.
40 CFR 264.91	(1) Requires owner/operator to implement compliance monitoring under 264.99 whenever hazardous constituents under 264.93 from a regulated unit are detected at a compliance point under 264.95; (2) Requires owner/operator to institute corrective action under 264.100 whenever the groundwater protection standard under 264.92 is exceeded; (3) Requires owner/operator to institute corrective action under 264.100 whenever hazardous constituents under 264.93 from a regulated unit exceed concentration limits under 264.94 in groundwater between the compliance point under 264.95 and the downgradient facility property boundary; and (4) in all other cases, requires owner/operator to institute detection monitoring under 264.98.
40 CFR 264.92	Requires owner/operator to comply with permit conditions designed to ensure that hazardous constituents detected in the groundwater from a regulated unit do not exceed concentration limits in the uppermost aquifer underlying the waste management area beyond the point of compliance during the compliance period.
40 CFR 264.93	Indicates that facility permit will specify the hazardous constituents to which the groundwater protection standard of 264.92 applies.
40 CFR 264.95 and 264.96	Provides that facility permit will specify the point of compliance at which the groundwater protection standard applies and at which monitoring must be conducted, as well as the compliance period.
40 CFR 264.97	Requires groundwater monitoring programs to satisfy Parts 264.98–264.100.
40 CFR 264.98–264.100	Establishes owner/operator responsibilities to develop detection monitoring, compliance monitoring and corrective action programs, respectively, for regulated units.
40 CFR 264.101	Requires applicants for RCRA permits to establish corrective action programs, but clarifies that the requirement does not apply to remediation waste management sites unless they are part of a facility subject to a permit for treating, storing or disposing of hazardous wastes that are not remediation wastes.
40 CFR 264, Subpart G, Closure and Post Closure (264.110–264.120)	Requires closure of hazardous waste management units in a manner that, among other things, protects groundwater. Requires closure and post-closure plans that include groundwater monitoring corrective action provisions, if applicable.
40 CFR 264.193, 264.221, 264.251, 264.272, and 264.301	Establishes unit-specific design and operating requirements for tanks, surface impoundments, waste piles, land treatment units and landfills, respectively, to prevent migration of hazardous wastes or hazardous constituents into groundwater.
40 CFR 264.551	Establishes requirements for CAMUs that were approved before April 22, 2002, or for which substantially complete applications (or equivalents) were submitted on or before November 20, 2000 (referred to as “grandfathered CAMUs”). Among other things, requires that groundwater monitoring provisions and closure/post-closure provisions that meet certain criteria be specified in a permit or order applicable to the CAMU.
40 CFR 264.552	Establishes requirements for CAMUs that are not grandfathered CAMUs. Allows certain wastes to be prohibited from placement in a CAMU. Prohibits placing liquids in a CAMU, except under specified conditions. Requires that a permit or order applicable to the CAMU include the specification of design, operation, treatment and closure requirements. Specifies minimum design requirements for CAMUs that include installation of a composite liner, but allows approval of alternative design and operating practices where justified. Specifies minimum treatment requirements for wastes prior to placement into a CAMU, but allows the treatment levels to be adjusted based on certain factors. Establishes requirements for groundwater monitoring, corrective action, and closure/post-closure care.

TABLE 32.4 Groundwater Protection Requirements for Hazardous Waste Treatment, Storage and Disposal Facility with Interim Status

Provision	Description
40 CFR 265.90	Requires owners/operators of interim status surface impoundments, landfills, or land treatment facilities that are used to manage hazardous waste to implement, and operate during the active lives of the facilities as well as during post-closure, groundwater monitoring programs that meet the requirements of 265.91, and comply with 265.92–265.94. Under certain circumstances, alternative groundwater monitoring requirements may be developed and set out in an approved closure or post-closure plan or in another enforceable document.
40 CFR 265.91 and 265.92	Establish requirements for groundwater monitoring systems and sampling and analysis plans at interim status units, respectively.
40 CFR 265.93	Requires more comprehensive, alternative groundwater monitoring when an owner/operator of an interim status unit assumes (or knows) that monitoring of indicator parameters in accordance with 265.91 and 265.92 would show statistically significant increases (or decreases in the case of pH) when evaluated under 265.93(b).
40 CFR 265.94	Establishes record keeping and reporting requirements for groundwater monitoring programs at interim status units.
40 CFR 265, subpart G, Closure and Post-Closure (265.110–265.120)	Establishes requirements for closure and post-closure care of interim status units in a manner protective of human health and the environment, including post-closure groundwater monitoring. Mandates corrective action if migration of hazardous constituents is detected.

proposal") required the collection of enough information to support determinations of whether corrective action would be needed and if so, to support selection and implementation of appropriate corrective action. EPA received numerous comments on the 1990 Subpart S proposal. As a result of this and other pressures and changing priorities, EPA never finalized the proposal, except for a small section dealing with corrective action management units (CAMUs) and temporary units (TUs) (58 FR 8658; February 16, 1993), which are units limited to handling remediation wastes. Notwithstanding, the Agency used the 1990 Subpart S proposal as guidance for many years, and a number of states still have corrective action programs based largely on it. Under the guidance in the 1990 Subpart S proposal, groundwater monitoring frequently is required to satisfy the proposed information collection requirements, and selected corrective actions often include groundwater remediation, with accompanying monitoring to evaluate success.

On May 1, 1996 (61 FR 19432), the Agency published an advance notice of proposed rule making, which introduced the "Subpart S Initiative." The Subpart S Initiative was designed to identify and implement improvements to the protectiveness, responsiveness, speed, and efficiency of the corrective action program. The Agency also discussed corrective action implementation and the evolution of the program since 1990, and set forth its goals and strategy for the future of the corrective action program. The 1996 advance notice provided guidance on areas of the program not addressed by the 1990 proposal, and replaced the 1990 proposal as the primary guidance for much of the corrective action program.

In 1999, EPA officially withdrew the 1990 Subpart S proposal and announced that it had determined that such regulations were not necessary to carry out the Agency's statutory corrective action duties (64 FR 54604; October 7, 1999). Rather, a strategy of reliance on existing regulations (including those provisions of the Subpart S proposal already promulgated), supplemented by guidance and enhanced training was revealed. Accordingly, EPA subsequently issued dozens of guidance documents and interpretive memoranda intended to facilitate corrective action. Of particular interest with respect to groundwater protection is the *Handbook of Groundwater Protection and Cleanup Policies for RCRA Corrective Action*, which was published in October 2001 and updated in April 2004. Topics addressed in this guidance document include: Groundwater protection and cleanup strategy; short-term protection goals; intermediate performance goals; final cleanup goals; groundwater cleanup levels; point of compliance; cleanup timeframes; source control; groundwater use designations; institutional controls; monitored natural attenuation; technical impracticability; reinjection of contaminated groundwater; performance

TABLE 32.5 Groundwater Protection Standards for Non-Municipal Non-Hazardous Waste Disposal Units That Receive Conditionally Exempt Small Quantity Generator (CESQG) Waste

Provision	Description
40 CFR 257.8 and 257.9	Establish criteria for locating non-municipal non-hazardous waste disposal units that receive CESQG waste with respect to floodplains and wetlands.
40 CFR 257.22–257.28	Requires groundwater monitoring systems for contaminant detection at non-municipal non-hazardous waste disposal units that receive CESQG waste. Establishes requirements for groundwater monitoring system design and sampling and analysis. Requires assessment monitoring program, as well as corrective measures evaluation, selection, and implementation, if contaminants are detected.

monitoring; and, completing groundwater remedies. As a guidance document, the *Handbook* does not impose legally binding requirements, and it expressly acknowledges that its provisions may not apply to a particular situation based upon the specific circumstances of the corrective action facility. The *Handbook* also recognizes that most states have their own requirements and policies aimed at cleaning up contaminated groundwater, which may vary from federal guidance and policies. Hence, professionals involved in groundwater monitoring and cleanups under RCRA Corrective Action will need to determine on a case-specific basis the requirements of the state where each cleanup is located.

32.5.2.2 Nonhazardous Waste

32.5.2.2.1 Introduction

Guidelines for state and regional solid waste plans are located in 40 CFR 255–258. Parts 255 and 256 cover the process by which states are expected to develop and implement solid waste management plans. Part 257 sets criteria for determining which solid waste disposal facilities and practices pose a reasonable probability of adverse effects on health or the environment. Such facilities and practices must be appropriately addressed in the state's waste management plans. Part 257 also contains criteria for ensuring that non-municipal non-hazardous waste disposal units that receive *conditionally exempt small quantity generator* (CESQG) waste do not adversely affect human health and the environment. Specifically, Part 257, Subpart B establishes restrictions on locating this type of disposal unit in floodplains and wetlands and establishes requirements for groundwater monitoring and implementation of corrective action at these units. Table 32.5 summarizes provisions of Part 257, Subpart B pertinent to groundwater protection. Finally, Part 258 establishes minimum national standards for location, operation, design, closure, post-closure care, and financial assurance of MSWLF units. It also delineates minimum national standards for corrective action and groundwater monitoring at MSWLFs, and therefore, is of particular interest with respect to groundwater protection issues. Table 32.6 summarizes provisions of Part 258 pertinent to groundwater protection.

32.5.2.3 Underground Storage Tanks

32.5.2.3.1 Introduction

In 1984, HSWA added Subtitle I, “Regulation of Underground Storage Tanks,” to RCRA as a result of Congress’ realization that, of several million USTs containing petroleum or hazardous substances in the United States, tens of thousands were leaking, and more would leak in the future. As leaking USTs can cause fires or explosions and can contaminate drinking water and groundwater, Subtitle I was designed to address these issues. EPA’s regulations implementing Subtitle I are located in Parts 280, 281, and 282. Parts 281 and 282 deal primarily with the process by which EPA (1) approves State UST programs, and (2) documents such approvals. Part 280, however, sets technical standards and corrective action requirements for owners and operators of USTs, and is of most interest with respect to groundwater protection planning.

TABLE 32.6 Groundwater Protection Guidelines for Municipal Solid Waste Landfill Units

Provision	Description
40 CFR 258, subpart B, Location Restrictions (258.10–258.19)	Establishes criteria for locating MSWLFs with respect to airports, wetlands, floodplains, seismic zones, fault areas, and unstable areas.
40 CFR 258, subpart C, Operating Criteria (258.20–258.29)	Establishes criteria for operating MSWLFs, including among other things, procedures for excluding the receipt of hazardous waste, cover material requirements, requirements for runon/runoff control systems, limitations on surface water discharges, and restrictions on placing liquids in a landfill.
40 CFR 258, subpart D, Design Criteria (258.40)	Requires design of MSWLFs to assure that concentrations of specified constituents do not exceed certain limits in the uppermost aquifer, and requires composite liners and leachate collection systems.
40 CFR 258, subpart E, Corrective action and Groundwater Monitoring (258.50–258.58)	Requires groundwater monitoring systems for contaminant detection at MSWLFs. Establishes requirements for groundwater monitoring system design and sampling and analysis. Requires assessment monitoring program, as well as corrective measures evaluation, selection, and implementation if contaminants are detected.
40 CFR 258, subpart F, Closure and Post-closure Care (258.60–258.74)	Requires all MSWLF units to prepare a closure plan, install a final cover system designed to minimize infiltration and erosion, and conduct post-closure care for 30 years, including monitoring the groundwater and maintaining the groundwater monitoring system, if applicable.

TABLE 32.7 Groundwater Protection Standards for Underground Storage Tanks

Provision	Description
40 CFR 280.20	Establishes performance standards for new USTs to prevent releases due to structural failure, corrosion, or spills, and overfills.
40 CFR 280.21	Requires upgrading of existing USTs.
40 CFR 280.30–280.33	Requires operating practices to prevent releases caused by spills, overfills, corrosion, incompatibility of content and tank materials, or repairs.
40 CFR 280.40–280.43	Requires new and existing UST systems to be equipped with a method, or combination of methods, of release detection, and establishes standards for such methods. Standards for groundwater monitoring are included.
40 CFR 280.50–280.52	Requires reporting and investigation of suspected releases.
40 CFR 280.53	Requires reporting and cleanup of spills and overfills.
40 CFR 280.60–280.67	Requires response and corrective action for confirmed releases, including reporting, initial abatement, initial site characterization, free product removal, site investigations for soil and groundwater contamination, preparation of corrective action plan, and public notice.
40 CFR 280.70–280.72	Establishes requirements for temporary and permanent closure, including measurements for the detection of releases.

32.5.2.3.2 Groundwater Protection at USTs

USTs are identified and classified by Part 280 according to five criteria: (a) physical characteristics; (b) exclusions; (c) deferrals; (d) contents (i.e., hazardous substance or petroleum); and (e) age (i.e., new or existing). Together, these criteria define how each UST is regulated. Table 32.7 briefly summarizes the UST groundwater protection provisions of Part 280.

32.5.3 RCRA Summary

RCRA, as amended by HSWA, attempts to protect groundwater quality through various programs. States are required to regulate land disposal of nonhazardous solid waste (Subtitle D). EPA is required to regulate underground storage tanks containing petroleum and hazardous substances (Subtitle I), as well as the generation, transportation, treatment, disposal, and release remediation of hazardous waste (Subtitle C).

32.6 Comprehensive Environmental Response, Compensation, and Liability Act (CERCLA)

32.6.1 CERCLA Legislative Background

32.6.1.1 Historical Perspective

In the late 1970s, many US citizens were becoming distressed about the dangers of sites contaminated with hazardous wastes. Perhaps the most famous case was Love Canal. Residents of this small New York community were evacuated after authorities learned that hazardous waste buried over a 25-year period had contaminated the groundwater and surrounding soils. Responding to concerns over Love Canal and similar sites, Congress passed CERCLA, popularly known as the Superfund law, in 1980. (The term “Superfund” refers to the trust fund that Congress established to pay for CERCLA cleanup and enforcement activities.) Because it covers all environmental media (groundwater, surface water, air, and soil), the scope of the Superfund law is broader than most environmental laws. CERCLA also differs from other environmental laws in that it requires EPA to move beyond its traditional regulatory authority and to provide response authority to clean up hazardous waste sites. In 1986, Congress passed the Superfund Amendments and Reauthorization Act (SARA), which reauthorized CERCLA for five years and expanded and strengthened many of its provisions. In 1990, Congress reauthorized CERCLA through 1994 but made no other significant changes. Since then, Congress has introduced several bills to reauthorize the program, but none has passed. In the meantime, many of the reauthorization proposals have been addressed through administrative reforms or court decisions, and annual appropriations have extended the program’s authority. In 2002, the President signed the Small Business Liability Relief and Brownfields Revitalization Act, which facilitated cleanup and revitalization of underutilized, contaminated sites.

32.6.1.2 Goals and Objectives

A primary goal of Superfund is to ensure protection of human health and the environment by preventing further contamination and cleaning up contamination at sites to levels called for in existing standards, or, where existing standards do not exist, to acceptable risk levels. Wherever technically practicable, groundwater is to be cleaned up to levels that correspond to those in existing standards, such as MCLs. Another goal is to maximize the involvement of potentially responsible parties in conducting site cleanups.

In 1989, EPA conducted the first in a series of evaluations to identify ways to improve the program. Goals resulting from this first evaluation included controlling acute threats immediately; monitoring sites over the long term; developing and using new technologies; and improving efficiency. Between 1993 and 1995, EPA initiated three rounds of administrative reforms designed to meet goals that included enhancing cleanup effectiveness and protectiveness and consistent program implementation. Goals articulated in the most recent administrative review, completed in 2004, included increasing the pace of site cleanups and controlling the risks to human health and the environment by remediating and restoring contaminated sites or properties to appropriate levels. Over the past few years, the goals and performance measures in Superfund have shifted from output-oriented measures (e.g., the number of cleanup completions) to results-oriented measures (e.g., controlling the migration of contaminated groundwater). There are two specific results-oriented goals for groundwater. The first is to control, by 2008, all identified unacceptable human exposures from site contamination to health-based levels or below for current land and groundwater use conditions at 84% (1,259) of the 1,494 Superfund human exposure sites (as of FY 2002). The second is to control, by 2008, the migration of contaminated groundwater through engineered remedies or natural processes at 65% (832) of the 1,275 Superfund groundwater exposure sites (as of FY 2002).

32.6.1.3 General Provisions

CERCLA provides Federal funding and enforcement authority to clean up hazardous waste sites that were created in the past. It also provides for response to hazardous spills and requires businesses to report releases of hazardous substances. Pertinent portions of this broad law that pertain to groundwater are summarized below.

32.6.1.3.1 §104 — Response Authorities

§104 authorizes EPA to respond when an actual or threatened *release* of a *hazardous substance* may endanger public health or welfare. CERCLA provides for two types of response actions. Removal actions are short-term responses that stabilize or clean up a hazardous site that poses a threat to public health or the environment. Removal actions include capping of contaminated soils to reduce migration of hazardous substances into groundwater and providing temporary alternate sources of drinking water to local residents. The second type of response action is a remedial action. Remedial actions are taken instead of or in addition to removal actions to minimize the migration of hazardous substances so that they do not cause substantial danger to present or future public health or the environment. Remedial actions include the study, design, and construction of longer term actions aimed at permanent remedy. Typical actions include constructing underground walls to control the movement of groundwater and pumping and treating of contaminated groundwater. Both removal and remedial actions must be conducted in accordance with the *National Contingency Plan (NCP)*.

32.6.1.3.2 §105 — National Contingency Plan

CERCLA requires EPA to revise and republish the NCP for the removal of oil and hazardous substances, originally prepared and published pursuant to §311 of the CWA. CERCLA requires that the plan reflect the responsibilities created by CERCLA, including the establishment of procedures for responding to releases of hazardous substances. The NCP is published as an EPA regulation in 40 CFR 300. It specifies procedures for identifying, removing, and remediating releases of hazardous substances. §105 also provides for the Hazard Ranking System (HRS). The HRS is a scoring system used to evaluate potential relative risks to human health and the environment from releases or threatened releases of hazardous substances. EPA uses the HRS to determine if a site should be on the National Priorities List (NPL). The NPL is EPA's list of the most serious hazardous waste sites identified for possible long-term remedial response using money from the Trust Fund. The NCP and the HRS are described further under CERCLA Regulations.

32.6.1.3.3 §121 — Cleanup Standards

§121 provides principles governing the cleanup of hazardous substances. §121(a) states that remedial actions are to be selected in accordance with the NCP. §121(b) states that remedies in which treatment permanently and significantly reduces the volume, toxicity, or mobility of the hazardous substances are preferred over remedies that do not provide such treatment. CERCLA requires that remedial actions are to be protective of human health and the environment, cost-effective, and utilize permanent solutions and alternative treatment technologies to the maximum extent practicable. In determining the degree of cleanup, CERCLA provides no standards of its own. However, §121(d)(2) requires site cleanups to attain standards from other Federal and State environmental programs that are applicable or relevant and appropriate requirements (ARARs). §121(d)(2)(A)(ii) states that such remedial actions shall require a level or standard of control which at least attains MCLGs established under the SDWA and water quality criteria established under the CWA, where such goals or criteria are relevant and appropriate under the circumstances of the release or threatened release. §121(d)(2)(B) states that in determining whether or not any water quality criteria under the CWA is relevant and appropriate under the circumstances for the release or threatened release, the designated or potential use of the groundwater must be considered. §121(d)(4) provides limited waivers for achieving ARARs.

§121(e) provides that CERCLA cleanups require no Federal, State, or local permits for the portion of any removal or remedial action conducted entirely onsite, where such remedial action is selected and executed in compliance with the rest of §121. §121(f) provides for "substantial and meaningful" State involvement in initiation, development, and selection of remedial actions to be undertaken in that state.

§107(a) provides that owners, operators, transporters, or generators potentially responsible for contamination problems at a Superfund site are liable for natural resource damages (NRD) resulting from a release of a hazardous substance. NRD are monetary compensations for injury to, destruction of, or loss of *natural resources*. Natural resources are land, fish, wildlife, biota, air, water, groundwater, drinking water supplies, and other such resources belonging to, or otherwise controlled by the United States; a state,

local, or foreign government; or Indian tribe. NRD are distinct from response costs, which are the costs of actions taken under the NCP to remove threats to human health and the environment caused by hazardous substance releases. §301 (c) of CERCLA requires the Department of Interior to promulgate regulations for assessing NRD resulting from hazardous substance releases. The regulations include two types of assessment procedures. "Type A" procedures are standard procedures for simplified assessments, and "Type B" procedures are "alternative protocols for conducting assessments in individual cases."

Unlike many environmental regulatory programs, CERCLA does not provide for State delegation. EPA and its ten Federal regions administer site cleanups financed by the Superfund trust fund. Before the Federal government initiates a fund-financed remedial action, the state in which the site is located must assure that it will provide for future maintenance of response actions and pay for 10% of all costs of remedial action, or 50% if the state had operated the facility undergoing cleanup. Many states have comparable programs that provide for the cleanup of contaminated sites that are not on the Federal NPL. In all cases, concerns of the state and local community are considered in cleanup decisions.

32.6.1.4 Issues and Outlook

When CERCLA was enacted 25 years ago, Congress expected that only a few hundred sites would require cleanup and that the program would require relatively modest funding. As of June 2005, there were 1242 sites on the NPL and 64 sites proposed to be added; 298 sites have been dropped. Over the past 10 years, new sites have been added to the NPL at the rate of about 28 per year. At sites with completed RODs, about 85% have required groundwater remediation. Assuming these trends continue, there will be ample opportunity to address groundwater contamination at Superfund sites well into the future. The Superfund program has many critics. Although Congress has struggled to modify and reauthorize CERCLA since 1991, no comprehensive bills have been passed. Key issues pertain to the liability system and to the remedy selection process. CERCLA's provision that parties are liable for cleanup costs for which they may not have been the cause can result in excessive litigation and other transaction costs that divert money from actual cleanup. Regarding remedy selection, many believe that the prescriptive, cumbersome, and confusing nature of the process has delayed cleanups. Minor amendments have been made to the law to help address some of the liability issues, and EPA has undertaken administrative reforms to respond to the remedy selection issues. In 2005, comprehensive reauthorization of CERCLA is not as high a priority as it was in the 1990s.

32.6.2 CERCLA Regulations

Superfund regulations are contained in the NCP (40 CFR 300) and in the Natural Resource Damage Assessment (NRDA) regulations (43 CFR 11).

32.6.2.1 National Contingency Plan

The NCP contains procedures for responding to discharges of oil and hazardous substances. Subpart E describes methods and criteria for assessing sites and determining the appropriate response to hazardous substance releases. Site assessment determines whether a site is eligible for cleanup funding and includes the following activities:

- Site Discovery — Identifies hazardous substance releases.
- Preliminary Assessment — Evaluates existing data to determine need for further action.
- Site Investigation — Assesses onsite conditions to determine need for HRS scoring.
- HRS scoring — Applies a mathematical approach to assess the relative risk of a site and determine its eligibility for the NPL. The HRS considers four exposure pathways: groundwater migration, surface water migration, soil exposure, and air migration. The groundwater migration pathway is based on three factor categories: likelihood of release, waste characteristics, and targets. The HRS evaluates these pathways and then combines them to produce an overall score for the site. Appendix A to Part 300 provides details on HRS scoring calculations; §3.0 of that appendix pertains to groundwater pathway calculations.

- NPL listing — Determines sites eligible for Superfund-financed remedial action. If the unitless score resulting from the HRS process is at least 28.5, the site is placed on the NPL. Once a site is placed on the NPL, the remedial process begins. The NCP provides general expectations to help identify and implement appropriate remedial actions. Some of these pertain to groundwater. For example, 40 CFR 300.420(a)(iii) expects EPA to use institutional controls such as water use restrictions to supplement engineering controls in limiting exposure to hazardous substances. However, institutional controls cannot substitute for active response measures (e.g., restoration of groundwater to its beneficial uses) as the sole remedy, unless such active measures are determined not to be practicable. EPA expects to return usable groundwater to its beneficial uses whenever practicable within a time frame that is reasonable; when restoration of groundwater to beneficial uses is not practicable, EPA expects to prevent further migration of the plume, prevent exposure to the contaminated groundwater, and evaluate further risk reduction.

The NCP prescribes the following specific activities for the remedial action process:

- Remedial investigation (RI) — Determines the nature and extent of the problem presented by a release. The RI includes field investigations to assess the characteristics of groundwater and other environmental media. A site-specific risk assessment is conducted to characterize the current and potential threats to human health and environment that may be posed by contaminants migrating to groundwater, surface water, and other pathways. The results of the baseline risk assessment help establish acceptable exposure levels for developing remedial alternatives in the feasibility study.
- Feasibility study — Develops and evaluates options for remedial action. Remediation goals are set to establish acceptable exposure levels that are protective of human health and the environment. For groundwater, MCLGs established under the SDWA that are set at levels above zero are to be attained by remedial actions for ground or surface waters that are current or potential sources of drinking water, where the MCLGs are relevant and appropriate. If an MCLG is determined not to be relevant and appropriate, the corresponding MCL is to be attained where relevant and appropriate. Where the MCLG for a contaminant has been set at zero, the MCL is to be attained by remedial actions for ground or surface waters that are current or potential sources of drinking water, where the MCL is relevant and appropriate.
- Remedy selection — Identifies the likely remedial alternative and provides for public comment. The Record of Decision (ROD) is the official report that documents the background information on the site and describes the chosen remedy and the justification for that remedy.
- Remedial design — Prepares technical plans for implementing the chosen remedy.
- Remedial action — Implements the remedial design.
- Operation and Maintenance — Ensures that the cleanup methods are working properly after a response action occurs. A groundwater restoration activity will be considered administratively complete when: (a) actions restore ground or surface water quality to a level that assures protection of human health and the environment; (b) actions restore groundwater or surface water to such a point that reductions in contaminant concentrations are no longer significant; or (c) ten years have elapsed, whichever is earliest. Ground or surface water measures initiated for the primary purpose of providing a drinking water supply and not for the purpose of restoring groundwater do not constitute treatment.

32.6.2.2 Natural Resource Damage Assessments

CERCLA authorizes the United States, individual States, Indian Tribes and foreign governments to act on behalf of the public as Natural Resource Trustees for natural resources under their respective trusteeship. Natural Resource Damage Assessments (NRDA) regulations (43 CFR 11) define the process whereby a natural resource trustee may pursue compensation on behalf of the public for injury to natural resources resulting from releases of hazardous substances. Individual NRDAs follow specific procedures for determining whether an injury has occurred to groundwater (and other) resources and for determining if the exposure pathway was through groundwater or another medium. Procedures also exist for testing

and sampling to determine injury and for quantifying baseline conditions and effects of the release on groundwater (and other) resources to determine the appropriate amount of compensation.

32.6.2.3 CERCLA Guidance

Guidance is designed to implement national policy, but it does not substitute for statutes or regulations, and, unlike regulations, does not require a public notice and comment process. Guidance provides direction but does not impose legally binding requirements on EPA, states, or the regulated community. EPA has issued guidance relevant to groundwater in the areas of monitored natural attenuation, presumptive remedies, vapor intrusion, and technical impracticability.

32.6.2.4 Monitored Natural Attenuation

Natural attenuation processes include natural physical, chemical, and biological process that act to reduce the mass, toxicity, mobility, volume, or concentration of contaminants in groundwater or soil. These processes include biodegradation, dispersion, dilution, sorption, volatilization, radioactive decay, and chemical or biological stabilization, transformation, or destruction of contaminants. As the understanding of these processes has increased, so too has the interest in using Monitored Natural Attenuation (MNA) as a component of groundwater treatment strategies. MNA refers to the reliance on natural attenuation processes that are part of a controlled and monitored cleanup approach to achieve site-specific remediation objectives in a reasonable time frame. Potential advantages of using MNA compared with active remediation include reduced potential for cross-media transfer of contaminants, reduced risk of human exposure to contaminants, reduced intrusion from fewer surface structures, and potentially lower overall remediation costs. Potential disadvantages can include increased time to achieve remediation objectives, toxicity and or mobility products that may be more hazardous than the parent compound, and the need for institutional controls to ensure long-term protection. In 1999, EPA issued guidance entitled *Use of Monitored Natural Attenuation at Superfund, RCRA Corrective Action, and Underground Storage Tank Sites*, which clarified the Agency's policy on the use of MNA for the cleanup of contaminated groundwater and soil. The guidance states that EPA considers MNA to be a means of addressing contamination under a limited set of site circumstances where its use meets the applicable statutory and regulatory requirements. It also said that a decision to implement MNA should include comprehensive site characterization, risk assessment where appropriate, and measures to control sources.

32.6.2.5 Presumptive Remedies

Since the 1980s, EPA has found that certain categories of sites have similar characteristics. By using information acquired from evaluating and cleaning up such sites, EPA has developed "presumptive remedies" to accelerate future cleanups of these sites. Presumptive remedies are preferred technologies for common categories of sites that are developed on the basis of historical patterns of remedy selection and EPA's scientific and engineering evaluation of performance data on technology implementation. In October 1996, EPA issued guidance entitled, *Presumptive Response Strategy and Ex-Situ Treatment Technologies for Contaminated Ground Water at CERCLA Sites, Final Guidance*. This guidance outlines a "phased approach" for addressing contaminated groundwater, which integrates the site characterization, early actions, remedy selection, design, implementation, and performance-monitoring phases. The strategy suggests ways to optimize the selected remedy to improve effectiveness and decrease costs and time. It identifies and describes presumptive technologies for treatment of extracted groundwater. These technologies include air stripping, granular/activated carbon, chemical/ultraviolet oxidation, aerobic biological reactors, chemical precipitation, ion exchange/adsorption, electrochemical methods, and aeration of background metals. Since issuing the guidance, EPA has been investigating additional ways to streamline groundwater response strategies. For example, for removing volatile organic compounds from groundwater, the Agency could modify its current practice of developing a unique design for each site, to developing standard designs that could be adapted to a particular site, water chemistry, and suite of chemicals.

32.6.2.6 Vapor Intrusion

Volatile chemicals in contaminated groundwater or soils can emit vapors that may migrate through subsurface soils. Vapor intrusion is the upward migration and entry of contaminants to indoor air from underlying groundwater or soil contamination sources. It is commonly associated with petroleum products, chlorinated solvents, and other volatile organic chemicals in groundwater and soil. High water tables and coarsely grained soils can increase the likelihood of intrusion. Vapor intrusion is of concern because in extreme cases, the accumulated vapors may cause safety hazards (e.g., explosion), acute health effects, or aesthetic problems (e.g., odors). Vapor intrusion has direct links to CERCLA cleanups. The potential for subsurface migration of vapors via the indoor air pathway, particularly at cleaned up sites that are being redeveloped, is a growing concern, and investigations are becoming increasingly common. In 2002, EPA issued *Draft Guidance for Evaluating the Vapor Intrusion to Indoor Air Pathway from Groundwater and Soils*, and many states are issuing their own guidance on the subject. EPA's guidance provides a tool for conducting a screening evaluation to determine whether the vapor intrusion exposure pathway is complete (i.e., that humans are exposed to vapors originating from site contamination) and, if so, whether it poses an unacceptable risk to human health. At the time of this writing, EPA is considering revising the guidance to, among other things, more accurately address the actual risks associated with vapor intrusion.

32.6.2.7 Technical Impracticability

The remediation of the most highly contaminated groundwater sites, such as those with dense nonaqueous phase liquids, is technically challenging. While the goal of groundwater cleanup at Superfund sites is to restore contaminated groundwater to ARAR-based levels where technically practicable, EPA studies indicate that complex site hydrogeology, contaminant characteristics, and other factors make complete restoration of some contaminated groundwaters technically impracticable with available technologies. Both CERCLA and the NCP allow for the waiving of ARAR-based levels for technical impracticability from an engineering perspective. In 1993, EPA issued *Guidance for Evaluating the Technical Impracticability of Ground-water Restoration*, which explains how EPA determines whether groundwater restoration at Superfund sites is technically impracticable, and, if so, the alternative measures that must be undertaken to ensure that a final remedy is protective. The guidance covers the types of technical data needed to determine technical impracticability, decision criteria, and alternative remedial strategies.

32.6.3 CERCLA Summary

In summary, CERCLA provides for the cleanup of groundwater and other media at old or abandoned sites contaminated by releases of hazardous substances. The law contains no cleanup standards of its own, relying instead on requirements from other Federal and State environmental programs such as the CWA and the SDWA. Many Superfund critics blame the law and its regulations for high cleanup costs and slow cleanup progress. Congress has been trying to reauthorize and improve the law since 1991.

Over the years, the scope of the Superfund program has expanded. Response actions aimed at rapid site cleanup to eliminate immediate threats to human health and the environment have expanded to include responses to terrorist activities and other non-waste site cleanup emergencies. In the early years, remedial actions focused on drum disposal sites and landfills; today they include sediments in harbors and mining sites. The universe of Superfund sites has also expanded, with many sites requiring multiple cleanup remedies. Many sites have moved from the assessment phase to the cleanup phase, and today, many are entering the post-construction completion phase, which often requires groundwater monitoring and institutional controls to ensure that the cleanup remains effective.

32.7 Other Laws

Two other laws are of interest in the area of groundwater regulation:

- The Surface Mining Control and Reclamation Act of 1977 (SMCRA), Public Law 95-87, 30 U.S.C. §1201 *et seq.*

- Federal Insecticide, Fungicide, and Rodenticide Act (FIFRA) of 1947, Public Law 80-104, 7 U.S.C. §136 *et seq.*

32.7.1 SMCRA

SMCRA provides general regulatory requirements governing surface coal mining and the surface impacts of underground coal mining. SMCRA requires the Office of Surface Mining (OSM) in the US Department of Interior to promulgate a federal program that controls surface coal mining and reclamation operations, and serves as a model for authorized state programs. The minimum general environmental protection performance standards established by the law include those directed at minimizing disturbances to the quality and quantity of water in groundwater systems. OSM has implemented general standards governing groundwater protection performance (30 CFR §715.17(h)) and water rights and replacement (30 CFR §715.17(i)). The groundwater protection performance standard requires that the recharge capacity of reclaimed lands must be restored to approximate pre-mining capacity. In addition, backfilled material must be placed on reclaimed land in a fashion that minimizes adverse impacts on groundwater flow and quality. Finally, groundwater monitoring must be conducted to determine the effects of mining and reclamation operations relative to the recharge capacity of reclaimed land as well as the water quantity and quality in groundwater systems at the mine area and in associated offsite areas. OSM's water rights and replacement standard is triggered when mining activities affect the underground water supply source of an owner with interest in real property who uses the water for domestic, agricultural, industrial, or other legitimate purposes. In the overall context of SMCRA implementation, the absence of historic data allowing interested parties to systematically track groundwater quality and quantity impacts from mining and reclamation activities has been identified as one of the principal issues.

32.7.2 FIFRA

FIFRA requires EPA to regulate the sale and use of pesticides in the United States through registration and labeling requirements. EPA is given the mandate to restrict the use of pesticides as necessary for the purpose of preventing unreasonable adverse effects on humans and the environment. FIFRA does not specifically mention groundwater. However, FIFRA does address concerns related to user risk, risk-benefit relationships, and consumption of food and drinking water.

Glossary

The following definitions are provided for the readers' better understanding and are not necessarily the same as the applicable legal definitions.

Absolute dominion The right of the landowner to capture the groundwater beneath the surface land. The doctrine is also known as the English rule or rule of capture.

ARARs Applicable or relevant and appropriate requirements include any standard, requirement, criteria, or limitation under any Federal environmental law, (e.g., CWA, SDWA, TSCA (Toxic Substances and Control Act), RCRA) and any promulgated standard, requirement, criteria, or limitation under a State environmental or facility siting law that is more stringent than any Federal standard requirement, criteria, or limitation.

CAMUs Corrective action management units are land areas, located within a RCRA-regulated facility, that have been designated by EPA or the authorized state for the purpose of managing remediation wastes generated from corrective action activities at that facility.

CESQG A generator is a conditionally exempt small quantity generator in a calendar month if he generates no more than 100 kg of hazardous waste in that month.

Conjunctive management The coordinated and integrated management of surface and groundwater resources.

Corrective action A corrective action is any action taken to remedy a hazardous release.

Correlative use The right of landowners to withdraw groundwater from a common pool according to their respective proportions of overlying land. The doctrine is also known as the common resource rule.

CSGWPP The Comprehensive State Ground Water Protection Program is an EPA program that attempts to focus State resource on prevention of contamination and remediation based on the relative use and value of groundwater.

Hazardous substance A hazardous substance includes but is not limited to any substance designated under the CWA, the Clean Air Act (CAA), TSCA, or any "hazardous waste" under RCRA.

Hazardous waste A hazardous waste is defined as any solid waste that is not excluded from RCRA Subtitle C regulation, and that exhibits a hazardous characteristic, is listed as hazardous waste under RCRA regulations, or is mixed with or derived from hazardous waste.

HRS The Hazard Ranking System is a scoring system used to evaluate potential relative risks to human health and the environment from releases or threatened releases of hazardous substances.

Injection well Practically speaking, an injection well is a hole in the ground that is deeper than it is wide and that receives wastes or other fluids. Under the UIC program, drain fields, leach fields, drywells, and sumps are considered to be injection wells even though traditionally they have not been thought of as wells.

LDRs Land disposal restrictions treatment standards are EPA regulations adopted under RCRA specifying concentration levels or methods of treatment that substantially diminish the toxicity or likelihood of migration of hazardous constituents from hazardous wastes placed into or onto the land.

MCLs Maximum contaminant levels are enforceable standards for drinking water quality. MCLs may be less stringent than MCLGs as they consider not only human health protection but also the availability of economically achievable treatment technology.

MCLGs Maximum contaminant level goals are nonenforceable targets set to protect human health.

MSWLF A municipal solid waste landfill is a discrete, publicly or privately owned area of land that receives household waste (e.g., garbage, trash and sanitary waste from homes, hotels, campgrounds, etc.), and may receive commercial and industrial solid waste, nonhazardous sludge, or hazardous waste from generators exempt from RCRA Subtitle C because they produce such small quantities.

NCP Officially known as the National Oil and Hazardous Substances Pollution Contingency Plan (40 CFR 300), the NCP outlines the responsibilities and authorities for responding to releases into the environment of hazardous substances and other pollutants and contaminants under the statutory authority of CERCLA and §311 of the CWA.

NPDES The National Pollutant Discharge Elimination System is a Federal program that issues wastewater discharge permits to all point source discharges to surface water.

NPL The National Priorities List is EPA's list of the most serious hazardous waste sites identified for possible long-term remedial response using money from the Superfund Trust Fund.

NRD Natural Resource Damages are monetary compensations for injury to, destruction of, or loss of natural resources.

NRDA Natural Resource Damage Assessments are processes whereby a trustee assesses the damages to natural resources resulting from a release of a hazardous substance under the Superfund law.

Natural resources Natural resources are the land, fish, wildlife, biota, air, water, groundwater, drinking water supplies, and other such resources controlled by Federal, State, local, or foreign governments, or Indian tribes.

Nonpoint sources Nonpoint sources of pollution are those that do not enter surface or groundwater through discrete pipes or conveyances. They include agricultural runoff, atmospheric deposition, groundwater inflow to stream beds, acid mine drainage, and some types of stormwater runoff.

POTW A publicly owned treatment works is a wastewater treatment plant owned by a state or local government.

Prescription Acquisition of a personal right to use water by reason of continuous usage.

Prior appropriation The landowner's right to use available water based on dates of appropriation pursuant to a government-administered permit system.

RCRA base program The RCRA base program is defined as all RCRA Subtitle C requirements imposed prior to enactment of the Hazardous and Solid Waste Amendments of 1984.

RI A Remedial Investigation is a process used under the National Contingency Plan to determine the nature and extent of the problem(s) presented by a release of a hazardous substance

ROD A Record of Decision is a public document that explains which cleanup alternative(s) will be used at a National Priorities List site.

Reasonable use The right of the landowner to capture the groundwater beneath the surface land, subject to making reasonable use of the resource. The doctrine is also known as the American rule.

Remediation wastes Remediation wastes are wastes managed for the purpose of implementing corrective action.

Restatement of Law A series of volumes authored by the American Law Institute that describe the law in a general area; for example, Restatement of the Law of Torts.

Release A release is any spilling, leaking, pumping, pouring, emitting, emptying, discharging, injecting, escaping, leaching, dumping, or disposing into the environment. Excluded from the definition of release are workplace exposures covered by OSHA, vehicular engine exhausts, radioactive contamination covered by other statutes, and the normal application of fertilizer.

Sole source aquifer A sole source aquifer is a sole or principal drinking water source for an area.

Solid waste A solid waste is any material that has been discarded by being abandoned, recycled, or considered inherently waste-like, and that is not excluded from RCRA Subtitle C requirements.

Source water protection program A program that protects drinking water supplies through evaluation of the activities that might cause undesirable pollutants to enter drinking water source waters and the system's susceptibility to contamination and implementation of controls.

SWMU A solid waste management unit is any discernable unit at which solid wastes have been placed at any time, irrespective of whether the unit was intended for the management of solid or hazardous wastes.

Takings The Fifth Amendment to the US Constitution prohibits the government from taking private property for public use without paying the owner just compensation. Regulatory takings issues involve the allegations by owners that the government's regulations restrict the use of their property and cause reductions in property values.

TUs Temporary units are tanks or container storage units, located within the boundaries of a RCRA-regulated facility, that have been designated by EPA or the authorized state for use in treating and storing remediation wastes generated from corrective action activities at that facility, and that will operate for no longer than one year.

IUC The Underground Injection Control program is a Federal program that regulates injection of fluids underground through injection wells.

USDW An underground source of drinking water is an aquifer that presently serves or could reasonably be expected to serve as drinking water supply.

USTs Under RCRA, USTs are underground storage tanks containing petroleum and hazardous substances.

Watershed management approach The watershed management approach is a framework for environmental management that focuses efforts to address the highest priority problems within hydrologically defined geographic areas, taking into consideration both ground and surface water flow.

Wellhead protection programs Wellhead protection programs are designed to impose controls on activities located in surface and subsurface areas surrounding a well that supplies a public water system, through which contaminants are reasonably likely to move toward and reach the well.

Further Information

Groundwater Allocation and Management

Henley, T., Arizona Department of Water Resources, *Groundwater Resources of the Lower Colorado Region*, in University of Colorado School of Law, Natural Resources Law Center, 2004 Summer Conference, Groundwater in the West, available at [http://www.colorado.edu/law/centers/nrlc/publications/Groundwater Conference/papers/Session5/henley.pdf](http://www.colorado.edu/law/centers/nrlc/publications/Groundwater%20Conference/papers/Session5/henley.pdf) (last visited September 13, 2005). Offers an excellent explanation of the need for regulation from a regulator's perspective.

Water Systems Council, *Who Owns the Water: A Summary of Existing Water Rights Laws*, Washington, D.C., 2002. Provides the source for the state-survey numbers displayed in Table 32.1.

Ferry, S., *Environmental Law*, Aspen Publishers, New York, NY, 2004. Discusses the rights associated with use of water resources in the surface and ground portions of the hydrological cycle.

Ashley, J.S. and Smith, J.A., *Western Groundwater Wars*, Forum for Applied Research and Public Policy, Knoxville, TN, pp. 33–39, spring 2001. Summarizes major groundwater rights doctrines, legal concepts, and future perspectives.

Government Accountability Office, *Freshwater Supply — States' Views of How Federal Agencies Could Help Them Meet the Challenges of Expected Shortfalls*, GAO-03-514, Washington, D.C., pp. 19–21, July 2003. Summarizes federal authorities to manage water resources.

Williams, S.M., *Overview of Indian Water Rights*, Water Resources Update, Universities Council on Water Resources, issue number 107, Carbondale, IL, pp. 6–8, spring 1997, available at http://www.ucowr.siu.edu/uploads/pdf/V107_A1.pdf (last visited May 7, 2005). Describes reservation, quantification and scope of use, and administration of Indian water rights.

Richardson, J.J., *Legal Impediments to Utilizing Groundwater as a Municipal Water Supply Source in Karst Terrain in the United States*, Environmental Geology, vol. 42, Springer, Berlin, pp. 532–538, 2002. Discusses groundwater use and taking issues.

CWA and SDWA

Readers are encouraged to visit the EPA Office of Water website at <http://www.epa.gov/water>. The CWA topics of interest are primarily located in the Office of Wetlands, Oceans, and Watersheds website at <http://www.epa.gov/owow>. The SDWA topics of interest are primarily located in the Office of Ground Water and Drinking Water website at <http://www.epa.gov/safewater/index.html>. The website offers a great deal of explanatory information and links to relevant reports.

RCRA and CERCLA

Readers are encouraged to visit the following EPA websites, which contain links to or searchable databases containing RCRA policy and guidance documents:

- <http://www.epa.gov/swerrims> (EPA's Office of Solid Waste and Emergency Response website).
- <http://www.epa.gov/rcreonline/> (EPA's RCRA Online database, which enables users to locate RCRA policy and guidance documents on topics including cleanup and groundwater monitoring).
- <http://www.epa.gov/epaoswer/hazwaste/ca/index.htm> (EPA's RCRA Corrective Action website).
- <http://www.epa.gov/epaoswer/osw/hazwaste.htm> (EPA's RCRA Hazardous Waste website).
- <http://www.epa.gov/epaoswer/osw/non-hw.htm> (EPA's RCRA Non-Hazardous Waste website).

Information on Superfund and links to laws, policies, and guidance and other documents are on EPA's Superfund website, <http://www.epa.gov/superfund>. EPA's CLU-IN website (<http://www.clu-in.org/>) is a source of on-line and downloadable information on site remediation and field analytical technologies, including those pertaining to ground water. The Federal Remediation Technologies Roundtable website (<http://www.frtr.gov>) was developed by a consortium of EPA and six other federal agencies to improve

the availability of innovative technologies and contains information on groundwater characterization and cleanup technologies.

The following documents concerning groundwater remediation and groundwater investigations at remediation sites are referenced in the text of this chapter.

Groundwater Remediation

EPA, *Use of Monitored Natural Attenuation at Superfund, RCRA Corrective Action, and Underground Storage Tank Sites*, OSWER Directive 9200.4-17P, April 21, 1999.

EPA, *Presumptive Response Strategy and Ex-Situ Treatment Technologies for Contaminated Ground Water at CERCLA Sites, Final Guidance*, OSWER Directive, EPA 9283.1-12, EPA/540/R-96/023, October 1996.

EPA, *Guidance for Evaluating the Technical Impracticability of Ground-Water Restoration*, OSWER Directive No. 9234.2-25, October, 1993.

Groundwater Investigation at Remediation Sites

EPA, *Handbook of Groundwater Protection and Cleanup Policies for RCRA Corrective Action*, EPA530-R-04-030, April 2004.

EPA, *Draft Guidance for Evaluating the Vapor Intrusion to Indoor Air Pathway from Groundwater and Soils*, November 2002, available at <http://www.epa.gov/correctiveaction/eis/vapor/complete.pdf>.

FIFRA

Lee, M., *Summary of Environmental Laws Administered by the EPA*, Congressional Research Service Report RL30022, January 12, 1999, available at <http://www.ncseonline.org/nle/crsreports/briefingbooks/laws/index.cfm> (last visited May 7, 2005). Provides a comprehensive overview of FIFRA.

33

Landfills

33.1	Introduction.....	33-1
33.2	Regulatory Considerations and Landfill Siting	33-2
	Subtitle D • Clean Water Act • Clean Air Act • State and Local Regulations • Landfill Siting • Hydrogeologic and Geotechnical Considerations	
33.3	Site Master Plan	33-5
33.4	Landfill Configuration.....	33-6
33.5	Conceptual Design.....	33-7
	Landfill Footprint • Subbase Grading • Sedimentation/Detention Basins • Borrow Areas • Final Cover Grading • Cell Layout and Filling Sequence • Other Factors	
33.6	Detailed Design	33-12
	Mechanical Properties of Waste • Liner System • Final Cover System • Berms • Storm Water Management System • Leachate Collection and Removal System • Gas Management System • Bioreactor Landfills • Alternative Final Covers	
33.7	Liner and Cover System Materials	33-22
	Soil Materials • Geosynthetic Materials	
33.8	Environmental Monitoring	33-25
	Groundwater Monitoring • Gas Monitoring	
33.9	Performance of Landfill Elements	33-26
	Geotechnical and Geosynthetics Analysis • Storm Water Management • Leachate Collection and Removal System	
33.10	Summary and Conclusions	33-27
	Glossary	33-27
	References	33-30
	Further Information	33-31

Beth A. Keister and
Pedro C. Repetto
URS Corporation

33.1 Introduction

Landfills may be classified in accordance with several criteria. The most common criteria used to classify landfills are type of waste, type of liner system, and geometric configuration. Based on current federal regulations, landfills are classified with respect to the type of waste. They contain: municipal solid waste, hazardous waste, and other types of solid waste. Federal regulations regarding these three types of landfills are published in Section 258, Section 264, and Section 257, respectively, of the Code of Federal Regulations. State regulations are, however, frequently more stringent than federal regulations. State regulations

frequently include specific regulations for landfills where other types of waste will be disposed, such as industrial waste, construction demolition debris, and residual waste.

Development of a landfill project requires proper attention to the diverse and, sometimes, conflicting array of regulatory, physical, and technical considerations relevant to siting and design. Two initial and essential steps for the development of a landfill project are site selection and preparation of a Site Master Plan. A thorough understanding of the physical setting of the site and the regulatory requirements for the design and permitting are of paramount importance for the successful development of a project.

Landfill designers must keep in mind that environmental regulations are intended to protect human health and the environment, and though they tend to be prescriptive, they do not necessarily incorporate all the requirements of a sound engineering design. Therefore, as part of the design process, appropriate engineering analyses must be performed to ensure adequate performance of the final landfill and all of its interrelated systems. These include leachate, landfill gas, and surface water management systems, to name a few to be discussed in this chapter.

33.2 Regulatory Considerations and Landfill Siting

The basic federal regulation for sanitary municipal solid waste landfills in the United States is the Environmental Protection Agency (EPA) Criteria for Municipal Solid Waste Landfills (MSWLF) (40 CFR 258), generally referred to as Subtitle D (of the Resource Conservation and Recovery Act). Other federal regulations that apply to MSWLFs have been promulgated under the Clean Water Act and the Clean Air Act.

The regulations for the design of hazardous waste landfills have been promulgated under Subtitle C of the Resource Conservation and Recovery Act. The regulations for the design of municipal and hazardous waste landfills are similar, except for the requirements for liner and final cover systems. The discussion presented in this chapter is applicable, in general, to both municipal and hazardous waste landfills, except for the details related to liner and final cover systems where this paper focuses on MSWLFs.

33.2.1 Subtitle D

The purpose of the Subtitle D regulations is to establish minimum national criteria for MSWLFs. As is common with many Federal regulations, the Federal Government has delegated the responsibility to implement their regulations to each individual state. Each state was required to submit to the EPA a plan for implementation of the Subtitle D regulations. Most states have received approval of their implementation plan and are designated “approved states.” Subtitle D regulations are organized into the following subparts: Location Restrictions, Operating Criteria, Design Criteria, Groundwater Monitoring and Corrective Action, Closure and Post-Closure Care, and Financial Assurance Criteria. The Location Restrictions specified under Subtitle D are discussed in the section below on landfill siting. Other Subtitle D requirements are discussed throughout the chapter.

33.2.2 Clean Water Act

The regulations applicable to MSWLFs promulgated under the Clean Water Act are the wetlands regulations and the National Pollutant Discharge Elimination System (NPDES) program. For MSWLF units, the NPDES program is applicable to discharges from storm water management systems and leachate treatment plants or ponds, as well as storm water discharges during construction. The NPDES program establishes discharge criteria and standards for the imposition of technology-based treatment requirements in permits issued under the Clean Water Act.

33.2.3 Clean Air Act

The EPA has promulgated standards and guidelines under the authority of the Clean Air Act entitled “Standards of Performance for New Stationary Sources and Guidelines for Control of Existing Sources: Municipal Solid Waste Landfills.” These standards and guidelines, promulgated in 1996, regulate MSWLF

emissions, commonly referred to as “landfill gas,” which is composed primarily of carbon dioxide, methane, and non-methane organic compounds (NMOCs). These gasses are produced as the waste decomposes and may be discharged from the gas collection system or directly from the waste mass itself.

As with Subtitle D, the responsibility of implementing the New Source Performance Standards (NSPS) lies with the EPA. However, the EPA has the ability to delegate authority to the State and in many cases, has done so. A set of rules are included in the federal rules specifically to deal with Emissions Guidelines (EG) for existing landfills; these rules require a State to submit plans that establish emission standards for an existing source when NSPS have been promulgated for a designated pollutant, such as landfill gas, and establishes minimum requirements as guidance.

The provisions of the NSPS and EG rule apply to all new landfills with a maximum design capacity equal to or greater than 2.5 million Mg or 2.5 m³. The provisions include criteria for determining landfill control requirements, design and operating specifications for control equipment, compliance schedules, and criteria for removal of controls. The provisions also include a series of monitoring, record keeping, and reporting requirements.

A three-tiered calculation process is applied to identify landfills that exceed 50 Mg (55 tons) of NMOC emissions annually. The NMOC emission rate has been selected as a surrogate for total landfill gas emissions because it is relatively easy and inexpensive to measure.

33.2.4 State and Local Regulations

State and local regulations vary significantly and may affect both the lateral and vertical limits of landfill development. These frequently include minimum setback distances from:

- Surface water bodies
- Sources of drinking water
- Public roadways
- Property lines
- Structures such as homes, schools, hospitals, and nursing homes
- Cultural resources such as historic parks, monuments, structures, and museums
- Recreational parks
- Designated wild and scenic rivers
- Coal seams or limestone outcrops
- Gas or oil wells

Other regulatory aspects that may affect the design or permitting are:

- Zoning
- Buffers
- Deed restrictions
- Special protection areas
- Traffic/access roads
- Height restrictions
- Groundwater separation
- Maximum/minimum final cover slopes

Typically, State regulations are administered by the State’s pollution control agency. Local regulations are most often found with the jurisdiction that has land use control, that is, the city, county or township.

33.2.5 Landfill Siting

33.2.5.1 State and local considerations

The majority of these local regulations control landfill siting. The siting phase of landfill projects is often the most labor and cost intensive and often involves some level of federal review even for states that have

been approved by EPA to administer their own programs. In Minnesota, for example, State rules identify the size of a proposed facility or facility expansion that requires the preparation of a state Environmental Impact Statement (EIS). This state process is analogous to the federal EIS process. During the siting phase, hydrogeologic and geotechnical investigations are completed as are master planning and conceptual design. Often several sites are included in initial planning and siting, each undergoing investigation and evaluation, until one site is identified as the most cost-effective site meeting regulatory approval.

33.2.5.2 Federal considerations

At a minimum, Subtitle D specifies the following location restrictions: Airport Safety, Floodplains, Wetlands, Fault Areas, Seismic Impact Zones, and Unstable Areas. A brief description of each of these restrictions is presented below.

1. *Airport Safety*: Existing units, new units, and lateral expansions located within 10,000 ft of any airport runway end used by turbojet aircraft, or within 5,000 ft of any airport runway end used only by piston-type aircraft must demonstrate that the facility does not pose a bird hazard to aircraft.
2. *Floodplains*: Existing units, new units, and lateral expansions located in 100-year floodplains must demonstrate that the facility does not restrict the 100-year flood flow, reduce the temporary water storage capacity of the floodplain, or result in the washout of solid waste so as to pose a hazard to human health or the environment.
3. *Wetlands*: This restriction requires that new units and lateral expansions shall not be located in wetlands, unless the following demonstrations are made:
 - Where applicable under Section 404 of the Clean Water Act, a demonstration must be made that no practicable alternative to the proposed landfill is available.
 - Construction and operation of the MSWLF unit will not cause or contribute to:
 - Violations of any applicable state water quality standard
 - Violate any applicable toxic effluent standard or prohibition under Section 307 of the Clean Water Act
 - Jeopardize the continued existence of threatened or endangered species or result in the destruction or adverse modification of a critical habitat protected under the Endangered Species Act of 1973
 - Violate any requirement under the Marine Protection, Research, and Sanctuaries Act of 1972 for protection of a marine sanctuary
 - To the extent required under Section 404 of the Clean Water Act or applicable state wetlands regulations, a demonstration that steps have been taken to attempt to achieve no net loss of wetlands (as defined by acreage and function). This is to be done by first avoiding impacts to wetlands to the maximum extent practicable, minimizing unavoidable impacts to the maximum extent practicable, and finally offsetting remaining unavoidable wetland impacts through all appropriate and practicable compensatory mitigation actions.
4. *Faults*: New units and lateral expansions may not be located within 200 ft of faults that have experienced displacement in Holocene time, unless it is demonstrated that a smaller setback distance will be protective of human health and the environment.
5. *Seismic Impact Zones*: New units and lateral expansions may not be located in seismic impact zones, unless it is demonstrated that the facility is designed to resist the maximum horizontal acceleration in lithified material for the site (Seed and Bonaparte, 1992).
6. *Unstable Areas*: Existing units, new units, and lateral expansions located in unstable areas must demonstrate that engineering measures have been incorporated into the design to ensure that the integrity of the structural components of the unit will not be disrupted. Unstable areas include areas subject to instability due to local soil or geologic conditions or conditions made potentially unstable due to adjacent development or human intervention.

33.2.6 Hydrogeologic and Geotechnical Considerations

To verify that the contemplated landfill development will not adversely affect existing groundwater quality and availability, regulations require that a number of aspects be adequately addressed. These include:

- Depth to groundwater
- Delineation of regional confined and unconfined aquifers
- Direction and rate of groundwater flow, including possible spatial and temporal variations
- Aquifer characteristics including thickness and saturated (and, as applicable, unsaturated) hydraulic conductivity
- Regional uses of the groundwater
- Background groundwater quality

A complete characterization of the hydrogeologic regime must include an understanding of the current and previous uses of the regional groundwater. A full understanding of the site aquifer characteristics and soil permeability is also needed. For example, an understanding of both saturated and unsaturated hydraulic conductivity may be required to accurately model how contaminants may be transported.

Many state regulations include specific requirements related to hydrogeologic and geotechnical field investigations, such as the minimum number and depth of test borings, test pits, piezometers, and monitoring wells. The number is based in part on the size of the proposed facility and on adjacent land uses. Depths are based on the proposed base grades (bottom of base liner), depth to the water table, and presence of additional lower aquifer units. Some regulations may provide guidance to use supplemental investigative techniques such as geophysics and other remote sensing methods. Some state regulations specify *in situ* hydrogeologic testing of aquifers using individual well or multi-well pumping tests to determine aquifer characteristics. These tests may be essential in determining the presence of an aquitard. Frequencies of groundwater level measurements may be specified to construct appropriate seasonal or regional groundwater potentiometric surface maps.

Often, specific soil laboratory tests and frequencies are specified to characterize subsurface soils or to evaluate suitability of potential borrow soil sources for construction, operation, and closure of the MSWLF unit. This can avoid the need for additional testing in the future.

Additional geotechnical factors that must be considered in the formulation of investigation programs include:

- Proposed height of landfill at full development and total depth of waste deposited
- Presence of near-surface bedrock
- Interrelated topographic/geotechnical considerations (e.g., substantial cuts or fills)
- Stability of natural and cut slopes
- The presence of soft, compressible, collapsible, or otherwise unsuitable foundation soils
- Availability of cover materials and uses for any excavated soils

33.3 Site Master Plan

Preparation of the Site Master Plan should consider the facilities required at the site, site-specific constraints (setbacks, buffer zones, wetlands, etc.), and optimal use of the land. The main objectives of preparing the Site Master Plan are: (1) to optimize the use of the land and resources that will be used for waste disposal; (2) to ensure that the relative locations of support facilities are convenient for landfill operation; and (3) to limit future changes to accommodate facilities that were not initially included in the original design and construction (Repetto and Foster, 1993).

The type of facilities required and their sizes depend mainly on the desired facility life (total airspace required), the expected waste stream, composition and intake rates, and specific regulatory requirements.

Typically, a Site Master Plan includes at least the following facilities:

- Waste disposal area
- Buffer zones
- Access, entrance, and internal roads
- Sedimentation/detention basin(s)
- Scale(s) and scale house
- Office building
- Maintenance building
- Stockpile area(s)
- Leachate loadout
- Landfill gas management system, including blower and flare or gas-to-energy plant
- Perimeter fence and gate
- Environmental monitoring equipment (groundwater monitoring wells and piezometers, landfill gas monitoring probes, surface water monitoring locations, settlement plates, leachate level monitoring wells or manholes, and others).

A Site Master Plan may also include facilities for recycling and composting, leachate pretreatment or treatment, vehicles storage, and on-site borrow.

33.4 Landfill Configuration

A typical landfill design includes the following: landfill footprint; subbase grading; final cover grading; cell layout; perimeter and intercell berms; liner system section; final cover section; storm water management system; leachate collection, recirculation, and removal system; gas management system; and, environmental monitoring system.

The interdependence between the various landfill elements and the sequence in which they are considered in the design plays a very important role in landfill design. The most common landfill types, from geometrical configurations are (Figure 33.1):

- Area fill: Landfilling progresses with little or no excavation. Normally used in relatively flat areas with a shallow groundwater table
- Above-ground and below-ground fill: The landfill consists of a two-dimensional arrangement of large cells that are excavated one at a time. Once two adjacent cells are filled, the area between them is also filled. Normally used in relatively flat areas without a shallow water table
- Valley fill: The area to be filled is located between natural slopes. It may include some below ground excavation
- Trench fill: This method is similar to the above-ground and below-ground configuration, except that the cells are narrow and parallel. It is generally used only for small waste streams.

The valley fill and trench fill configurations are shown primarily for historical purposes when dealing with existing landfills that have been filled over the last several decades and are, for the most part, now closed. Present day designs try to avoid valley locations to avoid troublesome surface water issues and associated geotechnical issues. Trench fills are avoided because the soils remaining between the trenches are essentially wasted airspace and waste streams are now larger than in the past. The new sites accommodate the existing terrain and use constructed berms, usually 6 to 8 ft in height, in the base of the landfill to delineate cell boundaries and serve as markers for construction of future adjacent cells. The base liner, including any geosynthetic layers, runs up the side of the berm and is buried into it. A styrofoam or plywood marker is used to indicate the edge of liners.

Cells are filled to predetermined heights and at predetermined slopes, usually in lifts of 6 to 10 ft. Waste is deposited by packer trucks and rolloff vehicles directly onto the “open face,” where loaders or compactors wait to compact it into place. Present day landfills can be more than 300 ft deep and cover hundreds of

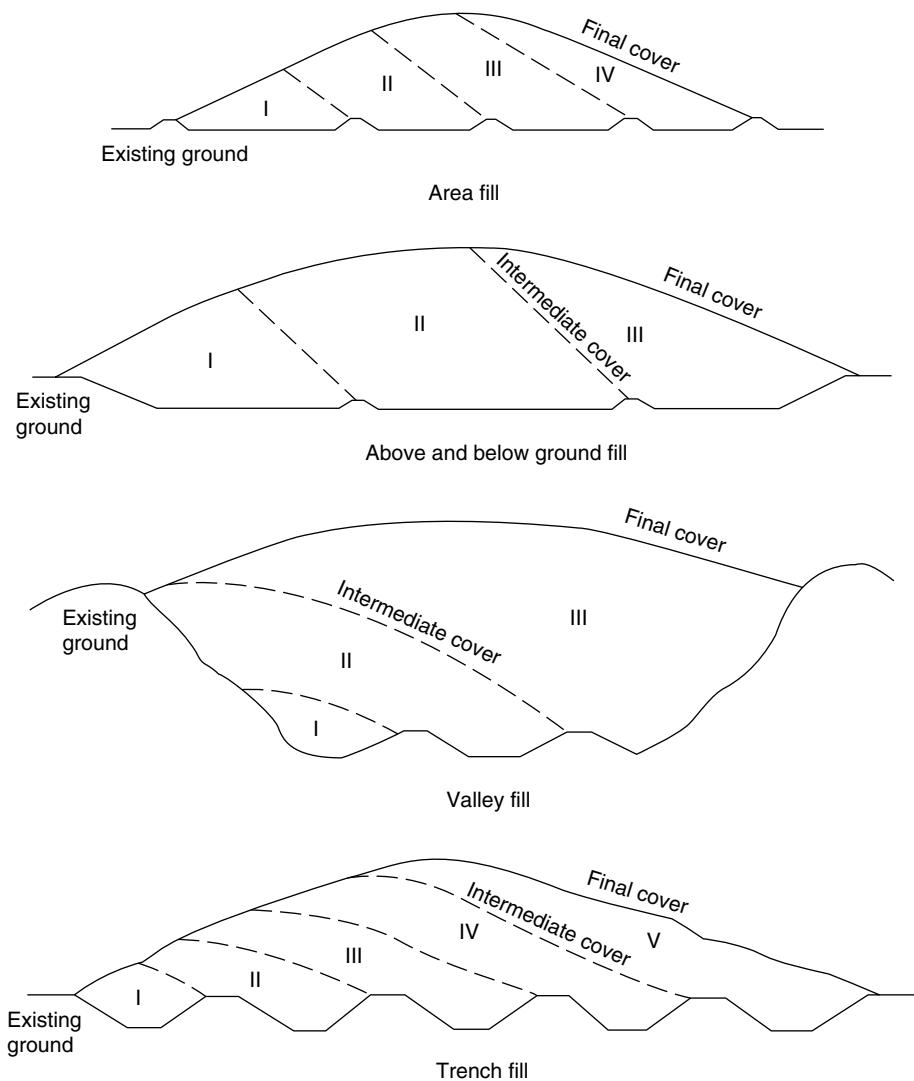


FIGURE 33.1 Landfill configurations.

acres. Waste acceptance rates are usually limited by traffic considerations and large new facilities include access roadways designed to allow placement of more than 2000 tons of waste daily.

33.5 Conceptual Design

The design of a landfill must take into account the interdependence of the various landfill elements, which can lead to several design iterations. To minimize design iterations, the design is generally subdivided into two stages: conceptual design and detailed design.

The conceptual design comprises the basic elements that define the cost of the project and its economic feasibility. These basic elements are those related to optimal use of the land, the airspace, and the soil balance. The airspace should include an allowance for daily, intermediate, and final covers. Soil balance is defined as the comparison between the volume of soil to be excavated and the volume of soil required for the entire project. Soil is excavated at landfill projects for cell development and for sedimentation/detention

basins. Excavated soil can be used for structural fill (subgrade fill, perimeter, and intercell berms); daily, intermediate and final covers; soil liners; sedimentation/detention basin berms; roads; and final land forms required by a particular end use. The soil balance applies not only to the total volume, but to soil of specific properties, as soil used for different purposes have different specifications.

The basic elements that must be included in the conceptual design are the landfill footprint, the subbase grading plan, the final grading plan, sedimentation/detention basins, borrow areas, stockpile areas, cell layout, and entrance and internal roads. The final grading plan and access roadways are often defined with a specific end use of the property in mind. Closed landfills are currently in use as parking facilities, golf courses, and various other innovative uses. Detailed design of the other elements, such as the leachate collection and removal system, the storm water management system, and the gas management system is not needed until the conceptual design has been completed.

Although the conceptual design must account for all the site constraints, end use plans, and regulatory requirements, it should be preliminary and simplified. The landfill elements included in a conceptual design are discussed below.

33.5.1 Landfill Footprint

The landfill footprint is generally selected by maximizing its area, considering the rest of the required facilities and site constraints (see Figure 33.2). Prime importance must be given to critical issues addressed during the permitting process (e.g., wetlands, groundwater issues, environmental monitoring requirements). At the conceptual design level, selection of the footprint does not need to be entirely accurate; for example, an allowance should be included for the width of perimeter berms, but the berms do not need to be designed.

The landfill footprint selected in the conceptual design stage will require some adjustments when other landfill elements are designed as part of the final design. However, if the conceptual design has been

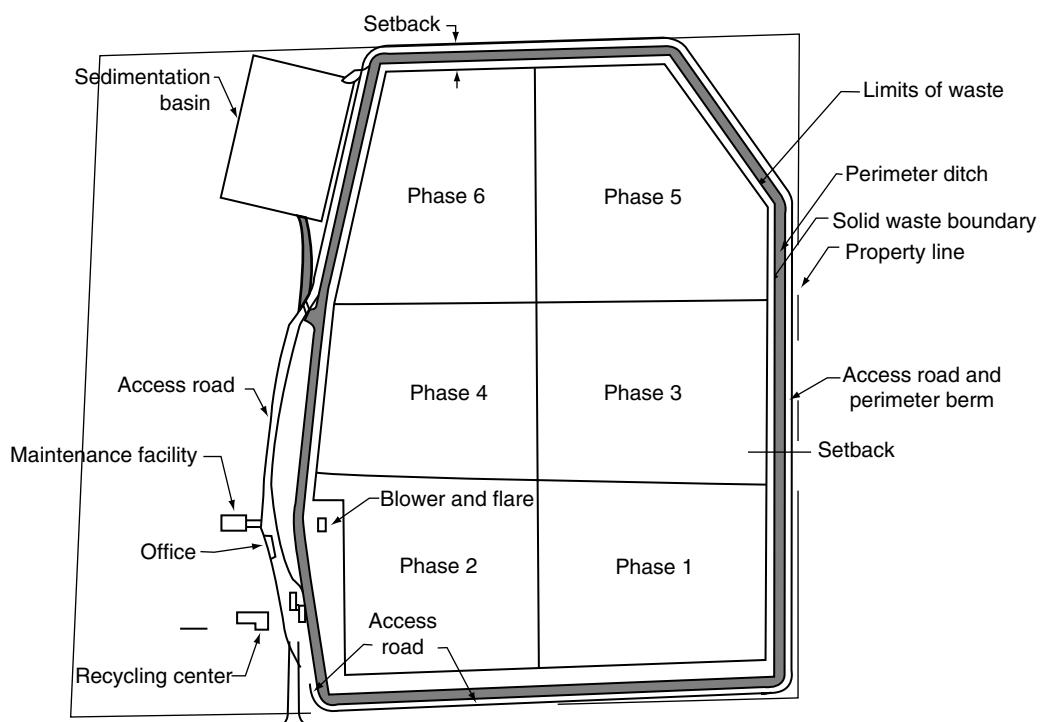


FIGURE 33.2 Landfill footprint and cell layout.

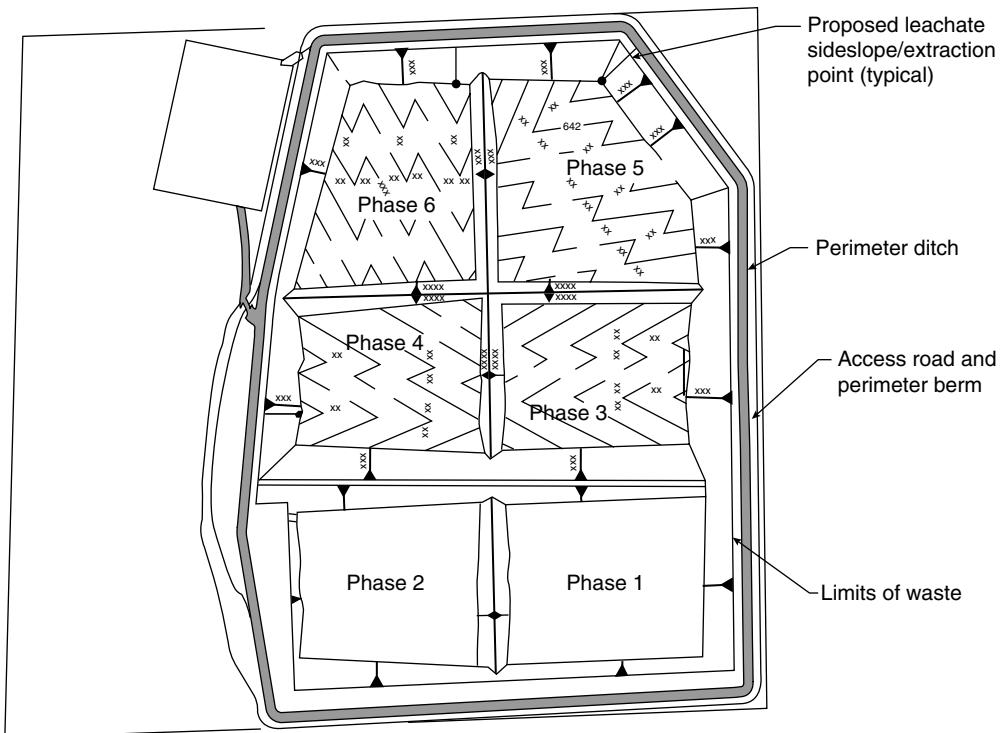


FIGURE 33.3 Subbase grading plan.

properly developed, the adjustments will not significantly modify the soil balance or the location of other facilities at the site.

33.5.2 Subbase Grading

Once the footprint has been defined, the next step is to prepare a subbase grading plan (see Figure 33.3), which identifies areas and depths of excavation, as well as areas and heights of fill. Selection of the landfill footprint and the subbase grading plan is an iterative process that depends significantly on the soil balance. At the conceptual design level, the subbase grading should be sufficient to obtain a preliminary estimate of excavation and fill volumes, and, together with the final cover grading, to estimate the airspace. It should consider vertical separation from groundwater, side slopes, and overall leachate collection system slope. However, at this level it need not include detailed subbase grading (leachate collection swales) or design of the intercell berms.

Subbase grades are generally set as deep as possible, so that airspace is maximized and soil is made available for landfill development. However, the footprint shape and extent, cell layout, excavation depth, and subbase design are strongly influenced by a number of factors, such as:

- *Depth to the water table or uppermost aquifer:* Some regulations require a minimum vertical separation from groundwater. Even if regulations do not require separation from groundwater, the cost of constructing base liner below the water table is usually prohibitive.
- *Depth to bedrock:* The cost of hard rock excavation negatively impacts the economic feasibility of a project. Additionally, excavated rock generally has little use in a landfill project unless it is processed (e.g., by crushing and sorting).
- *Stability of the foundation:* Areas previously mined or susceptible to sinkhole development may cause instability problems. In addition, areas with underlying peat or soft soils subject to

consolidation are unsuitable, as the base liner grades must remain as designed to accommodate leachate collection. Differential settlement of the base grade could damage geosynthetic liner components or leachate collection piping.

- *Site topography and stability of slopes:* In flat areas, only fill embankments and cut slopes for cell excavation need to be considered, whereas valley excavations can affect natural slopes. Excavation close to natural slopes may affect their stability and need to be considered in selecting the landfill footprint.
- *Stability of cut slopes:* The stability of cut slopes is a function of their height, inclination, groundwater conditions, and the strength and unit weight of the *in situ* soils. The inclination and height of the cut slopes affect the landfill airspace and the volume of soil to be excavated.
- *Permeability of natural soil:* In some cases a natural clay liner may be acceptable, which would eliminate the need for low permeability borrow for an engineered liner.
- *Soil balance:* An adequate soil balance allows a sufficient volume of soil to be available from on-site excavations and, at the same time, avoids excessive excavation that would have to be disposed.
- *Required airspace and available area:* Most landfill projects include requirements for airspace and daily waste acceptance rates that are predetermined based upon financial considerations and the population they are intended to serve.
- *Slope of the landfill subbase:* A minimum slope is required for the leachate collection system.

As discussed previously, perimeter and intercell berms do not need to be designed as part of the conceptual design. However, an estimate of their volume needs to be considered in the soil balance and subtracted from the available airspace.

33.5.3 Sedimentation/Detention Basins

Sedimentation/detention basins are used to collect, treat, and control the discharge of surface water collected on the site. The location and size of sedimentation/detention basins is one of the basic elements to be considered at the beginning of the design process, as they frequently occupy significant area and influence the landfill footprint. Sedimentation/detention basins also influence the soil balance, as they require excavation and the construction of berms. A detailed design of sedimentation basins is not required at the conceptual design level, as only their location, areal extent, and volume are important at this level. The depth to groundwater at sites with a high water table may limit the effective volume of the sedimentation basins. State and local regulations may also have a bearing on function and configuration of these facilities.

33.5.4 Borrow Areas

Borrow areas need to be evaluated as part of the conceptual design, as they have a significant influence on the landfill footprint and the soil balance. Borrow areas may be located on-site or off-site. This decision depends on-site topography, soil types available on-site, land availability, desired airspace, cost of imported borrow, and final site layout. Generally, soils from cell excavation are used as borrow material; however, the need for stockpiling for future use must be considered. The selection of on-site borrow areas may affect the landfill footprint, cell development, and the sequence of filling.

33.5.5 Final Cover Grading

At the conceptual design level, the final cover grading should be sufficient to obtain a preliminary estimate of the airspace. It should consider maximum and minimum regulatory slopes, any height restrictions, slope stability, and allowance for appropriate runoff control (see Figure 33.4). However, at this level it does not need to include design of the benches/channels for runoff control; the influence of the channels/benches in the airspace can be considered by assuming an average slope.

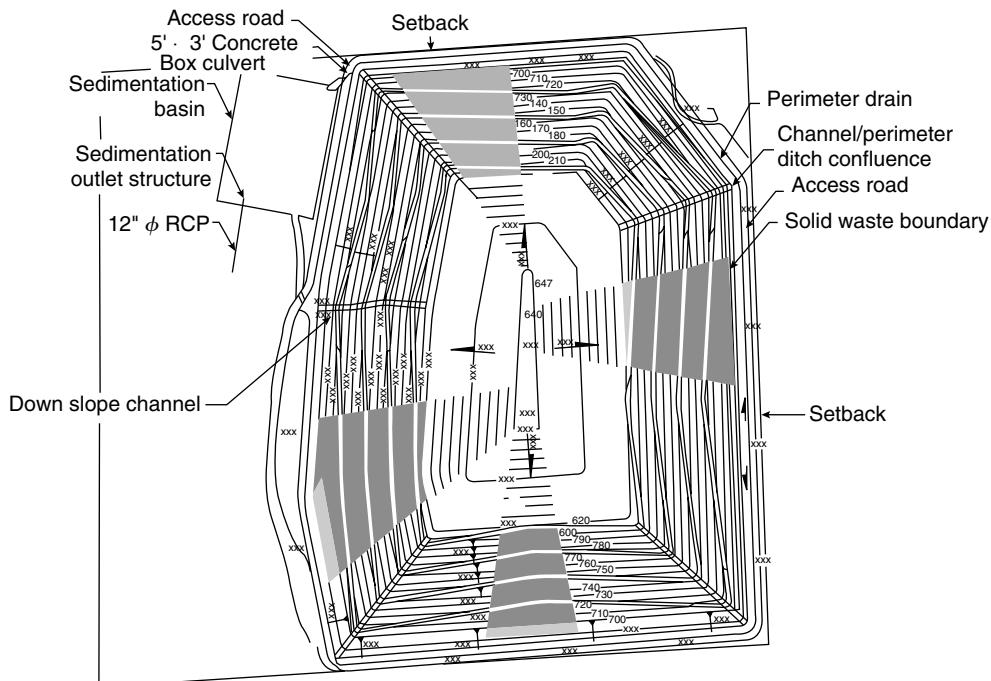


FIGURE 33.4 Final cover grading.

The final grading is generally designed as steep and high as possible, so that the maximum airspace is attained. However, both steepness and height need to be compatible with other requirements, such as regulatory restrictions regarding side slopes and height, overall stability of slopes and sliding of the final cover, and storm water management and erosion control.

33.5.6 Cell Layout and Filling Sequence

Once the overall landfill configuration (footprint, subbase grading, and final grading) has been decided, the landfill has to be divided into cells (see Figure 33.2). The objective of dividing the landfill into cells is to provide separate operational units by constructing one cell at a time, so that the initial capital investment is limited and the period of time over which the liner would be exposed is limited.

Cell sizes are selected based mainly on the waste stream and the desired life of each cell. In general, it is unusual for cells to be less than two acres or more than eight acres. The desired life of each cell is generally decided as a function of the desired schedule of construction operations. When analyzing the life of individual cells, it should be remembered that isolated cells form a single mound, while cells adjacent to a previously filled cell also add a wedge on the interim slope of the previous cell.

33.5.7 Other Factors

Other important factors to take into account in designing the cell layout are leachate collection and removal, and the filling sequence. Generally, the cells are designed with slope toward the landfill perimeter, so that leachate can be removed without crossing other cells. The cell layout plan should show all the leachate sumps and riser pipes.

It is generally advantageous to design the cell layout and the filling sequence so that excessive stockpiling of excavated soil is avoided and part of the final slopes are formed as early as possible. Once a part of

a final slope has been formed, final cover can be placed on that area. This reduces the cost of leachate treatment, as placement of final cover reduces contact water volume and infiltration into the waste.

33.6 Detailed Design

The detailed design includes refining the design of the basic landfill elements. The following sections discuss only the additional landfill elements not included in the conceptual design.

The main containment systems used in modern landfills are the bottom liner, the final cover, and the storm water management system. The function of the liner system is to prevent leachate from migrating into the subsoil. The primary function of the final cover is to minimize infiltration of precipitation into the refuse. The storm water management system controls runoff on the landfill itself to prevent erosion of the final cover, and run-on from the surrounding areas. Other types of containment systems associated with landfills are cut-off walls and blanket drains.

33.6.1 Mechanical Properties of Waste

Mechanical behavior of the waste mass, as it is deposited and after the landfill is filled to capacity, is a key factor in planning the hydrogeologic and geologic investigations, as well as in the final design of the facility and operation of the facility after it is constructed. Substantial work is being done by numerous entities to better characterize this markedly heterogeneous material (Fassett et al., 1994; Langer and Dixon, 2004). Unit weight, compressibility, shear strength, lateral stiffness and hydraulic conductivity have been studied in both the laboratory and *in situ*. An internationally vetted classification system and test standards are envisioned that will lead to the development of models using standardized input on waste composition to better predict behavior of the constructed system after placement. Such analysis will enhance the landfill design process by allowing a more indepth evaluation of interaction between liner designs, management systems designs, and surrounding geologic and hydrogeologic settings.

This is particularly important as leachate recirculation within landfills becomes more common. Liquid absorptive capacity of the waste mass is key to determining the amount of leachate recirculation that can occur initially and over time. Leachate application rates are dependent upon the hydraulic capacity of the liner drainage layer, the liquid absorptive capacity of the waste mass, and the hydraulic capacity of the waste mass. Of these three, only the hydraulic capacity of the liner system can currently be calculated with standard methods and techniques. Operation as a bioreactor landfill also necessitates evaluation of stability of the waste mass under saturated conditions.

Quantifying these characteristics in a standard manner will also allow better planning by facility operators who deal daily with ramifications of several properties of waste as it is received at the site, including compressibility to determine compaction rates, and its ability to accept and conduct liquids when leachate is recirculated.

33.6.2 Liner System

Liner systems are containment elements constructed under the waste to prevent or minimize infiltration of contaminated liquids into the subsoil or groundwater. The contaminated liquid, or leachate, may be part of the waste itself or may originate from precipitation that has infiltrated into the waste.

Landfill liner systems consist of multiple layers that fulfill specific functions. Landfill liner systems may consist, from top to bottom, of the following functional layers:

- *Protective layer:* This layer of soil, or other appropriate material, separates the refuse from the rest of the liner to prevent damage from large objects placed in the landfill.
- *Leachate collection layer:* This permeable layer collects leachate from the refuse and conveys it to sumps from where it is removed. Frequently, the functions of the protective layer and the leachate collection layer are integrated in one single layer of coarse granular soil.

- *Primary liner:* This primary low permeability layer (or composite of two different low-permeability materials in direct contact to each other) prevents or minimizes the movement of landfill leachate into the subsoil. Primary liners typically consist of low permeability soil or geosynthetics.
- *Secondary leachate collection layer* (or leakage detection layer): This high permeability (or transmissivity, if geosynthetic) layer is designed to detect and collect any leachate infiltrating through the primary liner. This layer is included only when a secondary liner is used. Collection layers typically consist of high permeability non-calcareous aggregate or rock, as well as several types of geosynthetics.
- *Secondary liner:* This is a second or backup low permeability layer (or composite of two different low-permeability materials in direct contact with each other). Not all liner systems include a secondary liner. Secondary liners typically consist of low permeability soil or geosynthetics.
- *Drainage layer:* In cases where the liner system is close to or below the water table, a high-permeability (or high-transmissivity, if geosynthetic) soil blanket drainage layer is generally placed under the liner system to control migration of moisture from the foundation to the liner system.
- *Subbase:* This bottom layer, generally site soil of intermediate permeability, separates the liner system from the natural subgrade or structural fill.

These layers are often separated by geotextiles to prevent migration of particles between layers, or to provide cushioning or protection of geomembranes.

Each of the liners (primary or secondary) may consist of one layer (clay, geomembrane, or geosynthetic clay liner [GCL]) or two adjacent layers of these materials, in which case it is called a “composite” liner. Liner systems comprising only the primary liner are called “single” liner systems. Liner systems comprising a primary and a secondary liner, with an intermediate secondary leachate collection layer, are called “double” liner systems. Sample liners are presented in Figure 33.5. The leachate collection layer(s) may be constructed of a permeable soil, a geonet, or a geocomposite. The geocomposite used for leachate collection layers consists of a geonet welded between two nonwoven geotextiles.

According to Subtitle D, the liner system must comply with either a design standard or a performance standard. The design standard under Subtitle D is a single composite liner comprised of a minimum 30-mil geomembrane placed over at least a 2-ft layer of compacted soil with a hydraulic conductivity of no more than 1×10^{-7} cm/sec. If the geomembrane is made of high density polyethylene (HDPE) the minimum thickness is specified to be 60 mils due to the effects of welding. The geomembrane must be placed in direct and uniform contact with the compacted soil.

Any alternative liner system design must demonstrate compliance with the performance standard. The performance standard requires demonstration that the maximum concentration levels of 24 chemical parameters listed in Subtitle D will not be exceeded in the uppermost aquifer at the relevant point of compliance. The relevant point of compliance must be located no more than 150 m from the waste management unit boundary and on land owned by the facility.

To perform the demonstration required for an alternative liner system design, it is necessary to conduct an assessment of leachate quality and quantity, leachate leakage through the liner system, and subsurface transport to the relevant point of compliance. The EPA has developed a modeling package to perform these analyses, called the Multimedia Exposure Assessment Model (MULTIMED, EPA, 1992), which is intended for sites where certain simplifying assumptions can be made.

33.6.3 Final Cover System

A low-permeability cover is a containment system constructed on top of the waste, primarily to control the infiltration of precipitation. Cover systems control infiltration by providing a low-permeability barrier and by promoting runoff with adequate grading of the final surface. Other functions of cover systems are to prevent contact of runoff with waste, to prevent spreading or washout of wastes, to reduce disease vectors, such as rodents, and to control the emission of gases from the decomposing refuse.

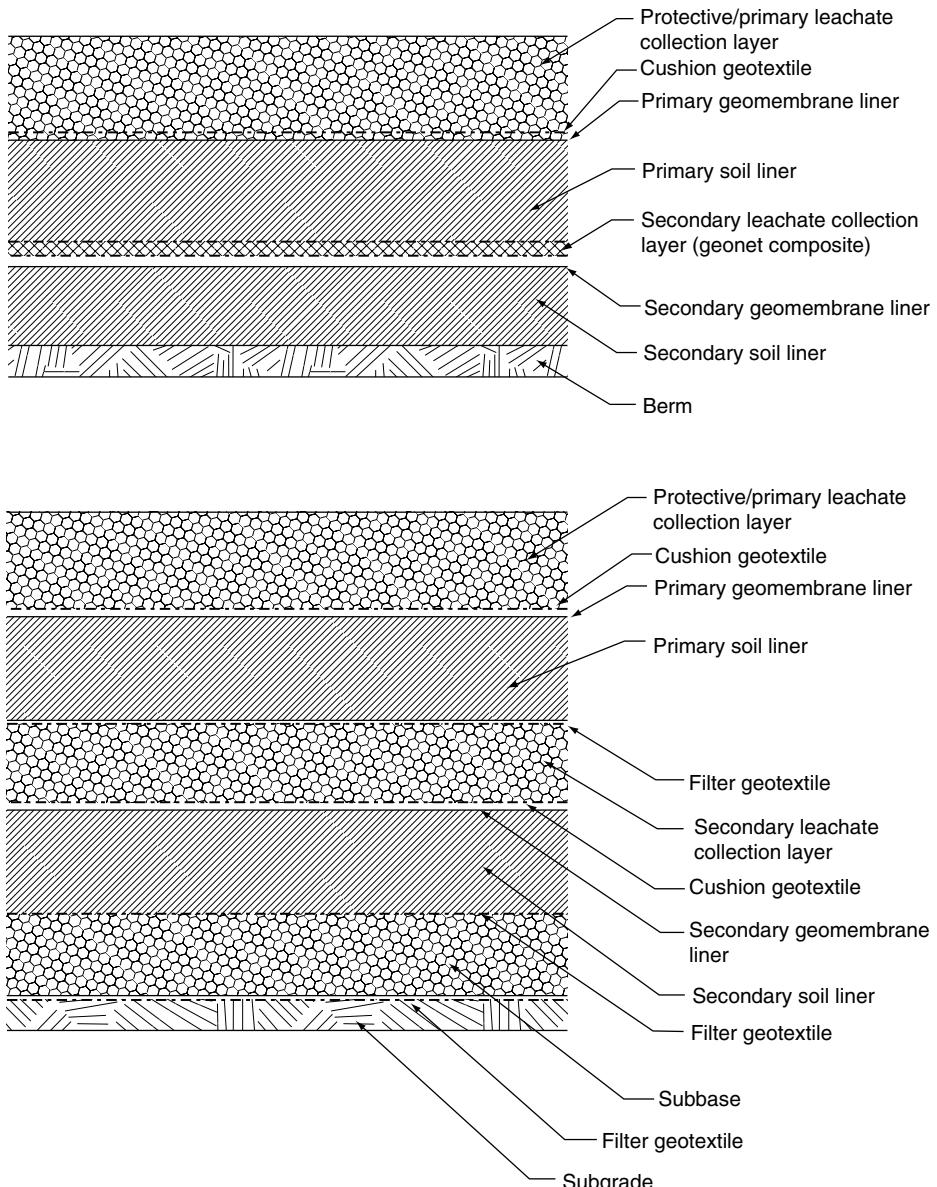


FIGURE 33.5 Double composite liner systems.

Cover systems also consist of multiple layers with specific functions related to the management of storm water. A low permeability cover system may include, from top to bottom, the following functional layers (see Figure 33.6):

- *Erosion (vegetative) layer:* This layer of soil is capable of supporting vegetation (typically grass) and has good resistance to erosion.
- *Infiltration layer:* The functions of this layer, also frequently called cover material, are to provide a minimum thickness of cover over the waste and to protect the underlying barrier layer from frost penetration.
- *Drainage layer:* This is a permeable layer whose function is to convey the water that infiltrates through the upper layers of the cover system to a perimeter collection system. A cover without

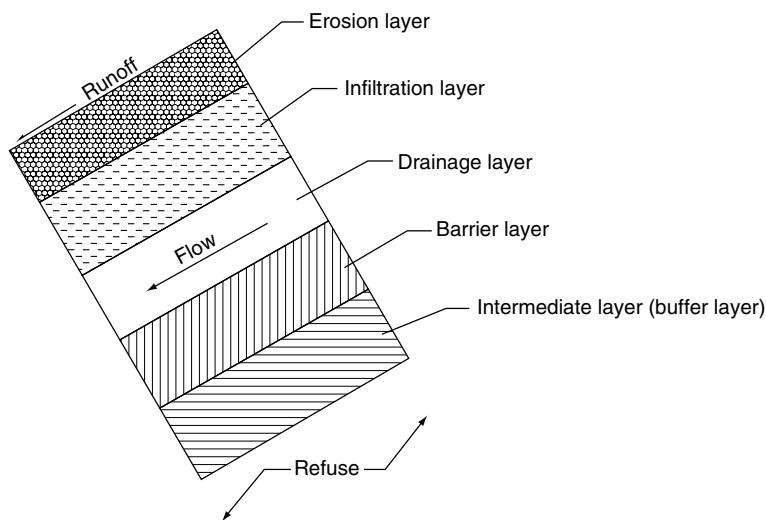


FIGURE 33.6 Final cover system functional layers.

a drainage layer is susceptible to damage occurring if the layers above the barrier layer become saturated. The drainage layer usually includes a toe drain at the toe of the final cover slope and may incorporate intermediate collector drains throughout.

- **Barrier layer:** This is a low-permeability layer (or layers of two different low-permeability materials in direct contact with each other), whose function is to minimize infiltration into the waste.
- **Intermediate layer:** This layer is placed primarily for operational reasons. It is placed directly upon the waste and is used as a subbase for the placement of the final cover system. It is often referred to as the buffer layer.

The erosion and infiltration layers are normally constructed of soil, as they must support vegetation and provide a minimum thickness for frost protection. The drainage layer can be constructed of a permeable soil, a geonet, or a geocomposite. The barrier layer can be constructed of compacted low-permeability soil, a geomembrane, a geosynthetic clay liner (GCL), or adjacent layers of two of these materials, in which case it is called a composite barrier layer. Some of these layers can also be separated by geotextiles.

Of the precipitation that falls on the cover, a part becomes surface water runoff, a portion infiltrates through the cover into the drainage layer, a part is retained in the cover soil, a small amount infiltrates into the waste through the barrier layer, and the remainder is lost through evapotranspiration. The runoff can produce erosion of the cover surface and must be controlled by means of a storm water management system. The infiltrated water primarily flows through the drainage layer, which must have adequate hydraulic capacity to convey the expected flow.

33.6.4 Berms

The design of landfills includes perimeter and intercell berms. The primary functions of the perimeter berms are to prevent the spreading of refuse outside of the disposal area during filling operations and to protect the landfill from storm water run-on from surrounding areas. The intercell berms provide separation between active cells and the rest of the landfill area. This allows construction of new cells to be performed simultaneously to the filling of active cells, and to prevent runoff from inactive cells and undeveloped areas from coming into contact with waste.

The perimeter and intercell berms are subject to driving forces produced by refuse slopes and, therefore, require structural strength. They also function to contain leachate; clayey soil works well in this regard. As long as the configuration and the material allow for adequate stability design, both clayey and granular structural fill materials can be used.

In addition to the landfill berms, sedimentation/detention basins may also include perimeter berms. These berms require low permeability in addition to structural strength.

33.6.5 Storm Water Management System

The storm water management system consists of two parts. One part controls runoff on the landfill itself to prevent erosion of the final cover. The other part is external to the landfill and its objective is to prevent run-on to the landfill from surrounding areas. Subtitle D specifies that storm water management systems be designed to collect and control at least the water volume resulting from the 24-h, 25-year recurrence interval design storm. Some state regulations also require verification of the outlet structures of sedimentation/detention basins for the 100-year recurrence interval design storm.

The runoff control system on the landfill consists generally of a system of benches or diversion berms located on the final cover (Figure 33.4). It is difficult to control runoff without benches, although it is possible for short slopes or on final covers flatter than 4H:1V. The vertical distance between the benches/channels is generally selected such that erosion of the final cover, calculated by the US Department of Agriculture Universal Soil Loss Equation, does not exceed two tons/acre/year as recommended by the EPA. An additional criterion to select the vertical distance between benches/channels is to prevent the surface water flow regime changing from sheet flow to shallow concentrated flow. By preventing this change in flow regime, the storm water has less energy with which it can mobilize soil. The final cover channels are designed with minimal longitudinal slope, more or less paralleling the final cover contours. Depending on the flow velocity, the channels may be lined with grass or rip-rap. A good rule of thumb for distance between diversion berms is 100 to 150 ft. Benches are typically spaced slightly farther apart.

The final cover channels generally discharge into gabion-lined or ungrouted rip-rapped downslope channels. A gabion lining or ungrouted rip-rap is generally used because of its ability to withstand settlements and the relatively high water velocities resulting from steep slopes. Sometimes pipes are used to convey the flow downslope; however, pipes are susceptible to damage from waste settlements. The downslope channels, in turn, discharge to other channels that convey the flow to sedimentation/detention basins. These other channels typically collect runoff from the final cover and run-on from the surrounding areas, and discharge to sedimentation basins that are often converted to detention basins after closure of the facility.

The design of the runoff control system on the final cover should be performed simultaneously to the detailed design of the final grading. At this stage, the need for a permanent access road on the final cover should also be evaluated. If such a road is required, it should be designed in conjunction with the runoff control system. These designs should also be coordinated with the design of the gas management system, as the location of gas vents or extraction wells may require the widening of benches in some areas.

33.6.6 Leachate Collection and Removal System

Landfills generate excess liquid as a byproduct of waste decomposition. This liquid, referred to as leachate, is collected by the leachate collection system (LCS), a series of pipes and trenches located at the bottom of the cell, immediately above the primary liner. The purpose of the LCS is to collect leachate from the waste and to convey it to sumps, where pumps remove it from the cell. This high-permeability collection layer oftentimes contains perforated pipes in swales surrounded by gravel (see Figure 33.7). The collection layer generally consists of noncalcareous gravel, and the perforated pipes are made of HDPE or PVC (see Figure 33.8); crushing of the collection pipes is a critical design factor. Subtitle D requires the LCS to be designed to maintain less than 1 ft of leachate head over the liner. The LCS transfers the leachate to sumps located at the lowest point of each cell, preferably adjacent to the toe of the inside slope of the perimeter berm. The leachate is typically pumped from the sump, up the side slope of the landfill, and into a leachate storage tank or to a treatment facility.

The leachate removal system typically consists of a riser pipe that parallels the inside slope of the perimeter berm commonly called a side slope riser (see Figure 33.9). The system requires a submersible

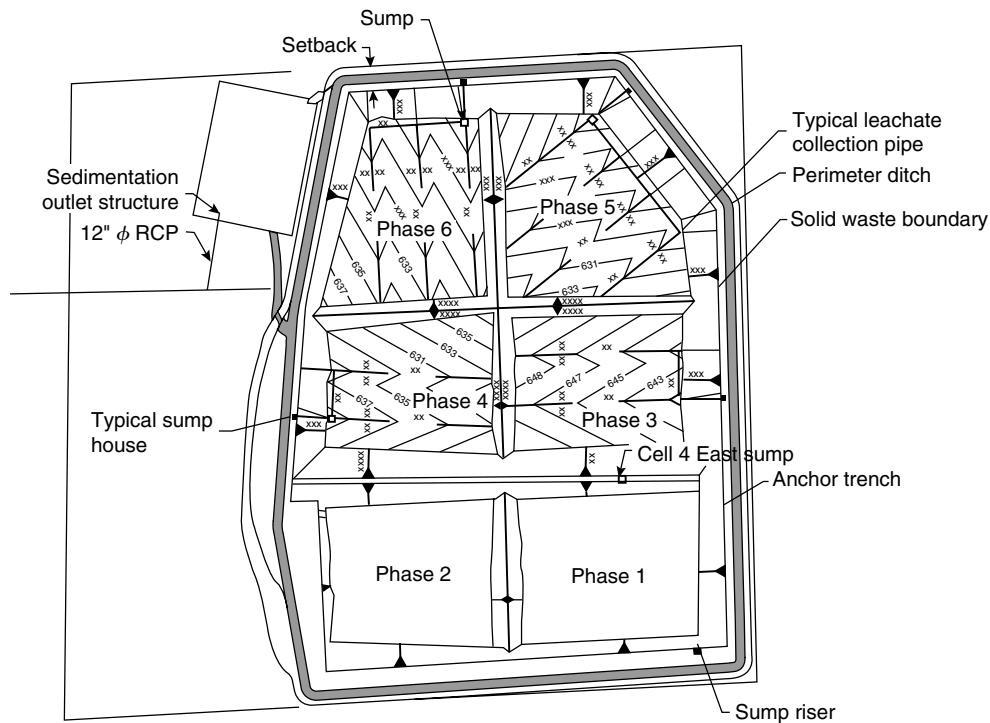


FIGURE 33.7 Leachate collection system.

pump to remove the leachate. Landfills constructed in the late 1980s and early 1990s may have accomplished leachate removal by means of a gravity flow pipe passing through the perimeter berm. This avoids the need for pumping, but requires a liner penetration that often affected liner performance; this is no longer a common practice (see Figure 33.9).

The leachate removed from the cells is generally conveyed by a header pipe to holding tanks or ponds. The header pipe may be gravity or force, depending on the site topography. The holding tanks or ponds are generally designed to hold 30 days of leachate production. Some regulations require double containment for leachate header pipes and holding tanks. Finally, the leachate may be discharged to an on-site treatment plant, or it may be piped or transported by truck to an off-site treatment facility.

The design of the final subbase grades should be performed simultaneously with the design of the leachate collection system. The modifications to the subbase grading prepared during the conceptual design phase should be such that the soil balance and the airspace should not change significantly.

33.6.7 Gas Management System

Landfills generate landfill gas as a product of decomposition of waste. Landfill gas is approximately 50% methane, 50% carbon dioxide and less than 1% trace contaminants referred to as Non-Methane Organic Compounds (NMOCs). Gas management systems can be passive or active. The type of system depends on state regulations, the size of the landfill, and site-specific issues related to odor control. As discussed under regulatory considerations related to the Clean Air Act, the NSPS and EG regulations establish that an active gas extraction system must be installed if the design capacity of the MSWLF exceeds 2.5 million tons and the total emission of NMOCs exceeds 55 tons per year.

When determining the landfill gas and the NMOC emission rate, the entire landfill must be considered, even if portions are closed. If installation of collection and control systems is required, these systems are required by rule to be installed in the entire landfill. Typical collection systems include gas extraction wells, trenches, and collection header piping and laterals. Typical control devices include flares, gas turbines,

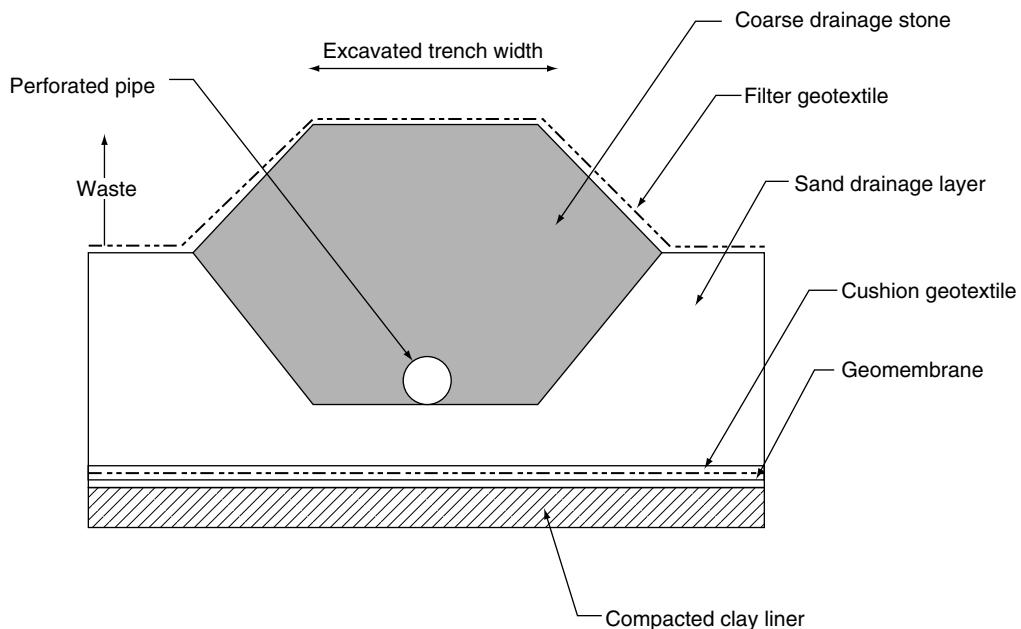


FIGURE 33.8 Leachate collection trench and collector pipe.

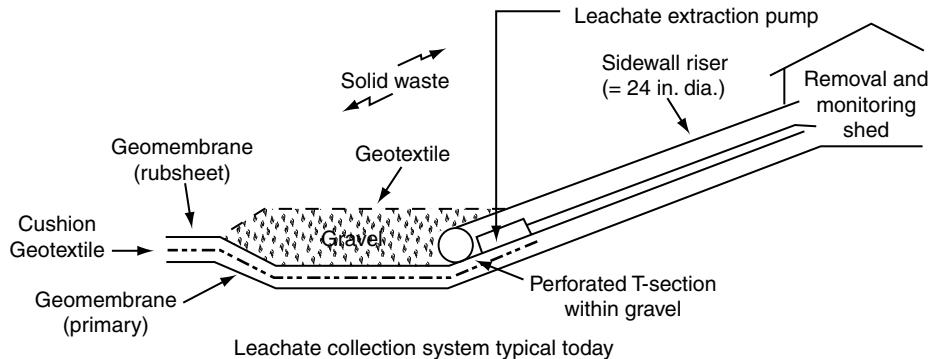
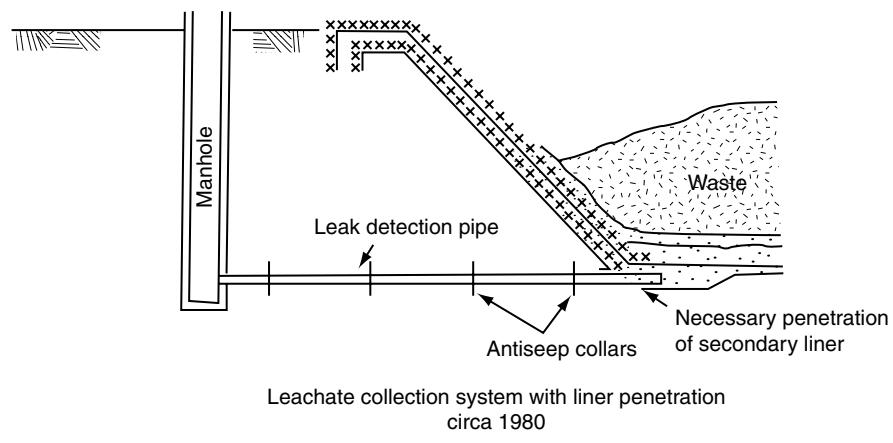


FIGURE 33.9 Leachate removal systems. (Courtesy Prof. R.M. Koerner.)

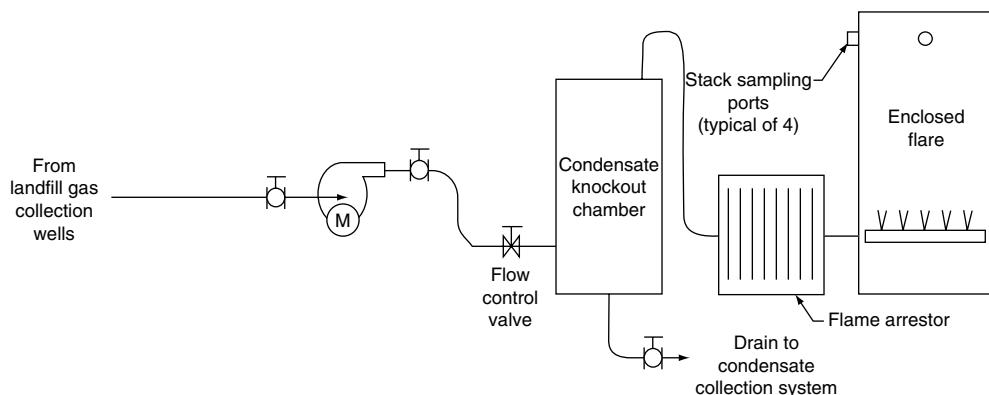


FIGURE 33.10 Landfill gas blower and flare schematic.

internal combustion engines, and boilers. All but flares provide the opportunity for energy recovery, and federal tax credits are periodically available to offset increased costs associated with power generation. A typical landfill gas blower and flare system is shown in Figure 33.10.

The NMOC emission rate is calculated as a function of the mass and age of the refuse, the gas generation constant (k), and the concentration of NMOCs (C_{NMOC}) allows the calculation of gas emissions to be based on conservative default values of k and C_{NMOC} provided by the EPA, or by using site-specific values determined from gas samples and pumping tests.

A passive system is designed to vent the gas by providing a path for gas flow without the use of additional mechanical components. A passive system generally consists of perforated pipes in rock-filled trenches embedded into the waste or a high permeability layer located within the waste mass or directly under the final cover to allow ventilation of the gas.

The primary components for an active system are gas extraction wells, trenches or layers, piping network, condensate collection traps, blowers or compressors with related power and control equipment, and disposal. From the piping network, the gas is transferred to the disposal point where the gas can be destroyed by means of open gas flares, or it can be utilized in energy generation, heating, or purified for other uses. When designing an active system, careful consideration should be given to the location of condensate traps and header pipes on the final cover, so that unnecessary iterations of the final grading plan are avoided.

33.6.8 Bioreactor Landfills

Many landfills are incorporating leachate recirculation and other operational changes that promote degradation of the waste mass at an accelerated rate. Subtitle D of RCRA allows for leachate recirculation in Section 258.28(a)(2), which provides for the placement of landfill gas condensate and leachate as the only allowable bulk noncontainerized liquid wastes to be placed in the landfill.

Leachate recirculation can be accomplished by several means, including spray irrigation, steam injection, vertical well injection, horizontal trench injection, and horizontal distribution layers. Horizontal injection trenches have pressure piping in the center to distribute the leachate to areas determined by sensors and monitored to be optimal for accepting additional moisture. Horizontal distribution layers operate under similar conditions, with pressure piping feeding geonet, stone, or other synthetic distribution layers. These two most common types of systems are used in landfills actively operating as bioreactor landfills. Spacing of the distribution system is dependent upon assumed hydraulic conductivity (horizontal and vertical) of the waste mass and the distribution medium, and the pressure used to distribute the leachate.

Bioreactor landfills show evidence of treatment of leachate over time. There are four phases of treatment during a process referred to as landfill organic stabilization. Assuming leachate starts as clean water that infiltrates into the waste mass, it starts with a pH of 7 and contaminant concentrations of zero. During

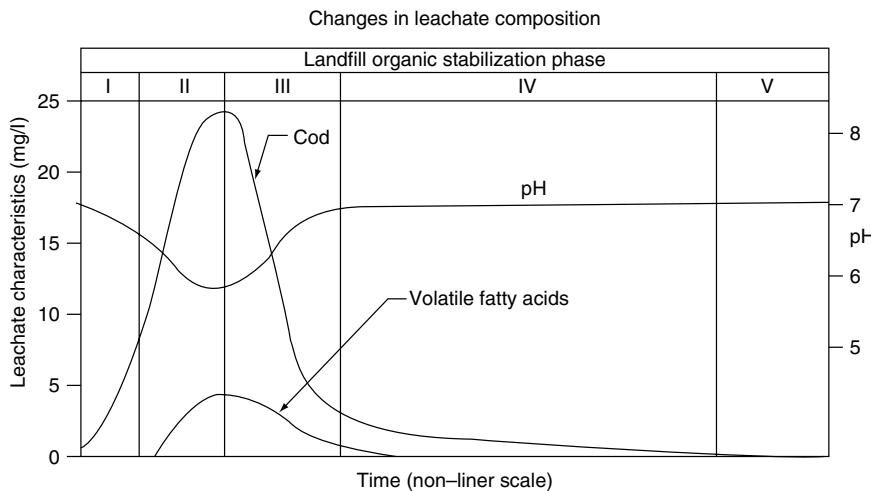


FIGURE 33.11 Changes in landfill leachate composition during landfill organic stabilization. (From Spear, J.W.Sr. (2004), Bioreactors and Leachate Recirculation Systems Design, Section 8 of Proceedings for Sanitary Landfill Design course offered through the University of Wisconsin, Madison, February 11–13, 2004.)

Phase I, the water begins to pick up contamination as it moves through the waste mass and because the contaminants are acidic, its pH drops below 7. During Phase II, the chemical oxygen demand (COD) and pH of the leachate reach extremes. During Phase III, the recirculated leachate serves as a catalyst to treatment of the organics in the landfill. Waste degrades and settlement of the landfill waste mass occurs. The level of contaminants in the leachate decrease and the pH again approaches 7. During Phase IV, this process is completed and the low levels of remaining contaminants in the leachate reach zero and the pH reaches 7. Some resources identify a fifth phase of operation where the leachate is essentially water percolating through biologically inert material. Figure 33.11 shows these five phases in sequence (Spear, 2004). Times for these phases vary depending on many factors and predictive models are not available at this point to identify either timing or the level of contamination relative to waste composition.

A similar process occurs with the landfill gas production. As leachate is recirculated, gas generation increases and peaks as the contaminants in the leachate are consumed. As shown in Figure 33.12 (Spear, 2004), gas production can reach twice the levels otherwise expected when leachate is recirculated. At present, the NSPS of the Clean Air Act do not address bioreactor landfills, but the increased production rates and other operational considerations will necessitate some additional regulatory consideration with regard to action levels and post-closure care periods.

The biology of bioreactor landfills is similar to traditional landfills, being largely dependent on anaerobic decomposition of organic matter and water into biodegraded organic matter, methane, carbon dioxide, and other gases. The limiting factor in the equation is moisture, thus the reliance on leachate recirculation and precipitation to promote the process. During initial stages, additional water is often permitted into the system based upon a “variance” or “pilot-study” allowed by regulatory agencies. Several states have issued state rules or guidance for leachate recirculation.

Bioreactor landfills have several recognized economic advantages over traditional landfills focused on safely entombing the waste. These include a delay in the need to perform leachate treatment, accelerated waste and leachate degradation, increased gas production rates, reduced post-closure care periods, all resulting in increased design capacity for the same air space and footprint. These factors change the economic models used to identify required design capacity, taking the designer back to the initial stages of the Master Planning process. As bioreactor technology reaches its full potential, landfills become more like traditional treatment facilities. Design of such facilities is in its infancy, requiring better methods to estimate landfill gas and leachate generation, to define interaction between landfill systems, and to define optimal waste type and intake rates.

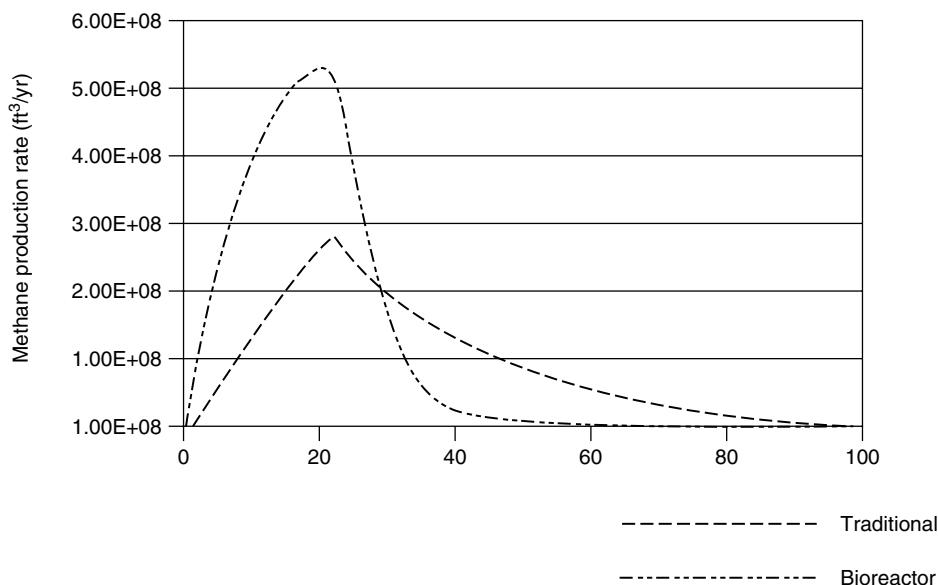


FIGURE 33.12 Methane production in traditional and bioreactor landfills. (From Spear, J.W.Sr. (2004), Bioreactors and Leachate Recirculation Systems Design, Section 8 of Proceedings for Sanitary Landfill Design course offered through the University of Wisconsin, Madison, February 11–13, 2004.)

To that end, EPA promulgated additional CAA regulations for MSWLF that require those sites operated as a landfill bioreactor and subject to existing CAA regulations (i.e., landfills containing over 2.5M³ of waste) to collect and control LFG emissions within 6 months of liquid additions (Thornloe, 2004). Using results from a study under way by the National Risk Management Research Laboratory (NRMRL), the EPA will be able to establish updated LFG emission factors for use with conventional landfilling operations. In addition, this research will establish emission factors for bioreactor operations. The research is being conducted in partnership with Waste Management, Inc. These data will be used to update the existing set of landfill gas emission factors, defined in AP-42 for EPA's landfill gas emissions model (LandGEM), used to define emissions associated with the NSPS and EG of the Clean Air Act.

33.6.9 Alternative Final Covers

The water balance models employed in designing a RCRA Subtitle D or Subtitle C final cover have been useful tools in identifying an alternative to these prescriptive cover systems. In arid or semiarid climates, water balances indicate that precipitation is so low that percolation never reaches the geosynthetic or clay barrier layer included in the final cover section. In regions where precipitation is less than 70% of potential evapotranspiration, water balance models such as UNSAT-H and the Hydrologic Evaluation of Landfill Performance (HELP) Model indicate that effective alternatives to the clay and geosynthetic layers may include capillary or other barrier systems (Khire et al., 2000).

The acceptance of Alternative Final Cover (AFC) design received a substantial boost when EPA undertook evaluation of 24 test covers at 11 sites in seven states. The Alternative Cover Assessment Program (ACAP) includes primarily sites in the dry western United States, but also includes sites in Nebraska, Iowa, and Georgia. More information on ACAP can be obtained at the EPA's ACAP web site (www.acap.dri.edu).

AFCs typically fall into two categories: Monolithic and Capillary Barrier final covers. Monolithic final covers rely on a loosely compacted fine-grained soil layer designed with a thickness that will allow it to store infiltration. The water balance is used to verify that the thickness of the cover is sufficient to store infiltration during months when vegetation is dormant and evaporation is lower.

Capillary Barrier systems rely on capillary breaks, created by placing a high permeability granular layer beneath a fine-grained soil that holds infiltration. AFCs can incorporate more than one capillary break and several layers.

The costs of these covers are substantially less than covers incorporating geosynthetics and granular or synthetic drainage layers with their associated collection systems, particularly if the required soils are readily available (Koerner and Daniel, 1992).

33.7 Liner and Cover System Materials

Materials used for the construction of liner and cover systems can be classified with respect to their origin and function. With respect to their origin they are classified into natural soils and geosynthetic materials. With respect to their function the classification is more complex and includes the following:

- Barrier layers
- Drainage layers
- Fill
- Vegetative layer
- Filtration/protection/cushion layers
- Tensile reinforcement

33.7.1 Soil Materials

33.7.1.1 Barrier Layers

Barrier layers constructed of natural soils may consist of clay or a soil-bentonite admixture. In the case of admixtures, the required bentonite content is selected based on laboratory permeability tests performed with different bentonite percentages.

The typical requirement for barrier layers is to have a permeability not exceeding 10^{-7} cm/sec. The main factors that influence the permeability of a given soil or soil-bentonite admixture are:

- *Compaction dry density*: For all other factors being constant, the higher the compaction dry density of a soil, the lower its permeability.
- *Compaction moisture content*: For all other factors being constant, the permeability of a given soil decreases with increasing compaction moisture content.
- *Compaction procedure*: It has been observed that specimens compacted to identical dry densities and moisture contents may exhibit different permeabilities if compacted using different procedures.
- *Consolidation pressure*: Barrier layers are subjected to high permanent loads while in service in the case of landfill liner systems. On the other hand, permanent loads are minimal on cover systems and an increase in dry density with time should not be expected. Therefore, permeability tests should be performed using a consolidation pressure consistent with the expected loads.

In landfills, the permeant is water in the case of covers or leachate in the case of liners. The flow of leachate through cohesive soils may produce changes in the permeability of the soil. These changes are evaluated by means of leachate compatibility tests. These are long-duration permeability tests, comparing effects on the liner material when distilled water is used initially as the permeant and then leachate from the facility (or a similar leachate) is used as the permeant.

33.7.1.2 Drainage Layers

Drainage layers constructed of natural soils consist of sands or gravels, and may be used in liners or covers. Drainage layers are designed to have a hydraulic capacity adequate to convey the design flow rate without significant head build up. The hydraulic capacity of a drainage layer is a function of its permeability, thickness, and slope. Frequently the transmissivity, defined as the product of the permeability times the thickness, is used to characterize a drainage layer.

33.7.1.3 Fill

The requirements for fill material includes the following general categories:

- *Grading fill:* There are no strict requirements for grading fills in which fill slopes are not constructed. Depending on the thickness of the fill and the loads to be applied on them, compressibility may become an important consideration.
- *Structural fill:* This category comprises fill used for elements such as intercell berms and perimeter berms in landfills, or perimeter berms for leachate ponds. When these elements will be subjected to significant lateral pressures, these berms are constructed of coarse granular soils. These berms are generally lined, so permeability is not an issue.
- *Water-containment berms fill:* Berms containing water-retaining structures, such as sedimentation and detention basin berms, require a combination of low permeability and high shear strength.

33.7.1.4 Vegetative Layer

Vegetative layers must be adequate to support vegetation, generally shallow-rooting grasses when synthetic or clay barriers are included, and must have adequate resistance to erosion. To support vegetation, the soil must contain sufficient nutrients. Nutrients can also be supplied by adding limestone or other fertilizers.

The erosion that a vegetative layer may suffer is a function of the soil type and the slope inclination and length. The soil loss is estimated by means of the Universal Soil Loss Equation, published by the US Department of Agriculture. The maximum allowable soil loss recommended by the EPA for landfill covers is 2 tons/acre/year.

33.7.2 Geosynthetic Materials

There are a large number of geosynthetic materials used in environmental applications and many different tests to characterize their properties. A comprehensive discussion of geosynthetic materials is presented in another chapter of this Handbook. Therefore this section presents a brief summary of the most common types of geosynthetic materials generally used in landfills. For a detailed discussion on the applications of geosynthetic materials, product data, and manufacturers, the Specifier's Guide of the Geotechnical Fabrics Report (Industrial Fabrics Association International (IFAI), updated annually) is recommended. The Geosynthetics Research Institute (GRI), Drexel University, has defined specifications for most common geosynthetic materials and test methods. ASTM methods and standards have also been written for many soils and geosynthetic materials used in landfill applications.

In general, two types of tests are included in geosynthetic specifications: conformance tests and performance tests. Conformance tests are performed prior to installing the geosynthetics, to demonstrate compliance of the material with the project specifications. Performance testing is performed during the construction activities, to ensure compliance of the installed materials and the installation procedures with the project specifications.

33.7.2.1 Barrier Layers

Geosynthetic barrier layers may consist of geomembranes, also called flexible membrane liners, or GCLs.

The most common polymers used for geomembranes are HDPE, linear low density polyethylene (LLDPE), polypropylene, ethylene propylene diene monomer (EPDM) and polyvinyl chloride (PVC). Selection of the polymer is based primarily on chemical resistance to the substances to be contained. The polymer most widely used for landfill liner systems is HDPE, as it has been shown to adequately resist most landfill leachates. In the case of cover system barrier layers, flexibility is frequently an important selection factor, as landfill covers are subjected to significant settlements. LLDPE geomembranes are generally more flexible than HDPE and therefore are frequently used for cover systems.

GCLs consist of a thin layer of dehydrated bentonite placed between two geotextiles or bonded to a geomembrane. The geotextiles are fixed to the bentonite layer by means of adhesives, needle-punched fibers, or stitches.

33.7.2.2 Drainage Layers

Geosynthetic drainage layers used as part of liner and cover systems may consist of geonets or geocomposites.

Polymers used for geonets are polyethylene, HDPE, and medium density polyethylene (MDPE). Geotextiles of various types and weights are attached to geonets to manufacture geocomposites.

In addition to the properties of the geonet or geocomposite, the properties of the geotextiles attached to the geonet are generally specified separately as discussed below. Seaming of geonets is performed by means of nylon ties, which are not intended to transfer stresses. For geocomposites, the lower geotextiles are generally overlapped, the geonets are joined with nylon ties, and the upper geotextiles are sewn.

33.7.2.3 Geotextiles

Geotextiles may be used in landfills to perform several different functions. The most important are:

- *Separation:* Consists of providing separation between two different materials to prevent mixing. A typical application is placement of a granular fill on a soft subgrade.
- *Reinforcement:* Applications of geotextiles for reinforcement are identical to those of geogrids, discussed below.
- *Filtration:* The geotextile is designed as a filter to prevent migration of soil particles across the geotextile.
- *Drainage:* In-plane capacity to convey flow or transmissivity.
- *Cushioning/Protection:* The geotextile serves to separate a geomembrane from a granular soil, to prevent damage to the geomembrane.

The polymers most commonly used to manufacture geotextiles are polypropylene and polyester. With respect to their structure, geotextiles are classified primarily as woven and nonwoven. Each of these types of structure is in turn subdivided depending on the manufacturing process.

33.7.2.4 Geogrids

Geogrids may be used in landfill liners and covers to provide tensile reinforcement if needed, as well as other types of reinforcement. Typical applications in landfill projects include:

- *Reinforcement of slopes and embankments:* Potential failure surfaces would have to cut across layers of geogrid. The resisting forces on the potential failure surface are the shear strength of the soil and the tensile strength of the geogrid.
- *Reinforcement of retaining structures:* In this application, some of the soil pressure that would act against the retaining wall is transferred by friction to the part of the reinforcing geogrid layer adjacent to the wall, while the rest of the geogrid layer provides passive anchorage.
- *Unpaved roads:* The stiffness of the geogrid allows distribution of loads on a larger area and prevents excessive rutting.
- *Reinforcement of cover systems:* A potential mode of failure of relative steep covers is sliding of a soil layer on the underlying layer (veneer-type sliding). To control this type of failure, a geogrid layer is embedded within the unstable soil layer to provide a stabilizing tensile force. The opposite side of the geogrid must be anchored in an anchor trench or have sufficient length to develop sufficient passive resistance to restrain its displacement.
- *Bridging of potential voids under liner systems:* When liner systems are constructed on existing waste (vertical expansions) or in areas where sinkholes may develop, geogrids are used within the liner system to allow bridging of voids.

Geogrids are made of polyester, polyethylene, and polypropylene. If exposed to waste or leachate, selection of the polymer is based on chemical resistance, as in the case of geomembranes. Depending on the direction of greater strength and stiffness, geogrids are classified as uniaxial or biaxial.

33.8 Environmental Monitoring

Landfill projects must include groundwater and landfill gas monitoring systems. If there are important water bodies nearby, surface water monitoring may also be required.

33.8.1 Groundwater Monitoring

The groundwater monitoring system specified by Subtitle D consists of a sufficient number of wells installed at appropriate locations and depths to yield groundwater samples from the uppermost aquifer at the site. A number of wells must provide samples representative of groundwater reaching the site (background) and others of the quality of groundwater passing the relevant point of compliance (downgradient).

The monitoring plan should include procedures for:

- Frequency of sampling
- Sample collection
- Sample preservation and shipment
- Analytical procedures
- Chain-of-custody control
- Quality assurance and quality control
- Statistical method for data evaluation

Groundwater elevations must be measured in each well immediately prior to purging, each time groundwater is sampled. Two types of monitoring programs are included in Subtitle D: detection monitoring and assessment monitoring.

Detection monitoring is required semi-annually at all monitoring wells during the active life of the facility and post-closure period, for 62 constituents listed in Appendix I of Subtitle D. Four independent samples from each monitoring well must be collected and analyzed for the Appendix I constituents during the first semi-annual event. At least one sample from each well must be collected and analyzed during subsequent semi-annual sampling events.

If a statistically significant increase over background concentration for one or more of the Appendix I constituents occurs, and no demonstration can be made to disassociate the increase from the MSWLF, an assessment monitoring program must be initiated within 90 days.

Assessment monitoring includes sampling and analysis for 213 constituents listed in Appendix II of Subtitle D within 90 days of triggering the assessment monitoring program and annually thereafter. At least one sample from each downgradient well must be collected and analyzed during each sampling event. If any Appendix II constituents are detected, four independent samples from background and downgradient wells must be collected and analyzed to establish the background for the constituents. Additionally, a groundwater protection standard must be established for each Appendix II constituent detected. Within 90 days of detection of Appendix II constituents, and on a semi-annual basis thereafter, all wells must be resampled and analyzed for Appendix I constituents and for the Appendix II constituents that were detected. At least one sample from each well (background and downgradient) must be collected during each sampling event. The facility may return to detection monitoring if all of the detected Appendix II constituents are at or below background levels for two consecutive sampling events.

If subsequent monitoring indicates concentrations of Appendix II constituents above background but below groundwater protection standards, the facility must continue assessment monitoring. If one or more Appendix II constituents are detected at statistically significant levels above the groundwater protection standard, the following actions must be initiated within 90 days: characterize the nature and extent of the release; install at least one additional monitoring well; notify all adjacent landowners or residents over the contaminant plume; and initiate assessment of corrective measures.

33.8.2 Gas Monitoring

The gas monitoring program generally consists of performing methane concentration measurements by means of a portable gas probe indicator. Depending on state regulations, permanent gas monitoring wells or probes may also be required as part of the design.

Subtitle D specifies the following limits with respect to gas monitoring:

- Methane concentrations within on-site structures must not exceed 25% of the lower explosive limit (LEL)
- Methane concentration at the site boundary must not exceed the LEL

33.9 Performance of Landfill Elements

33.9.1 Geotechnical and Geosynthetics Analysis

Liner and cover systems typically consist of multiple layers of soil and geosynthetic materials. This creates a complex system of materials with different strengths, moduli, and interface frictions in which geotechnical and geosynthetic analyses must be conducted to evaluate their performance. These materials are susceptible to damage by tensile stresses and strains in the geosynthetic materials and tensile cracks in the earth materials. Geotechnical analyses normally conducted for the design of landfills include overall slope stability, stability under seismic loading, sliding of the liner system on berm slopes during construction, sliding of the final cover, settlement of the landfill bottom, settlement of the final cover, and bearing capacity. Overall slope stability and stability under seismic loading are frequently some of the most critical and are briefly discussed below.

Overall slope stability analyses are performed to evaluate the stability at critical times during the filling sequence and at the end of filling. The stability analyses include slip surfaces and wedges through underlying soils, surrounding soils, refuse, and interfaces between layers of the liner system. A detailed discussion of the potential failure modes has been published by Mitchell and Mitchell (1992). Geotechnical properties of refuse and interface friction between layers of the liner system are unconventional geotechnical parameters that present significant uncertainties; reviews of available data have been published by Landva et al. (1984), Martin et al. (1984), Mitchell et al. (1990a, 1990b), Landva and Clark (1990), Singh and Murphy (1990), Fassett et al. (1994), Kavazanjian et al. (1995), and Singh (1995). Due to the diversity of materials, strain compatibility of shear strength parameters should be an important consideration in these analyses.

The analysis of landfill stability under seismic loading is required under the location restrictions specified in Subtitle D for sites located in seismic impact zones. The seismic loading in the landfill is generally determined by wave propagation analysis. The stability under seismic loading is then analyzed by means of pseudo-static slope stability analysis or calculation of seismically induced displacements. Detailed discussion of the method for seismic analysis of landfills has been published by Repetto et al. (1993a, 1993b), Augello et al. (1995), and Bray et al. (1995).

Performance of geosynthetic materials generally includes analysis of tensile stresses and strains, leachate compatibility, durability, puncture resistance, permittivity and clogging of geotextiles, and transmissivity of geonets and geocomposites. Methods to evaluate the performance of geosynthetic materials have been developed and are discussed in detail by Koerner (1994).

33.9.2 Storm Water Management

Subtitle D and State regulations typically require the design of run-on and runoff control systems for the 25 year, 24 hour storm event. Methods outlined by the US Department of Agriculture, Soil Conservation Service in the publication entitled "TR55: Urban Hydrology for Small Watersheds" (USDA 1986) and "TR-20: Computer Program for Project Formulations — Hydrology" are generally used to calculate flow

rates for the design and size of storm water control structures. Several commercial computer programs based upon these methods are available with user-friendly interfaces.

Finally, Manning's Equation for open channel flow is used to analyze channel capacity. The base width, side slopes, and longitudinal slope for the channels can be selected at this time. In general, the minimum slope used for the channels on the final cover is 2% to allow for settlement and subsidence.

33.9.3 Leachate Collection and Removal System

The size and spacing of the leachate collection pipes are dependent upon the amount of leachate generated. That can be calculated using two methods: the water budget (or balance) method and the HELP model. The water budget method, based upon the interrelationship of precipitation, evapotranspiration, runoff, soil moisture storage, and infiltration as discussed by Viessman et al. (1989), is a manual computation that can be performed using a spreadsheet type setup. This method does not, however, allow the consideration of a geomembrane component in the cover.

Conversely, the HELP model, developed by the US Army Corps of Engineers (EPA 1994), allows the inclusion of geomembranes within the cover or liner. This program requires similar input parameters to the water budget method, yet the program includes significant default data (e.g., climatological for 102 cities, soil and vegetation types, etc.). The results of the program include runoff, evapotranspiration, vertical drainage through the liner, percolation through the layers, peak daily leachate generation rates, average annual and monthly leachate generation rates, and lateral drainage from layers above the liner. The program output also provides data to select pipe spacing, which is most dependent upon the head of leachate over the primary liner. Once the leachate generation rates have been developed and the collection pipe spacing has been selected, the size and slot parameters for the pipes are calculated.

33.10 Summary and Conclusions

Development of a landfill project is a complex process that requires a thorough understanding of the physical setting of the site, relevant regulatory requirements, and engineering principles to ensure adequate performance of the various landfill elements. The design of a landfill must consider the interdependence of airspace, soils balance, geotechnical stability, current and ultimate land use plans, environmental monitoring and operation of the gas, surface water, and leachate management systems. With all of these interdependent elements, design is an iterative process, initiated with the Master Plan, followed by the Concept Plan for permitting purposes, and finally Detailed Design for construction documents. Once constructed, a monitoring program is needed to verify environmental compliance of the facility.

Glossary

The definitions presented in this section have been extracted from the Code of Federal Regulations (CFR), 40 CFR 258, *EPA Criteria for Municipal Solid Waste Landfills*, and 40 CFR 261, *Identification and Listing of Hazardous Waste*. Definitions not available in the federal regulations were obtained from the Virginia Solid Waste Management Regulations.

Agricultural Waste All solid waste produced from farming operations or related commercial preparation of farm products for marketing.

Commercial Solid Waste All types of solid waste generated by stores, offices, restaurants, warehouses, and other nonmanufacturing activities, excluding residential and industrial wastes.

Construction/Demolition/Debris Landfill A land burial facility engineering, constructed, and operated to contain and isolate construction waste, demolition waste, debris waste, inert waste, or combinations of the above solid wastes.

Construction Waste Solid waste that is produced or generated during construction, remodeling, or repair of pavements, houses, commercial buildings, and other structures. Construction wastes

include, but are not limited to, lumber, wire, sheetrock, broken brick, shingles, glass, pipes, concrete, paving materials, and metal and plastics if the metal or plastics are part of the materials of construction or empty containers for such materials. Paints, coatings, solvents, asbestos, any liquid, compressed gases or semi liquids, and garbage are not construction wastes.

Cover Material Compactable soil or other approved material, which is used to blanket solid waste in a landfill.

Debris Waste Wastes resulting from land-clearing operations, including, but not limited to, stumps, wood, brush, leaves, soil, and road spoils.

Demolition Waste Solid waste that is produced by the destruction of structures and their foundations and includes the same materials as construction wastes.

Garbage Readily putrescible discarded materials composed of animal, vegetal, or other organic matter.

Hazardous Waste The definition of hazardous waste is fairly complex and is provided in 40 CFR Part 261, *Identification and Listing of Hazardous Waste*. The reader is referred to this regulation for a complete definition of hazardous waste. A solid waste is classified as hazardous waste if it is not excluded from regulations as hazardous waste; it exhibits characteristics of ignitability, corrosivity, reactivity, or toxicity as specified in the regulations; or it is listed in the regulations (the regulations include two types of hazardous wastes: those from nonspecific sources and those from specific sources).

Household Waste Any solid waste (including garbage, trash, and sanitary waste in septic tanks) derived from households (including single and multiple residences, hotels, motels, bunkhouses, ranger stations, crew quarters, campgrounds, picnic grounds, and day-use recreation areas).

Industrial Solid Waste Solid waste generated by a manufacturing or industrial process that is not a hazardous waste regulated under Subtitle C of RCRA. Such waste may include, but is not limited to waste resulting from the following manufacturing processes: electric power generation; fertilizer/agricultural chemicals; food and related products/by-products; inorganic chemicals; iron and steel manufacturing; leather and leather products; nonferrous metals manufacturing/foundries; organic chemicals; plastics and resins manufacturing; pulp and paper industry; rubber and miscellaneous plastic products; stone, glass, clay, and concrete products; textile manufacturing; transportation equipment; and water treatment. This term does not include mining waste or oil and gas waste.

Industrial Waste Landfill A solid waste landfill used primarily for the disposal of a specific industrial waste or a waste that is a by-product of a production process.

Inert Waste Solid waste that is physically, chemically, and biologically stable from further degradation and considered to be nonreactive. Inert wastes include rubble, concrete, broken bricks, whole bricks, and blocks.

Infectious Waste Solid wastes defined as infectious by the appropriate regulations.

Institutional Waste All solid waste emanating from institutions such as, but not limited to, hospitals, nursing homes, orphanages, and public or private schools. It can include infectious waste from health-care facilities and research facilities that must be managed as an infectious waste.

Leachate A liquid that has passed through or emerged from solid waste and contains soluble, suspended, or miscible material removed from such waste.

Liquid Waste Any waste material that is determined to contain free liquids.

Liner A continuous layer of natural or synthetic materials beneath or on the sides of a storage or treatment device, surface impoundment, landfill, or landfill cell that severely restricts or prevents the downward or lateral escape of hazardous waste, hazardous waste constituents, or leachate.

Litter Any solid waste that is discarded or scattered about a solid waste management facility outside the immediate working area.

Monitoring All methods, procedures, and techniques used to systematically analyze, inspect, and collect data on operational parameters of the facility or on the quality of air, groundwater, surface water, and soils.

Municipal Solid Waste Waste that is normally composed of residential, commercial, and institutional solid waste.

Municipal Solid Waste Landfill Unit A discrete area of land or an excavation that receives household waste, and that is not a land application unit, surface impoundment, injection well, or waste pile. A municipal solid waste landfill unit may also receive other types of RCRA Subtitle D wastes, such as commercial solid waste, nonhazardous sludge, small-quantity generator waste, and industrial solid waste. Such a landfill may be publicly or privately owned.

Putrescible Waste Solid waste that contains organic material capable of being decomposed by microorganisms and causing odors.

Refuse All solid waste products having the character of solids rather than liquids and that are composed wholly or partially of materials such as garbage, trash, rubbish, litter, residues from cleanup of spills or contamination, or other discarded materials.

Release Any spilling, leaking, pumping, pouring, emitting, emptying, discharging, injection, escaping, leaching, dumping, or disposing into the environment solid wastes or hazardous constituents of solid wastes (including the abandonment or discarding of barrels, containers, and other closed receptacles containing solid waste). This definition does not include any release that results in exposure to persons solely within a workplace; release of source, by-product, or special nuclear material from a nuclear incident, as those terms are defined by the Atomic Energy Act of 1954; and the normal application of fertilizer.

Rubbish Combustible or slowly putrescible discarded materials that include but are not limited to trees, wood, leaves, trimmings from shrubs or trees, printed matter, plastic and paper products, grass, rags, and other combustible or slowly putrescible materials not included under the term *garbage*.

Sanitary Landfill An engineered land burial facility for the disposal of household waste, which is located, designed, constructed, and operated to contain and isolate the waste so that it does not pose a substantial present or potential hazard to human health or the environment. A sanitary landfill also may receive other types of solid wastes, such as commercial solid waste, nonhazardous sludge, hazardous waste from conditionally exempt small-quantity generators, and nonhazardous industrial solid waste.

Sludge Any solid, semisolid, or liquid waste generated from a municipal, commercial, or industrial wastewater treatment plant, water supply treatment plant, or air pollution control facility exclusive of treated effluent from a wastewater treatment plant.

Solid Waste Any garbage (refuse), sludge from a wastewater treatment plant, water supply treatment plant, or air pollution control facility, and other discarded material, including solid, liquid, semi-solid, or containing gaseous material resulting from industrial, commercial, mining, and agricultural operations and from community activities. Does not include solid or dissolved materials in domestic sewage, or solid or dissolved materials in irrigation return flows or industrial discharges that are point sources subject to permit under 33 U.S.C 1342, or source, special nuclear, or by-product material as defined by the Atomic Energy Act of 1954, as amended.

Special Wastes Solid wastes that are difficult to handle, require special precautions because of hazardous properties, or the nature of the waste creates management problems in normal operations.

Trash Combustible and noncombustible discarded materials. Used interchangeably with the term *rubbish*.

Vector A living animal, insect, or other arthropod that transmits an infectious disease from one organism or another.

Washout Carrying away of solid waste by waters of the base flood.

Yard Waste That fraction of municipal solid waste that consists of grass clippings, leaves, brush, and tree prunings arising from general landscape maintenance.

References

- Augello, A.J., Bray, J.D., Leonards, G.A., Repetto, P.C., and Byrne, R.J. (1995), Response of Landfills to Seismic Loading, Proceedings of GeoEnvironment 2000, Special Geotechnical Publication, ASCE.
- Bray, J.D., Augello, A.J., Leonards, G.A., Repetto, P.C., and Byrne, R.J. (1995), Seismic Stability Procedures for Solid-Waste Landfills, *Journal of Geotechnical Engineering*, ASCE, 121, 139–151.
- Fassett, J.B., Leonards, G.A., and Repetto, P.C. (1994), Geotechnical Properties of Municipal Solid Wastes and Their Use in Landfill Design, Proceedings of the WasteTech '94, Charleston, SC, National Solid Waste Management Association.
- Geosynthetic Research Institute, Drexel University, GRI Test Methods and Standards, www.geosynthetic-institute.org
- Industrial Fabrics Association International, Geotechnical Fabrics Report — Specifier's Guide, December.
- Kavazanjian, E., Jr., Matasovic, N., Bonaparte, R., and Schmertmann, G.R. (1995), Evaluation of MSW Properties for Seismic Analysis, Proceedings of the GeoEnvironment 2000, Special Geotechnical Publication, ASCE.
- Khire, M.V., Benson, C.H., and Bosscher, P.J. (2000), Capillary Barriers: Design Variables and Water Balance, *Journal of Geotechnical and Geoenvironmental Engineering*, August 2000, pp. 695–708.
- Koerner, R.M. (1994), *Designing with Geosynthetics*, 3rd Ed., Prentice Hall, Englewood Cliffs, NJ.
- Koerner, R.M. and Daniel, D.E. (1992), Better Cover-Ups, *Civil Engineering*, May, Vol. 62(5), pp. 55–57.
- Landva, A.O. and Clark, J.I. (1990), Geotechnics of Waste Fills, ASTM, STP 1070, *Geotechnics of Waste Landfills — Theory and Practice*, A. Landva and G.D. Knowles, ed., American Society for Testing and Materials, pp. 86–106.
- Landva, A.O., Clark, J.I., Weisner, W.R., and Burwash, W.J. (1984), *Geotechnical Engineering and Refuse Landfills*, Proceedings of the 6th National Conference on Waste Management in Canada, Vancouver.
- Langer, U. and Dixon, N. (2004), Mechanical Properties of MSW: Development of a Classification System. 4th BGA Geoenvironmental Engineering Conference, Startford-upon-Avon, UK, Thomas Telford, London, pp. 267–274.
- Martin, R.B., Koerner, R.M., and Whitty, J.E. (1984), Experimental Friction Evaluation of Slippage Between Geomembranes, Geotextiles and Soils, Proceedings of the International Conference on Geomembranes, Denver, pp. 191–196.
- Mitchell, R.A. and Mitchell, J.K. (1992), Stability Evaluation of Waste Landfills, Proceedings of the ASCE Specialty Conference on Stability and Performance of Slopes and Embankments — II, Berkeley, CA, June 28–July 1, pp. 1152–1187.
- Mitchell, J.K., Seed, R.B., and Seed, H.B. (1990a), Kettleman Hills Waste Landfill Slope Failure. I: Liner System Properties, *Journal of Geotechnical Engineering*, ASCE, 116, 647–668.
- Mitchell, J.K., Seed, R.B., and Seed, H.B. (1990b), Stability considerations in the design and construction of lined waste repositories, ASTM STP 1070, *Geotechnics of Waste Landfills — Theory and Practice*, A. Landva and G.D. Knowles, ed., American Society for Testing and Materials, pp. 207–224.
- Repetto, P.C., Bray, J.D., Byrne, R.J., and Augello, A.J., (1993a), Applicability of Wave Propagation Methods to the Seismic Analysis of Landfills, Proceedings of the WasteTech '93, National Solid Wastes Management Association, California, January.
- Repetto, P.C., Bray, J.D., Byrne, R.J., and Augello, A.J., (1993b), *Seismic Analysis of Landfills*, 13th Central Pennsylvania Geotechnical Seminar, Hershey, PA, April.
- Repetto, P.C. and Foster, V.E. (1993), Basic Considerations for the Design of Landfills, Proceedings of the First Annual Great Lakes Geotechnical/Geoenvironmental Conference, The University of Toledo, Toledo, OH.
- Seed, H.B. and Bonaparte, R. (1992), Seismic Analysis and Design of Lined Waste Fills: Current Practice, Proceedings of the ASCE Specialty Conference on Stability and Performance of Slopes and Embankments — II. Berkeley, CA, June 28–July 1, pp. 1521–1545.

- Singh, S. (1995), Geotechnical Considerations in the Strength Characterization of Municipal Solid Waste, Proceedings of the 10th Pan-American Conference on Soil Mechanics and Foundation Engineering, Guadalajara, Mexico, November.
- Singh, S. and Murphy, B. (1990), Evaluation of the Stability of Sanitary Landfills, ASTM STP 1070, Geotechnics of Waste Landfills — Theory and Practice, A. Landva and G. D. Knowles, ed., American Society for Testing and Materials, pp. 240–258.
- Spear, J.W.Sr. (2004), Bioreactors and Leachate Recirculation Systems Design, Section 8 of Proceedings for Sanitary Landfill Design course offered through the University of Wisconsin, Madison, February 11–13, 2004.
- Thornloe, Susan A. (2004), U.S.EPA's Field Test Programs to Update Data on Landfill Gas Emissions, Air Pollution Prevention and Control Division, Office of Research and Development, USEPA, Research Triangle Park, NC.
- United States Code of Federal Regulations, Title 40 (1991), Protection of the Environment, Part 258, EPA Criteria for Municipal Solid Waste Landfills.
- United States Department of Agriculture (1986), Soil Conservation Service, Engineering Division, TR-55: Urban Hydrology for Small Watersheds.
- United States Environmental Protection Agency (1992), Center for Environmental Assessment Modeling, The Multimedia Exposure Assessment Model (MULTIMED), Computer program available at: <http://www.epa.gov/ceampubl/mmmedia/multim1/>
- United States Environmental Protection Agency (1994), Office of Research and Development, The Hydrological Evaluation of Landfill Performance (HELP) Model, User's Guide for Version 3 (EPA/600/R-94/168a) and Engineering Documentation for Version 3 (EPA/600/R-94/168b). This software is available through the *Site Characterization Library, Volume 1, Release 2* (EPA report number 600/C-98/001) through the EPA's National Service Center for Environmental Publications (800-490-9198) (<http://www.epa.gov/ncepi>).
- Viessman, Jr., W., Lewis, G., and Knapp, J. (1989), *Introduction to Hydrology*, 3rd Edition, Harper & Row Publishers, New York.

Further Information

- Geosynthetic Research Institute, www.geosynthetic-institute.org
- International Geosynthetics Society, www.geosyntheticsociety.org
- North American Geosynthetics Society, www.nagsigs.org
- USACE Geosynthetics, www.geosynthetica.net/tech_docs/USACEdocs.asp
- USEPA. 1993. Office of Research and Development. *Quality Assurance and Quality Control for Waste Containment Facilities, Technical Guidance Document*. EPA/600/R-93/182.
- USEPA. 1993. Office of Research and Development. *Proceedings of the Workshop on Geosynthetic Clay Liners*. EPA/600/R-93/171
- USEPA. 1993. Office of Research and Development. *Report of Workshop on Geomembrane Seaming*. EPA/600/R-93/112.
- USEPA. 1988. Risk Reduction Engineering Laboratory. *Guide to Technical Resources for the Design of Land Disposal Facilities*. EPA/625/6-88/018.
- USEPA. 1986. Office of Solid Waste and Emergency Response. *Technical Guidance Document: Construction Quality Assurance for Hazardous Waste Land Disposal Facilities*. EPA/530-SW-86-031.
- USEPA. 1982. Office of Solid Waste and Emergency Response. *Evaluating Cover Systems for Solid and Hazardous Waste*. EPA/SW-867.

34

Evapotranspirative Cover Systems for Waste Containment

34.1	Introduction.....	34-1
34.2	Unsaturated Soil Hydraulic Concepts Relevant to Evapotranspirative Cover Performance	34-3
	Water Flow through Unsaturated Porous Media •	
	Unsaturated Soil Hydraulic Properties	
34.3	Types of Evapotranspirative Covers	34-9
	Monolithic Covers • Capillary Barriers • Anisotropic Barriers	
34.4	Relevant Issues for Design of Evapotranspirative Covers.....	34-13
	Design Strategies • Performance Criteria and Regulatory Issues • Important Design Variables • Numerical Modeling Issues	
34.5	Performance Monitoring of Evapotranspirative Covers	34-16
	Monitoring Strategies • Regulatory Issues Specific to Compliance Demonstration • Lysimetry • Monitoring of Moisture and Suction Profiles • Monitoring of Meteorological Variables and Overland Runoff	
34.6	Case Studies	34-19
	Site Survey • OII Landfill • Rocky Mountain Arsenal	
	Glossary.....	34-27
	References	34-27
	Further Information	34-31

Jorge G. Zornberg and
John S. McCartney
University of Texas at Austin

34.1 Introduction

The management of hazardous and municipal wastes may have important implications on groundwater quality. The current trend in waste management in the United States involves its disposal and isolation within containment facilities to minimize human and environmental contact. One of the key engineered components in municipal and hazardous waste containment systems is the cover system. The cover system is the surface culmination of a landfill, so it interacts closely with the atmosphere. Cover design and analysis involves concepts from various disciplines such as geotechnical engineering, environmental

engineering, soil science, agriculture engineering, climatology, biology, and hydrology. Integration of concepts from these various disciplines presents important challenges to researchers, designers, regulators, and stakeholders. This is particularly the case in evapotranspirative covers, the understanding of which relies heavily on the quantification of atmospheric processes at the land surface and of water flow through unsaturated soil.

The water balance components used to quantify the conservation of water mass in an engineered cover may include evaporation and plant transpiration (together referred to as evapotranspiration), precipitation, overland runoff, soil moisture storage, lateral drainage, and basal percolation. Basal percolation, an important variable to quantify the overall performance of landfill covers, is the volume of water that exits the lower boundary of the engineered cover with time. The water that cannot be removed from the cover by evapotranspiration or lateral drainage reaches the underlying waste mass, possibly mobilizing contaminants that may eventually reach the groundwater. Accordingly, one of the primary objectives of a landfill cover system is to control basal percolation. Additional objectives of landfill covers include accommodating differential settlements without compromising its performance, and controlling landfill gas release. In addition, the cover should remain stable under static and seismic conditions, minimize long-term maintenance, allow land re-use, and provide an aesthetic appearance.

The design of final cover systems for new municipal and hazardous waste containment systems in the United States is prescribed by the US Resource Conservation and Recovery Act (RCRA) Subtitles D and C, respectively. Federal- and state-mandated cover systems for municipal and hazardous waste landfills have endorsed the use of resistive barriers. Resistive cover systems involve a liner (e.g., a compacted clay layer) constructed with a low saturated hydraulic conductivity soil (typically 10^{-9} m/sec or less) to reduce basal percolation. Figure 34.1(a) shows the water balance components in a resistive system, in which basal percolation control is achieved by maximizing overland runoff. To enhance cover performance and lower construction costs, RCRA regulations allow the use of alternative cover systems if comparative analyses and field demonstrations can satisfactorily show their equivalence with prescriptive systems. Evapotranspirative covers are alternative systems that have been recently proposed and successfully implemented in several high-profile sites. Evapotranspirative covers are vegetated with native plants that survive on the natural precipitation and have been shown to be stable over long periods of time. Figure 34.1(b) shows the water balance components in an evapotranspirative cover system. Evapotranspiration and moisture storage are components that influence significantly the performance of this system. Internal lateral drainage may also be a relevant component in some cover types (capillary barriers on steep slopes). The novelty of this approach is the mechanism by which basal percolation control is achieved: an evapotranspirative cover acts not as a barrier, but as a sponge or a reservoir that stores moisture during precipitation events, and then releases it back to the atmosphere as evapotranspiration or lateral drainage. Silts and clays of low plasticity are the soils most commonly used in evapotranspirative covers, as they can store water while minimizing the potential for cracking upon desiccation.

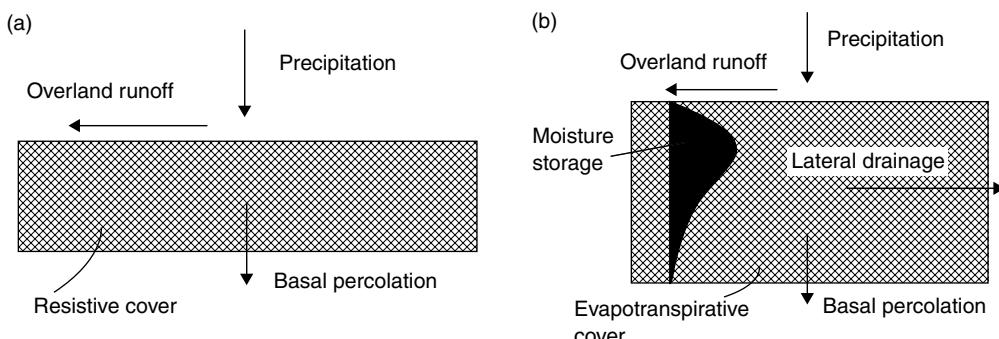


FIGURE 34.1 Water balance components: (a) in a resistive barrier and (b) in an evapotranspirative cover

Additional advantages of evapotranspirative covers over typical clay barrier systems include minimizing desiccation cracking, ease of construction, and low maintenance. Also, effective evapotranspirative covers can be constructed with a reasonably broad range of soils, which can lead to significant cost savings. The adequacy of evapotranspirative cover systems has been verified by field experimental assessments (Anderson et al., 1993; Dwyer, 1998), and procedures have been developed for quantitative evaluation of the variables governing their performance (Khire et al., 1999; Zornberg et al., 2003). The scope of this chapter includes: (i) basic hydraulic concepts for unsaturated soils relevant to evapotranspirative covers, (ii) types of evapotranspirative covers, (iii) design and modeling issues, (iv) field monitoring approaches, and (v) case studies.

34.2 Unsaturated Soil Hydraulic Concepts Relevant to Evapotranspirative Cover Performance

34.2.1 Water Flow through Unsaturated Porous Media

Designing a truly impermeable barrier (i.e., one leading to zero basal percolation) should not be within an engineer's expectations. Instead, the objective of the cover should be to minimize the basal percolation of water to acceptable levels. Quantification of the basal percolation is challenging as it involves analysis of water flow through unsaturated soils subject to complex atmospheric boundary conditions. Both air and water are present in the voids of unsaturated soils. The relative amounts of water and air in the soil, typically quantified on a volumetric basis, highly influence the soil hydraulic behavior. Figure 34.2 illustrates some of the common phase relationships used for the analysis of water flow processes in unsaturated soils. The volumetric moisture content θ , is defined as the ratio between the volume of water and the total control volume. The porosity n , which is the ratio between the volume of voids and the total control volume, corresponds to the volumetric moisture content at saturation (i.e., $n = \theta_s$). The degree of saturation S , commonly used to normalize the moisture content of a soil is the ratio between the volumetric moisture content and the porosity. Finally, the volumetric air content is the difference between the porosity and the volumetric moisture content.

In an unsaturated soil, water is held within the pores against the pull of gravity by a combination of adsorptive and capillary pressures (Olson and Langfelder, 1965). Adsorptive pressures are present in soils due to electrical fields and short-range attractive forces (van Der Waal forces) that tend to draw water toward the soil particles in highly plastic clays, where the net negative charges on water dipoles and surface of clay particles interact with the cations in the pore water. The capillary pressure is quantified as the difference between the pore air pressure and the pore water pressure. Water is a wetting fluid for most soil particles, which implies that the air-water menisci between individual soil particles are convex, tensioned membranes. Accordingly, the air pressure is greater than the water pressure, which has a negative magnitude. The adsorptive and capillary pressures are typically considered together as a single

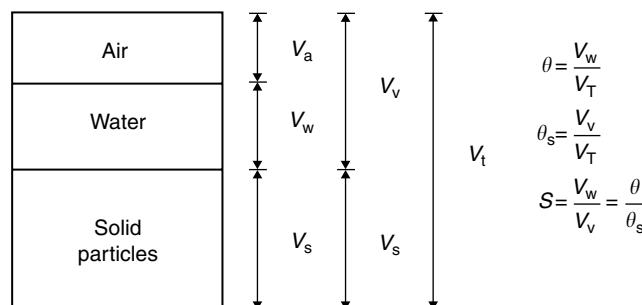


FIGURE 34.2 Volumetric phase diagram for unsaturated soils.

variable, referred to as the matric suction, ψ , which has units of pressure (kPa). The capillary rise in a pipette provides an analogy useful to assess the influence of pore sizes on the matric suction, given by the following expression

$$\psi = P_a - P_w = h_c \rho_w g = \frac{2\sigma_{aw} \cos \gamma}{R} \quad (34.1)$$

where P_a is the pore air pressure, P_w is the pore water pressure, h_c is the height of capillary rise in a pipette of radius R , ρ_w is the density of water, g is the acceleration of gravity, σ_{aw} is the surface tension between water and air, and γ is the wetting contact angle (typically 10° for quartz minerals). Equation 34.1 assumes that air is under atmospheric pressure ($P_a = 0$) and indicates that the suction is inversely proportional to the pore radius. Accordingly, for the same volumetric moisture content, a fine-grained soil (with comparatively small pore radii) will have a higher suction than a coarse-grained soil. The relationship between moisture content and suction is thus related to the pore size distribution of the soil as discussed in Section 34.2.2.

Flow of water through soils is driven by a gradient in the hydraulic energy, which is quantified by the fluid potential (energy per unit mass of water). The fluid potential is given by an expanded form of Bernoulli's equation:

$$\Phi = gz + \frac{1}{2} \left(\frac{v}{n} \right)^2 + \frac{-\psi}{\rho_w} + \frac{P_o}{\rho_w} \quad (34.2)$$

where Φ is the fluid potential, z is the vertical distance from the datum, v is the water discharge velocity, n is the porosity, and P_o is the osmotic suction. In Equation 34.2, the four terms on the right-hand side correspond to the potential energy, the kinetic energy, the energy due to the water pressure ($\psi = -P_w$ if $P_a = 0$), and the energy due to osmosis. The discharge velocity (v/n) is comparatively small, leading to a negligible kinetic energy component. The osmotic suction is typically considered constant throughout an evapotranspirative cover, and consequently does not lead to a contribution to the hydraulic gradient. As in the case of water-saturated soils, Darcy's law for unsaturated soils indicates that flow is driven by the gradient in total hydraulic potential. However, the available pathways for water flow in unsaturated soil decrease as the moisture content decreases (or as suction increases). This is quantified by the hydraulic conductivity function $K(\psi)$, which accounts for the decrease in conductivity with increasing suction (or decreasing moisture content). The K -function is discussed in Section 34.2.2.2. The discharge velocity through a soil in the vertical direction z can be estimated using Darcy's law and Equation 34.2, as follows:

$$v = \frac{Q}{A} = -\frac{K(\psi)}{g} \frac{\partial \Phi}{\partial z} = -K(\psi) \frac{\partial}{\partial z} \left(1 - \frac{1}{\rho_w g} \frac{\partial \psi}{\partial z} \right) \quad (34.3)$$

where Q is the volumetric flow rate, and A is the area of soil perpendicular to the flow direction. Figure 34.3 shows a control volume of thickness dz for one-dimensional (1-D) water flow through a soil layer with thickness L using the base of the soil layer as datum. The continuity principle in this control volume can be expressed by:

$$\frac{\partial \theta}{\partial t} = -\frac{\partial v}{\partial z} \quad (34.4)$$

where the left-hand side represents the change in moisture storage in the control volume, and the right-hand side represents the change in flow rate across the control volume. Substitution of Equation 34.3 into Equation 34.4 leads to the governing equation for 1-D flow through unsaturated porous materials, referred to as Richards' equation:

$$\frac{\partial \theta}{\partial \psi} \frac{\partial \psi}{\partial t} = \frac{\partial}{\partial z} \left[K(\psi) \left(1 - \frac{1}{\rho_w g} \frac{\partial \psi}{\partial z} \right) \right] \quad (34.5)$$

Richards' equation is a coupled, nonlinear parabolic equation, which can be solved using finite differences or finite elements. Numerical solutions to Richards' equation can be challenging because the constitutive function [$K(\psi)$ and $\theta(\psi)$] are highly nonlinear and may have undefined or zero derivatives. Further,

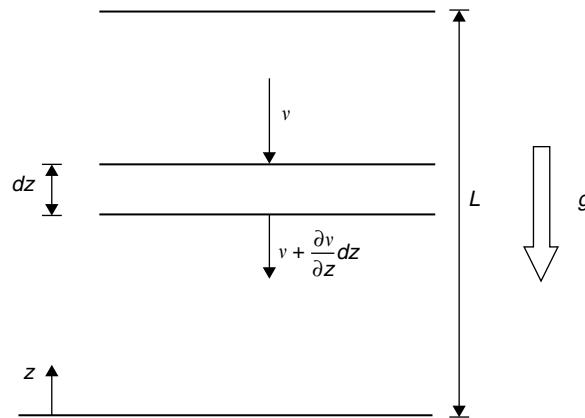


FIGURE 34.3 Control volume for vertical flow through an evapotranspirative cover

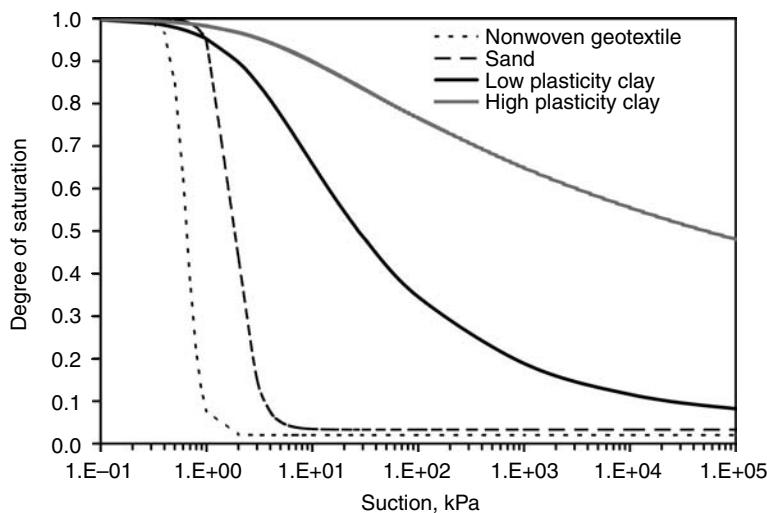


FIGURE 34.4 Typical SWRCs for different geotechnical materials

selection of boundary conditions may not be straightforward for several practical problems. Boundary conditions typically used for evapotranspirative covers include atmospheric flux boundary conditions at the surface (infiltration or evaporation), and unit hydraulic gradient (i.e., zero change in suction with depth) at the base. A discussion on the boundary conditions for evapotranspirative covers may be found in Fayer and Jones (1990). Additional difficulties for solving water flow problems for unsaturated soils arise when considering moisture removal from plant roots, desiccation cracking, animal intrusion, and volumetric changes.

34.2.2 Unsaturated Soil Hydraulic Properties

34.2.2.1 Soil Water Retention Curve

The moisture storage of the soil is an important performance variable of evapotranspirative covers. The moisture storage is typically quantified using the relationship between volumetric moisture content and soil suction, referred to as the Soil Water Retention Curve (SWRC). Figure 34.4 shows the SWRCs for different geotechnical materials. The coarser materials (sand and geotextile) show a highly nonlinear

response, with a significant decrease in moisture content (or degree of saturation) in a narrow range in suction. The fine-grained materials (silt and clay) show a more gradual decrease in moisture content with increasing suction. The nonlinearity observed in these relationships is partly caused by the range of pore size distributions for these materials. An important characteristic in a SWRC is the air entry value. During initial drying of a fully saturated soil specimen water does not flow from the soil until the suction corresponding to the air entry value is reached. When this suction is reached, air enters the specimen and the moisture content decreases. Once the air entry value is reached, the moisture content drops from saturation to a value that remains approximately constant with increasing suction. This low moisture content value is often referred to as the residual moisture content. The residual condition occurs because the water becomes occluded (or disconnected) within the soil pores, with no available pathways for water flow to occur.

The SWRC for a given material is not only sensitive to the pore size distribution, but also the soil mineralogy, density, and pore structure (Hillel, 1988). The SWRC can show significantly different wetting and drying paths, a phenomenon referred to as hysteresis (Topp and Miller, 1966; Kool and Parker, 1987). During drying, the largest pores drain first, followed by the smaller pores. During wetting, the smaller pores fill first, but the presence of large pores may prevent some from filling. Also, wetting of a dry media often leads to entrapment of air in the larger pores, which prevents saturation of the media unless positive pressure is applied to the water. Air entrapment causes the wetting path to be relatively flat for high suctions (e.g., over 100 kPa), with a steep increase in volumetric moisture content at lower suctions.

Several techniques are available to determine the SWRC experimentally (Klute et al., 1986; Wang and Benson, 2004). Two main groups of techniques have been used to define the SWRC. The first group of techniques (“physical” techniques) involves an initially water-saturated material from which water is slowly expelled by imposing a suction to a specimen boundary. Flow continues until it reaches a condition at which the moisture content and suction are in equilibrium. The most commonly used physical technique is the axis translation technique. A common test that is based on this technique is the pressure plate test (Figure 34.5(a)), which involves placing a soil specimen on a high air-entry ceramic plate and applying air pressure to the specimen. The air pressure causes the pore water to pass through the ceramic plate since the water pressure on the effluent side of the plate is kept at atmospheric pressure (zero). At equilibrium, the air pressure corresponds to the suction. The outflow volume is measured using a constant head Mariotte bottle. This approach is repeated for successively higher pressures that gradually dry the specimen, after which the pressure may be decreased to measure the wetting behavior. At the end of the testing, the gravimetric moisture content may be measured destructively, and the moisture content at each pressure increment can be back-calculated from the outflow measurements. Additional details can be found in the ASTM standard for SWRC determination (ASTM D6836 2002). Another technique, the hanging column, is shown in Figure 34.5(b). This test also involves a ceramic plate, but connects the bottom of the plate to a manometer tube. A negative water pressure is imposed on the water level in the ceramic plate by holding the manometer tube beneath the plate.

The second group of techniques (“thermodynamic” techniques) involves allowing water to evaporate from a specimen in a closed chamber under controlled relative humidity. The relative humidity is controlled by allowing water to evaporate from a saturated salt solution placed within the chamber, as shown in Figure 34.5(c). Another commonly used thermodynamic technique is the chilled mirror

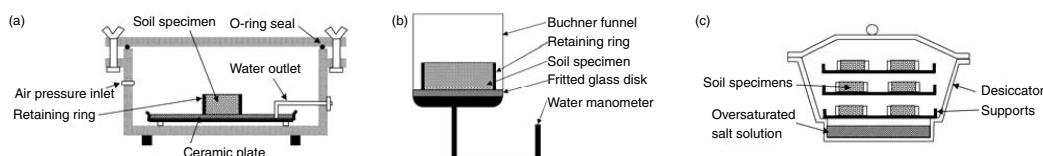


FIGURE 34.5 Conventional methods to determine the SWRC: (a) pressure plate; (b) hanging column; and (c) saturated salt solutions.

hygrometer (Wang and Benson, 2001). This device infers the total soil suction (matric and osmotic) by measuring the vapor pressure in the soil, which is related to the temperature at which moisture condenses on a mirror. When condensation occurs, a change in the optical properties of the mirror is detected. In general, physical techniques are used for relatively low suctions (e.g., under 1500 kPa), while thermodynamic techniques are used for higher suctions.

Conventional techniques to define the SWRC require significant time to obtain limited data. For example, determination of the SWRC for a high-plasticity clay specimen may take several months. Also, conventional determination methods require the use of several specimens and destructive measurement of moisture content. Problems specific to SWRC testing involve diffusion of air across porous ceramics, lack of control of volume change during drying and wetting (e.g., Cabral et al., 2004), and inability to impose a stress state representative of field conditions.

Centrifugation can be used to alleviate shortcomings of conventional characterization of the SWRC. Centrifugation increases the body forces on a porous media, accelerating fluid flow as time increases quadratically with g-level. Centrifuges were first used in the early 1930s to define the SWRC by soil scientists and petroleum engineers (Gardner, 1937; Hassler and Bruner, 1945). Saturated specimens can be placed upon a saturated ceramic plate that conducts only liquid. During centrifugation, the increased body force causes water to exit the specimen through the ceramic while air enters the surface of the specimen. The suction profile within the specimen can be defined if the bottom boundary is maintained saturated (zero suction). The suction distribution obtained at equilibrium (i.e., when flow ceases) is:

$$\psi(r) = \frac{\rho_w \omega^2}{2} [r_0^2 - r^2] + \psi(0) \quad (34.6)$$

where r_0 is the outer radius of the centrifuge specimen, r is the distance toward the center of rotation with a datum at the outer radius, ω is the angular velocity, and $\psi(0)$ is the suction at the bottom boundary of the specimen (zero if a saturated ceramic plate is used as the bottom boundary condition). Analytical techniques can be used to associate the average moisture content (measured destructively) with the suction at the soil surface (Forbes, 1994).

A SWRC is typically quantified by fitting experimental data to power law, hyperbolic, or polynomial functions (Brooks and Corey, 1946; van Genuchten, 1980; Fredlund and Xing, 1994). Although the Brooks and Corey (1946) model is able to represent a sharp air entry suction, the van Genuchten (1980) model is most commonly used in numerical analyses because it is differentiable for the full range of suctions. The van Genuchten model is given by:

$$\theta = \theta_r + (\theta_s - \theta_r) [1 + (\alpha \psi)^N]^{-(1-(1/N))} \quad (34.7)$$

where θ_r is the residual moisture content, θ_s is the saturated moisture content (porosity), and α (units of kPa^{-1}) and N (dimensionless) are fitting parameters. Preliminary estimates of the SWRC could be obtained using databases that rely on the granulometric distribution of soils (Fredlund and Xing, 1994).

34.2.2.2 Hydraulic Conductivity Function

The relationship between hydraulic conductivity and suction, also referred to as the K -function, provides a measure of the increased impedance to moisture flow with decreasing moisture content. The saturated hydraulic conductivity K_s provides a measure of the minimum impedance to moisture flow through soil. Figure 34.6 shows K -functions for different geotechnical materials. Near saturation, the coarser materials (sand and geotextile) have high hydraulic conductivity, while the fine-grained materials (silt and clay) have lower hydraulic conductivity. However, as the soil dries, the coarse materials end up being less conductive than the fine-grained materials. That is, as the fine-grained materials can retain more water in the pores as suction increases, they still have available pathways for water flow, and are thus more conductive than coarser materials. The superior performance in arid climates of evapotranspirative covers relative to conventional resistive covers can be attributed to the lower unsaturated hydraulic conductivity of the selected cover soils.

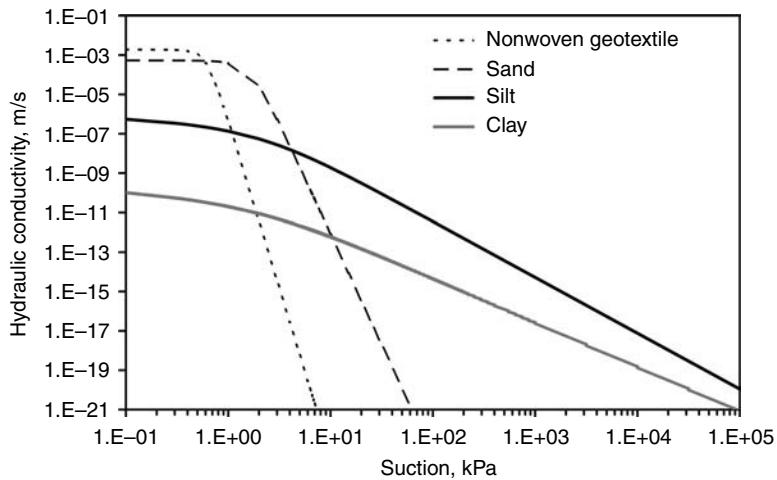


FIGURE 34.6 Typical K -functions for different geotechnical materials.

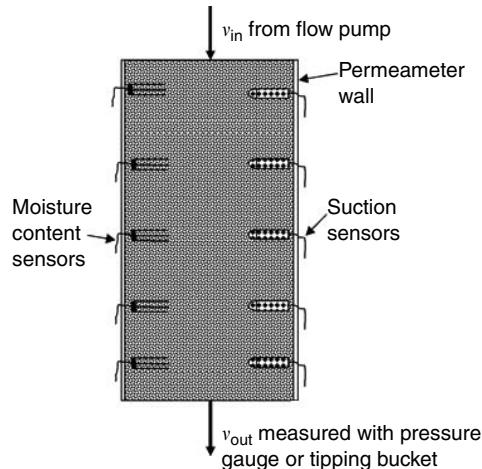


FIGURE 34.7 Parameter set up for measurement of K -function.

Conventional methods used to define the K -function are costly, time consuming, and prone to experimental error. Accordingly, the K -functions (e.g., such as those in Figure 34.6) are often predicted based on pore size distributions such as van Genuchten–Mualem model (1980), as follows:

$$K(\theta) = K_{\text{sat}} \sqrt{\frac{\theta - \theta_r}{\theta_s - \theta_r}} \left[1 - \left(1 - \left(\frac{\theta - \theta_r}{\theta_s - \theta_r} \right)^{1-N} \right)^{1/(1-N)} \right]^2 \quad (34.8)$$

$K(\psi)$ can be defined by substituting Equation 34.7 for θ into Equation 34.8. Other predictive relationships for the K -function are given by Burdine (1953) and Campbell (1974).

Several techniques have been proposed for direct determination of the K -function in the laboratory (Benson and Gribb, 1997). Conventional techniques to measure the K -function involve flow of liquid through a specimen confined within a permeameter. Flow is applied using either ceramic plates or flow pumps. Figure 34.7 shows a typical permeameter setup used to measure the hydraulic conductivity (Meerdink et al., 1996; Lu and Likos, 2005; McCartney et al., 2005b). Permeameters have differed in

specimen confinement and size, control of boundary conditions, and availability of instrumentation. The K -function can be estimated using steady or transient flow processes. During steady moisture infiltration with a deep water table, a unit hydraulic gradient (e.g., $i = 1$) is typically observed in the soil profile, sufficiently far from a water table boundary. Accordingly, suction does not change with depth and water flow is driven only by gravity. In this case, the hydraulic conductivity equals the imposed steady-state discharge velocity. Additional points are obtained by changing the imposed flow. During transient flow processes, the suction and moisture content profiles are measured as a function of depth and time, and the K -function can be estimated using the instantaneous profile method (Watson, 1966; Olson and Daniel, 1979; Meerdink et al., 1996). While techniques based on transient processes yield more information about the K -function, steady state techniques provide more reliable information. As for the SWRC, conventional techniques used to define the K -function require significant time to obtain limited data. Problems specific to K -function testing include boundary effects on the flow process, difficulties in uniformly distributing water from flow pumps to the specimen, and tedious testing procedures.

To alleviate these shortcomings, centrifuge testing has been used to define the K -function in geotechnical projects involving the design of ET covers (Zornberg et al., 2003; Dell'Avanzi et al., 2004). Nimmo et al. (1987) developed the Internal Flow Control Steady-State Centrifuge (IFC-SSC) method, which uses a system of reservoirs to control the fluid flow rate and suction at the upper and lower surfaces of a specimen. Conca and Wright (1994) developed the Unsaturated Flow Apparatus (UFA), which uses a sophisticated rotary joint to a low fluid flow rate into the specimen. The UFA uses open-flow centrifugation, which does not impose a suction value on the specimen. For steady state conditions, the SSC and UFA use Darcy's law to determine the K -function:

$$K(\psi) = \frac{-v}{[-(1/(\rho_w g))(d\psi/dz) - ((\omega^2 r)/g)]} \quad (34.9)$$

where v is the imposed discharge velocity. Points on the K -function curve are defined using Equation 34.9 after reaching steady state flow conditions. The SSC and UFA do not allow the direct monitoring of the relevant variables (suction, moisture, discharge velocity) in-flight during testing. If the suction gradient in Equation 34.9 is assumed to be negligible, the hydraulic conductivity becomes inversely proportional to ω^2 . The SSC and UFA centrifuges must be periodically stopped to measure the specimen mass to ensure steady state flow, and the moisture content must be measured destructively at the end of the test.

Although the SSC and UFA allow a faster testing time than conventional K -function testing methods, their shortcomings led to the development of an improved centrifuge device, referred to as the Centrifuge Permeameter for Unsaturated Soils (CPUS) (McCartney and Zornberg, 2005a). This device incorporates the use of a low-flow hydraulic permeameter and a high-g centrifuge capable of continuously, non-destructively, and non-intrusively measuring suction, moisture content, and fluid flow rate in a single specimen during centrifugation. Accordingly, CPUS allows an expedited determination of both the SWRC and K -function from a single specimen in a single test. Measuring the SWRC and K -function during a flow process is consistent with actual flow problems, unlike conventional techniques such as the axis-translation technique. Figure 34.8(a) shows a view of the CPUS centrifuge and Figure 34.8(b) shows a schematic view of the CPUS permeameter and its instrumentation layout. A special low-flow fluid union is used to supply fluid from the stationary environment to the rotating specimen within the centrifuge. CPUS facilitates the use of experimentally obtained, rather than theoretically derived hydraulic properties in the design of evapotranspirative cover systems.

34.3 Types of Evapotranspirative Covers

34.3.1 Monolithic Covers

Monolithic covers are evapotranspirative covers that consist of a single soil layer placed directly over the waste (Zornberg et al., 2003). Figure 34.9a shows a schematic view of a monolithic soil cover. The soil layer

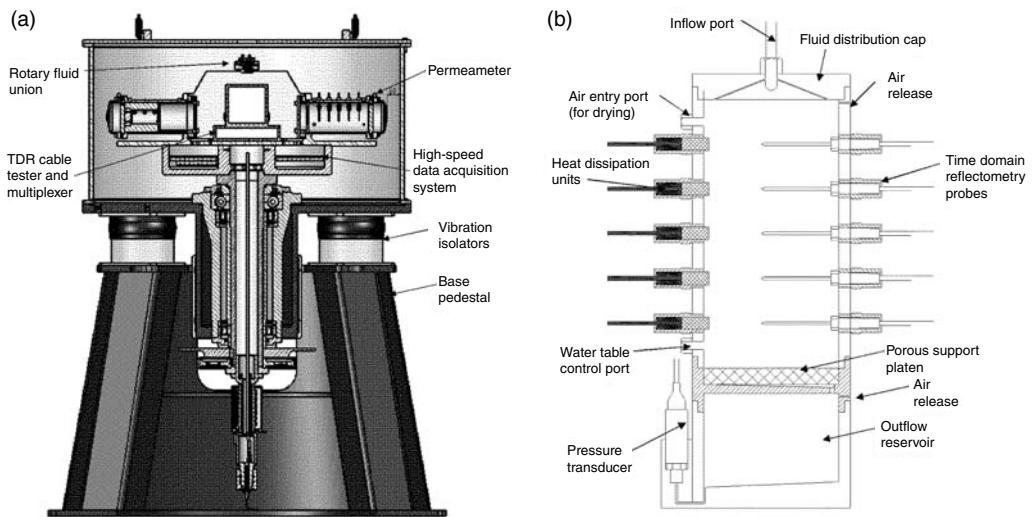


FIGURE 34.8 Centrifuge parameter for unsaturated soils: (a) centrifuge layout and (b) Permeameter.

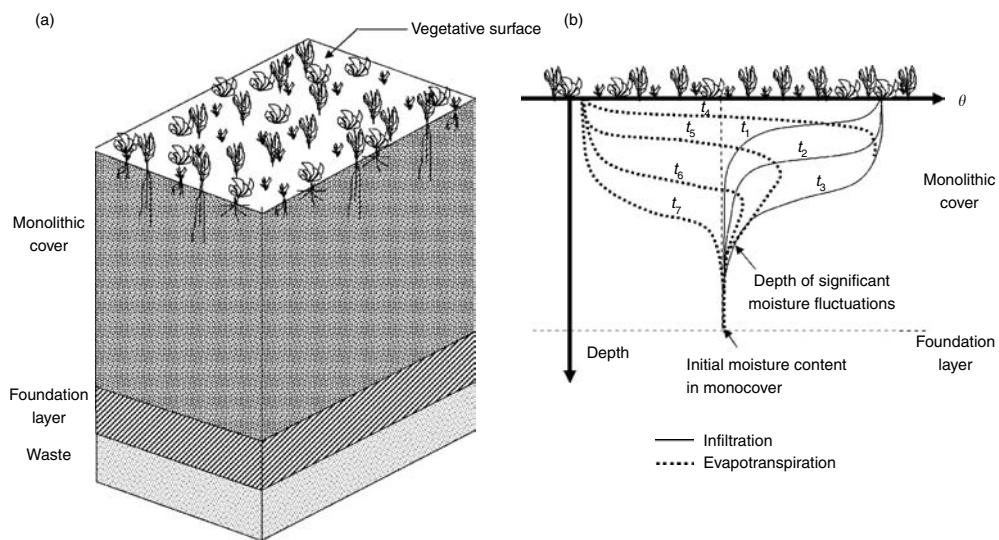


FIGURE 34.9 Monolithic cover: (a) soil profile and (b) typical seasonal moisture content fluctuations.

acts both as a substrate for vegetation and a hydraulic barrier. A foundation layer consisting of the same soil type is typically used to provide a level surface above the waste. Early research focused on investigation of the long-term behavior of natural soil layers in arid regions, assuming that the behavior is analogous to that of an engineered monolithic cover (Waugh et al., 1994). These studies found that moisture content fluctuations in natural analogues in recent geologic history are typically confined to the upper few feet of soil, indicating the adequacy of monolithic covers as an acceptable long-term solution to waste disposal.

The major aspects in monolithic covers are the proper characterization of the hydraulic properties (K -function and SWRC) of the soils as well as the determination of the appropriate thickness of the engineered soil cover. Figure 34.9b shows schematic moisture profiles, illustrating typical seasonal fluctuations in a properly performing monolithic cover. The moisture profiles illustrate wetting during infiltration events and subsequent drying due to evapotranspiration. Although some moisture fluctuations

are expected to reach the base of the cover in extreme events, most of the moisture fluctuations are expected to take place only within the upper portion of the cover. Monolithic cover design requires selection of the cover thickness and soil moisture storage necessary to keep the basal percolation below a minimum allowable (design) value, given the expected weather conditions for a site.

The soil moisture storage of a cover can be calculated as the integral of the volumetric moisture content profile with depth. The upper bound on soil moisture storage depend on the shape of the SWRC. The greater the moisture retained in a soil for the suction values expected in the field, the greater the moisture storage. A parameter that has been used to quantify the moisture storage is the field capacity, which is defined as the threshold moisture content value above which the soil no longer retains water by capillarity under the effects of gravity (Zornberg et al., 1999). When water is added to a soil that is at field capacity, drainage occurs. The field capacity may be obtained from infiltration tests, although a generally accepted value for silt and low plasticity clay soils is the moisture content corresponding to a suction of 33 kPa (Nachabe, 1998; Meyer and Gee, 1999). The storage capacity of a monolithic cover can be preliminarily estimated by multiplying the volumetric moisture content at field capacity (obtained using the SWRC) by the cover thickness.

34.3.2 Capillary Barriers

Capillary barriers are evapotranspirative covers that consist of a layered soil system typically involving a fine-grained soil (silt, clay) placed over a coarse-grained material (sand, gravel, nonwoven geotextile). Capillary barriers use the contrast in hydraulic properties between the fine- and coarse-grained soils to enhance the ability of the fine-grained material to store moisture (Shackelford et al., 1994; Stormont and Anderson, 1999; Khire et al., 1999; 2000). Figure 34.10a and Figure 34.10b show SWRCs and K-functions for sand (coarse-grained component), and low plasticity clay (fine-grained component). The capillary break concept relies on the continuity of suction, even at the interface between two different materials. Figure 34.10a shows that when a clay–sand system is at an initially high suction (e.g., 1000 kPa), the clay has a degree of saturation of 0.2 while the underlying sand is at residual moisture content. Figure 34.10b indicates that at this high suction, the clay has a hydraulic conductivity of approximately 1×10^{-13} m/sec while the sand has an even higher impedance to flow. Consequently, if moisture infiltrates into the fine-grained material after a precipitation event and reaches the interface with the coarse-grained material it can only progress into the coarse-grained material at a very slow rate. Consequently, water will accumulate at the interface until the suction at the interface reaches a value at which the hydraulic conductivity of the coarse grained material is no longer below that of the fine-grained material (3.0 kPa in Figure 34.10b). This suction value is referred to as the breakthrough suction. A breakthrough suction of 3.0 kPa is significantly below the suction corresponding to field capacity (typically considered at 33 kPa for clay), which indicates that the degree of saturation in the clay will be relatively high (95%) when breakthrough occurs. For suction values less than that corresponding to field capacity, water would have drained downwards by gravity had the capillary break not been present. When the breakthrough suction is reached, leakage is observed in the coarse-grained layer at a rate approaching that corresponding to the saturated hydraulic conductivity of the barrier layer.

Figure 34.10c shows a schematic view of a capillary break cover. Similar to the monolithic cover, the soil layer acts both as a substrate for vegetation and a hydraulic CRL (or capillary retention layer, CRL). However, a coarse-grained material (capillary break layer, CBL) is placed over the foundation material to create a capillary break at the interface between the CRL and the CBL). Figure 34.10d shows schematic moisture profiles illustrating the expected seasonal fluctuations in a properly performing capillary break cover. Unlike the monolithic cover, the moisture front can reach the bottom of the barrier layer without resulting in basal percolation, provided that the moisture at the interface does not exceed the breakthrough value. An important benefit of capillary breaks is that the moisture storage within the fine-grained soil can exceed its freely draining state (field capacity). Consequently, more water can be stored in a capillary break cover than in a monolithic cover of equivalent thickness. Alternatively, a thinner fine-grained layer can be used in a capillary break cover to obtain the same moisture storage as in a monolithic cover. The magnitude

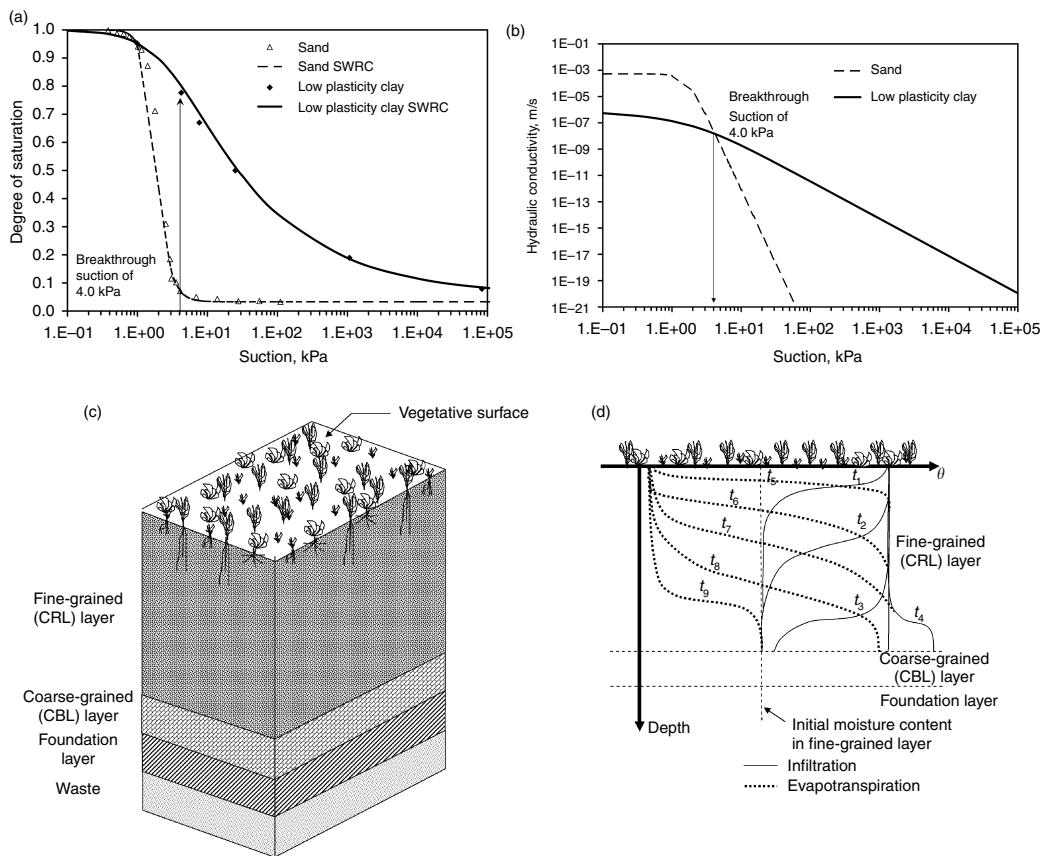


FIGURE 34.10 Capillary barrier details: (a) conceptual SWRC illustration of the capillary break effect; (b) conceptual K-function illustration of capillary break effect; (c) capillary break soil profile; and (d) typical seasonal moisture content fluctuations.

and long-term changes in the conditions leading to breakthrough are aspects under current investigation, especially in inclined covers (Parent and Cabral, 2004). Inclined covers may experience lateral diversion, which leads to greater moisture contents in the lower portion of a slope, resulting in a greater likelihood for breakthrough. Conservatism should be used in capillary barrier design, as these barriers have typically been reported to experience breakthrough during spring snowmelt when plant evapotranspiration is at a minimum (Khire et al., 1999; 2000). In addition, preferential flow through larger pores may lead to significant variability in the breakthrough suction (Kampf and Holfelder, 1999).

34.3.3 Anisotropic Barriers

Anisotropic barriers are similar to capillary barriers, but their design accounts for the internal lateral drainage through one or more drainage layers resulting from the inclination of the cover (Stormont, 1995; Bussiere et al., 2003; Parent and Cabral 2005). Figure 34.11a shows a schematic view of an anisotropic barrier. Anisotropic layers typically involve a soil vegetation substrate overlying a coarse-grained drainage layer, which are in turn underlain by a primary barrier layer and a second coarse-grained layer to provide a capillary break. The coarse-grained drainage layer functions both as a capillary break to the vegetation substrate and as a drainage layer for breakthrough water. The water collected by the drainage layer, along with moisture migrating laterally within the vegetation substrate is typically diverted to a ditch before a significant amount of water can infiltrate into the primary barrier. Figure 34.11b shows moisture profiles

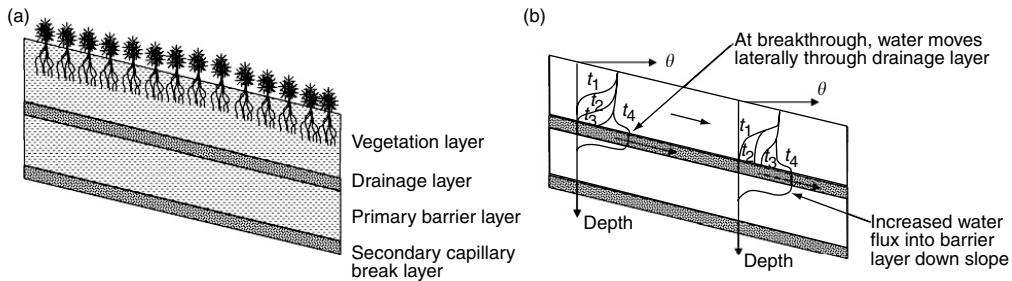


FIGURE 34.11 Anisotropic barrier: (a) soil profile and (b) typical moisture content fluctuations during infiltration.

illustrating the expected trends along the length of an anisotropic barrier during a wet season. The capillary break at the drainage layer increases the moisture storage capacity of the vegetative substrate. Infiltration of water into the barrier layer generally occurs toward the toe of the slope due to water accumulation from infiltration and lateral drainage. Design of anisotropic covers is more complex than that of monolithic or capillary break covers due to the need to quantify the required hydraulic properties for the layered profile as well as the volume of laterally diverted water. Field comparisons between the performance of test-scale capillary break, anisotropic, and monolithic covers performed by Dwyer (1998) indicate that the anisotropic cover performed well compared to the other covers for the same weather conditions over a five-year monitoring program. However, its construction was the most difficult of the three covers.

34.4 Relevant Issues for Design of Evapotranspirative Covers

34.4.1 Design Strategies

Despite the conceptual simplicity of evapotranspirative covers, their design is not trivial and typically involves numerical and field demonstration components. The design of an evapotranspirative cover involves the identification of performance criteria that are used to evaluate the suitability of different design alternatives. This phase typically requires close interaction between the designer, site stakeholders, and local regulators. The goal is to identify the criteria that will provide for human safety, constructability, and cost effectiveness. The design process typically involves identification of different cover design alternatives. This includes: (i) identification of suitable local borrow soil sources, (ii) characterization of the hydraulic properties of pre-selected soils under different placement conditions, (iii) determination of the cover geometry, (iv) identification of site-specific critical meteorological conditions, and (v) identification of suitable plant communities and vegetative cover properties. Design typically involves use of numerical models to predict the cover performance and field demonstrations aimed at evaluating the different design alternatives. Performance monitoring programs have been typically implemented after construction to verify the adequacy of the selected design alternative. Recent developments in evapotranspirative cover design have employed decision analysis, statistical characterization of design variables, and cost estimations to identify the optimal combination of performance variables. Such design approaches may help optimize the collection of laboratory data, identify if a site-specific field testing program is necessary, and quantify the risk of failure associated with design alternatives.

34.4.2 Performance Criteria and Regulatory Issues

Evapotranspirative covers are considered alternative covers and, as such, their design involves comparison of their performance with that of prescriptive cover systems (i.e., equivalency demonstration). Development of suitable performance criteria for ET cover systems generally involves equivalence demonstration with prescriptive covers (Morris and Stormont, 1997; McCartney and Zornberg, 2002; Albright et al., 2004). Because of the site-specific interactions between an evapotranspirative cover and the

local climate and environment, performance criteria for evapotranspirative covers must account for site-specific conditions. This section outlines different types of performance criteria that have been put forth for evapotranspirative covers. The type of performance criteria will impact both the design procedures and the methods of post-closure monitoring.

Determination of a percolation criterion has been approached from two different perspectives, at least in recent evapotranspirative cover projects in the United States. The first involves defining a *quantitative* maximum value of basal percolation (e.g., expressed in mm/year) that should not be exceeded. The second involves a *comparative* approach aiming at achieving a basal percolation value in the evapotranspirative cover that is smaller than that through a prescriptive cover under the same weather conditions.

The maximum basal percolation value (quantitative percolation criterion) is typically defined in agreement with regulatory agencies, yet it should be based on actual performance data from prescriptive type covers or on the results of verified numerical models. A quantitative criterion establishes that the basal percolation through the evapotranspirative cover, P_e (mm/year), should be less than the Maximum Quantitative Percolation Value, MQPV (e.g., mm/year), deemed to satisfy equivalence, as follows:

$$P_e < \text{MQPV} \quad (34.10)$$

where MQPV is a predefined criterion and P_e is determined from field monitoring and from numerical simulations. The choice of a quantitative percolation criterion implies that the cover must be designed for a wide range of possible meteorological conditions to ensure fulfillment of the criterion for worst case scenarios. On the other hand, if the MQPV is selected conservatively, the cover design may lead to unrealistically high material and construction costs.

A comparative percolation criterion for evapotranspirative covers involves defining the maximum ratio between the basal percolation through an evapotranspirative cover and that through a prescriptive cover. This percolation criterion recognizes that the performance of the evapotranspirative cover should be compared to that of a resistive cover under the same meteorological conditions. Specifically, the basal percolation through the evapotranspirative cover (P_e) should be less than that through the prescriptive cover (P_p), affected by the Maximum Comparative Percolation Ratio (MCPR) established for the project. That is, equivalence is deemed satisfied when the following condition is fulfilled:

$$P_e < \text{MCPR} \cdot P_p \quad (34.11)$$

where MCPR is dimensionless and P_e and P_p are basal percolation values (e.g., in mm/year). Adopting an MCPR of 1.0 implies that the evapotranspirative cover should perform better than a prescriptive cover under the same weather conditions. The methods used to define basal percolation values P_e and P_p may involve the use of a numerical model or of monitored field test plots. The first approach may not account for factors that are not adequately modeled using unsaturated flow models such as desiccation, surface settlement, animal burrowing, and fingering. The second approach may result in additional costs and may not account for critical weather conditions and long-term factors.

Performance criteria may also involve interpretation of the moisture profiles in the cover rather than quantification of percolation values. Specifically, quantitative criteria can involve comparison of the moisture storage within the cover with a certain threshold value. Similarly, criteria may involve comparison of the moisture at the base of the cover with a certain threshold moisture value (that can in turn correspond to a maximum value of $K(\theta)$ calculated from Equation 34.8). Ideally, performance criteria would involve interpretation of both basal percolation and moisture profile information (McCartney and Zornberg, 2002).

34.4.3 Important Design Variables

34.4.3.1 Soil Hydraulic Properties

The soil hydraulic properties relevant to the performance of evapotranspirative covers include the saturated hydraulic conductivity (K_s), the soil-water retention curve (SWRC), and the K -function.

Determining the required soil hydraulic properties is a challenging aspect of evapotranspirative cover design. Target values of the hydraulic properties are typically selected based on the required cover performance and on the expected weather conditions at the site. An important difficulty in the methods used to evaluate hydraulic properties in the laboratory is their high sensitivity to sample preparation and testing procedures. For example, the relative compaction of the soil and its compaction moisture content will typically affect the soil hydraulic properties. Finally, the differences between hydraulic properties determined in the laboratory and those in the field add to the difficulty in soil characterization. In addition to specimen preparation variability, the hydraulic properties in the field may change with time due to hysteresis (different behavior during wetting and drying), soil cracking, settlement, plant or animal intrusion, and erosion.

34.4.3.2 Cover Geometry

Cover geometry variables include the thickness of the cover, as well as the cover inclination. As the performance of the soil cover depends on its capacity to store liquid, a thicker cover will have a higher moisture storage capacity. The slope of the cover will impact its stability, the amount of water drained laterally, as well as the amount of overland runoff during precipitation events. Some evapotranspirative covers have been constructed with significantly high slope inclinations (as high as 1.5H:1V). This is the case of the OII landfill in Pomona, California (see Section 34.6.2). Steep slopes may require stabilization to account for seismic and static stability.

34.4.3.3 Critical Meteorological Variables

Critical meteorological conditions are site specific and play a significant role in the selection of the required soil properties, cover thickness, and slope inclination. A general guideline is that evapotranspirative covers are suitable for regions in which the potential evapotranspiration is greater than the precipitation. An important aspect of the design process involves identification of a worst-case precipitation record. Multi-year records must be obtained to consider the patterns in cover performance over both wet and dry years. In addition to the precipitation, the solar radiation, minimum and maximum air temperatures, wind speed, relative humidity, and percentage cloud cover are often used in numerical models to calculate the potential evapotranspiration. These variables are generally available from local or national weather station databases.

34.4.3.4 Vegetation Parameters

Vegetative cover properties determine the amount of water that can be removed from the cover by transpiration. Water uptake by plants has been quantified empirically as a function of variables such as the wilting point suction, leaf area index, evapotranspiration partitioning model parameters, and root density distribution. Although researchers have been incorporating plant behavior in numerical models since the 1950s, the quantification of these variables is still unclear. The wilting point suction is the suction above which the plants are unable to draw water from the soil. Plants will typically stop transpiring and grow dormant under such conditions. The leaf area index is the leaf area coverage per unit ground area, and is typically correlated with the moisture demand of plants. A plant with more foliage will be able to photosynthesize faster, leading to more transpiration. Evapotranspiration can be estimated using the Penman-Monteith model (Monteith, 1965), which requires knowledge of the daily maximum and minimum temperatures, dew point temperature, solar radiation, percent cloud cover, wind speed, and precipitation (Zornberg and McCartney, 2003). However, once the evapotranspiration has been estimated, it must be partitioned into evaporation and transpiration components for use in numerical models (Fayer et al., 1992). The Ritchie model can be used to correlate the relative contribution of the plant transpiration with the leaf area index. The root density distribution depends on the particular plant, although it is typically measured to avoid plants with roots deeper than the cover thickness. A diverse group of native shrubs and grasses is recommended. Selection

of plants with different heights also prevents intrusion of burrowing animals while providing erosion control.

34.4.4 Numerical Modeling Issues

Analysis of the performance of alternative covers has been performed using: (i) unsaturated flow analyses, and (ii) simplified water balance analyses (Albright et al., 2002). Unsaturated flow analysis entails solving Richards' equation for certain surface boundary conditions (e.g., water infiltration, overland runoff, evaporation, transpiration) and bottom boundary conditions (e.g., unit hydraulic gradient, seepage face). The governing equations are solved using numerical techniques such as the finite element method or the finite difference method. Relevant outputs from these analyses include the transient moisture redistribution and basal percolation. Four commonly used Richards' equation-based codes that have been used in the analysis of evapotranspirative covers are UNSAT-H (Fayer and Jones, 1990), HYDRUS-1D (Simunek et al., 1998), LEACHM (Hutson and Wagenet, 1992), and Vadose/W (GeoSlope, 2004). The current version of UNSAT-H allows modeling of hysteresis in the SWRC, and was observed to have a close fit to observed data (Khire et al., 1999). HYDRUS-1D is more user-friendly and has built-in libraries for soil hydraulic properties and initial estimates for most performance variables. LEACHM has the distinct advantage of being an open-source code, while being particularly useful for conducting sensitivity analyses that allow identification of relevant variables (Zornberg et al., 2003). VADOSE/W is a commercial code that considers fully coupled heat and mass transfer with vapor flow, vegetation, gas diffusion and ground freezing in a finite element formulation in one or two dimensions. It solves a modified form of the Richards' equation with pressure as the dependent variable, not water content, which improves solution stability. Despite their ability to solve Equation 34.5 subject to complex boundary conditions, it is still challenging to use Richards' equation-based codes for predicting small basal percolation values with a high level of accuracy. It should be noted that the mass balance errors for numerical models may be of the same order of magnitude as basal percolation values of relevance for the analysis of evapotranspirative covers.

Analyses involving water balance use the conservation of mass at the soil surface to estimate the basal percolation through the cover. This approach treats the soil layer as a sponge, able to hold a certain maximum amount of water (at the field capacity suction) against the pull of gravity. Basal percolation is calculated as the volume of moisture in the cover that exceeds the field capacity moisture storage after accounting for moisture removed by evapotranspiration and lateral drainage. Water balance codes that have been used for evapotranspirative cover include the Hydrologic Evaluation of Landfill Performance (HELP) (Schroeder et al., 1994), which is distributed by the US Environmental Protection Agency (EPA), and the Erosion Productivity Impact Calculator (EPIC) (Williams et al., 1984). A significant limitation of the water balance approach is that it does not consider transient moisture redistribution, which plays an important role in moisture removal from the cover by evapotranspiration. Also, as the water balance approach considers that the only driving mechanism for water flow is gravity, it implicitly considers that the hydraulic gradient is always equal to 1.0.

34.5 Performance Monitoring of Evapotranspirative Covers

34.5.1 Monitoring Strategies

In the past decade, there has been a significant effort to expand the evapotranspirative cover knowledge base by constructing full-scale field test plots involving the monitoring of basal percolation and moisture profiles (Benson et al., 1999, 2001; Dwyer, 1998; O'Kane et al., 1998; Zornberg and McCartney, 2003). Monitoring schemes allow direct quantification of the response of evapotranspirative covers to atmospheric conditions. The field monitoring program should be consistent with the performance criteria used for the cover design. Different technologies have been considered to evaluate the basal percolation, moisture profiles, suction profiles, and meteorological variables, as discussed further in the following sections.

34.5.2 Regulatory Issues Specific to Compliance Demonstration

The selection of a quantitative or comparative percolation criterion for an evapotranspirative cover project (Section 34.4.2) may have significant implications on the compliance demonstration for the cover. As mentioned, the threshold basal percolation (MQPV) to be used in a quantitative percolation criterion is difficult to define. However, once the MQPV has been defined the monitoring program used for compliance demonstration phase using a quantitative percolation criterion is reasonably straightforward, as it involves monitoring the basal percolation to verify that it remains below the MQPV. Yet, limitations in field monitoring and numerical modeling must be understood to correctly interpret the data collected for compliance demonstration. On the other hand, the monitoring program used for compliance demonstration when using a comparative percolation criterion for the design may not be straightforward. For example, an approach could involve continued comparison monitored in the basal percolation values of two covers (prescriptive and alternative) over the lifetime of the evapotranspirative cover. Accordingly, the time period used for compliance demonstration is a key factor. For instance, compliance could be based on either the average basal percolation value during a time period or on the maximum basal percolation amounts associated with specific precipitation events.

34.5.3 Lysimetry

Gravity lysimeters are the most common tools for directly monitoring basal percolation. They typically consist of a drainage layer underlain by a hydraulic barrier (Benson et al., 1999, 2001), such as a geocomposite or gravel drainage layer placed on top of a geomembrane. The geocomposites used in gravity lysimeters consist of a geonet sandwiched between two nonwoven geotextiles. A geonet is a thin polymeric sheet with slotted grooves that provide high lateral transmissivity. Geotextiles are polymeric fabrics used as filters, protection layers, and drainage layers. Geomembranes are polymeric sheets that have equivalent hydraulic conductivity values on the order of 10^{-15} m/sec. When a soil cover is placed above the lysimeter, it is intended that percolation through the cover soil will reach the geocomposite, and be transmitted down-slope to a collection bin. For effective performance, it is important to avoid that the presence of the lysimeter interferes with water flow (i.e., by introducing a capillary break).

The main advantage of lysimeters is that they can be constructed to cover large areas (10 to 100 ft on side), which allows a spatial averaging of the water flow through a potential system of saturated preferential pathways in the cover (areas with lower density, cracks, animal burrows, or decayed plant roots). However, lysimeters have several shortcomings, the most significant being that they provide little insight into reasons for poor or adequate cover performance. Another shortcoming is that, despite their high transmissivity and permittivity when saturated, the geotextile component of a lysimeter may cause a capillary break when in contact with unsaturated soils (Stormont and Morris, 2000; Zornberg and McCartney, 2003; McCartney et al., 2005b). A capillary break at the lysimeter–soil interface distorts the suction and moisture content profiles in an evapotranspirative cover and may lead to significant underestimation of basal percolation. In addition, lysimeters create barriers to upward flow from lower depths (Scanlon and Milly, 1994). This is typically caused by water vapor flow in response to thermal gradients, and may lead to overestimation of the basal percolation at the site.

34.5.4 Monitoring of Moisture and Suction Profiles

As the overall performance of an evapotranspirative cover system relies on its ability to store moisture until it is removed by evapotranspiration, moisture content or suction profiles can be monitored to assess why the evapotranspirative cover performs adequately (or not). Continued monitoring of *in situ* soil volumetric moisture content profiles is important in many geoenvironmental engineering and hydrological projects. In particular, monitoring of soil volumetric moisture content can provide relevant feedback on the migration of moisture through evapotranspirative covers.

Time Domain Reflectometry (TDR) technology has been used to measure the volumetric moisture content in evapotranspirative covers (Dwyer, 1998; O’Kane et al., 1998; Siddiqui et al., 2000; Albright and

Benson, 2002; Zornberg and McCartney, 2003). TDR involves measuring the velocity of an electromagnetic pulse applied to a transmission line that terminates in a shielded probe placed within the soil mass (Topp et al., 1980; Siddiqui et al., 2000; Suwansuwat and Benson, 1999). The pulse is reflected due to changes in impedance along the transmission line-probe system (e.g., the beginning of the probe and the end of the probe). The velocity of the reflected pulse is affected by the dielectric constant of the water within the soil mass, which is an order of magnitude greater than that of air and soil particles. The bulk dielectric constant of the soil mass, calculated from the velocity of the reflected pulse, can then be correlated with the soil volumetric moisture content. Conventional TDR systems use a cable tester to generate high frequency electromagnetic pulses (~10 to 100 GHz) and measure the reflected waveform (with a resolution on the order of nanoseconds). Although conventional TDR systems are generally adequate for a wide range of soil types, their accuracy decreases for high moisture contents, saline soils, and highly conductive clays. Their shortcomings include the relatively high cost of probes and cable tester systems and comparatively complicated installation procedures to prevent probe damage.

Water content reflectometer (WCR) probes have often been used as an alternative to conventional TDR probes (Dwyer, 1998; Albright and Benson, 2002; Chandler et al., 2004). WCR probes infer the moisture content by measuring the dielectric content of the soil, similar to TDRs. However, WCRs use smaller electronic circuitry placed within the probe itself that generate a lower frequency electromagnetic pulse (~40 MHz) (Chandler et al., 2004). WCR probes have lower power requirements and allow longer cable lengths than conventional TDRs. In addition, WCR probes can use conventional field dataloggers instead of cable testers, which makes them attractive for field applications. Despite these advantages, the use of comparatively low frequencies may lead to decreased resolution in volumetric moisture content measurements and to correlations that are more sensitive to the soil electrical conductivity and temperature than TDRs (Kim and Benson, 2002).

Suction measurements can also be made to complement moisture content evaluations. Concurrent suction and moisture content monitoring can provide data suitable to interpret evapotranspirative cover performance. Specifically, the suction and moisture content measurements can provide information for *in situ* determination of the SWRC. This can be used to interpret the SWRC during wetting and drying, interface phenomena like capillary breaks, and optimization of the SWRC to be used in numerical models. Suction measurements have been conducted in the field using several types of techniques, such as tensiometers (Ridley and Burland, 1995; Sisson et al., 2002; Tarantino and Mongiovì, 2003) and heat dissipation units (HDUs) (Flint et al., 2002; Nichol et al., 2003). Tensiometers consist of a pressure transducer with a porous ceramic stone filter. Due to continuity of suction, the suction within the ceramic stone will come into equilibrium with that of the surrounding soil. As water is drawn from the ceramic stone, the pressure transducer will measure a negative pressure value. HDUs consist of a heating unit and a thermocouple placed within a ceramic stone. The HDU infer the soil suction by applying a heat pulse to the heating unit in the ceramic stone, and measuring the transient changes in temperature of the ceramic stone using the thermocouple. The thermal properties of the ceramic are related to its moisture content, with a wet ceramic having higher thermal conductivity than a dry ceramic. Accordingly, the thermal response of the ceramic to the applied heat pulse can be calibrated against the suction in the ceramic, which is the same as the suction in the soil.

34.5.5 Monitoring of Meteorological Variables and Overland Runoff

Site-specific monitoring of meteorological variables is important to proper interpretation of evapotranspirative cover performance, as these variables vary on a regional and local scale. Precipitation is generally measured using tipping bucket rain gauges. These gauges collect water in a two-sided bucket with a capacity of approximately 4 ml. The bucket is placed on a pivot so that water flowing into the gauge is funneled into one side of the bucket. After the capacity of the bucket is reached, the weight of the water causes the bucket to tip on the pivot, spilling the water and allowing water to be funneled into the other side of the bucket. The gauge records the time of each tip. Overland runoff has been typically measured using geomembrane swales anchored on the soil surface on the perimeter of the cover. Difficulties have

been encountered in the use of such swales due to thermal expansion of the exposed geomembranes. Other variables typically measured on the site include wind speed and direction using an anemometer, temperature, and solar radiation.

34.6 Case Studies

34.6.1 Site Survey

Table 34.1 shows a survey of evapotranspirative covers in the United States that have implemented field monitoring programs. A total of 53 evapotranspirative covers were identified at 48 sites in the United States. Among these, 19 of the sites have undergone or are undergoing construction or closure of an evapotranspirative cover, while the rest involve compliance demonstration using test covers. The number of sites with performance monitoring shown in this table reflects the growing number of alternative cover construction projects because of their technical and economic benefits. The table also provides the water balance variables monitored at each site and the study period. Although information is limited for some of the private covers, it appears that most covers have used lysimetry and moisture profile monitoring as part of compliance demonstration programs. Two of the sites, the OII landfill in Monterey Park, California, and the Rocky Mountain Arsenal, near Denver, Colorado, are discussed next.

34.6.2 OII Landfill

The first evapotranspirative cover system approved by the US EPA for a hazardous waste Superfund site was in 1997 (Zornberg et al., 2003). The cover was constructed at the OII Superfund site in Monterey Park, California, approximately 16 km east of downtown Los Angeles. The design of this cover involved five phases that were undertaken to define the cover layout configuration. The design phases included: (i) evaluation of a baseline evapotranspirative cover, (ii) equivalence demonstration of the baseline cover by comparison with the basal percolation through a prescriptive cover, (iii) evaluation of the sensitivity of different design parameters (e.g., cover thickness, soil characteristics, rooting depth, and potential use of irrigation schemes) on the basal percolation through the cover, (iv) use of compilation of the results of this analysis as basis for the design of the final evapotranspirative cover, and (v) equivalence demonstration using soil-specific hydraulic properties of each cover soil.

The criterion used for the design of the cover system at the OII Superfund site was comparative, and required that the basal percolation through the proposed evapotranspirative cover be less than the basal percolation through a prescriptive, resistive cover. That is, the Maximum Comparative Percolation Ratio (M CPR) at this project was 1.0. The prescriptive cover, defined by a consent decree, consisted of a 1200-mm thick system, which included a 300-mm thick vegetative layer, a 300-mm thick clay layer, and a 600-mm thick foundation layer. The vegetative and foundation layers both were considered to have a saturated hydraulic conductivity of 1×10^{-6} m/sec, and the clay layer to have a saturated hydraulic conductivity of 1×10^{-8} m/sec.

A laboratory testing program was implemented to characterize the candidate borrow soils using soil specimens remolded under different compaction and moisture conditions. The experimental program included determination of hydraulic, shear strength, desiccation potential, and edaphic properties. Table 34.2 summarizes the laboratory test results for one of the candidate borrow soils measured in the laboratory for use in the equivalence demonstration. Following identification of the candidate soil borrow sources and determination of their hydraulic properties, soil-specific equivalence demonstrations were performed for the proposed evapotranspirative cover. Soil-specific parameters used in the unsaturated flow analyses included the SWRC, saturated hydraulic conductivity and K -function (obtained by centrifugation). In addition, soil-specific information from compaction tests was used in the analyses to define the soil placement conditions (unit weight and moisture content) that would optimize the performance of the cover.

TABLE 34.1 Survey of ET Cover Case Studies

Site name	Cover type	Cover location	Cover details				Water balance variables measured				Study period		
			Meteorological variables	Percolation	Water content	Suction	Runoff	Start date	End date	Full scale or test section			
Greater Wenatchee Regional Landfill ^a	Capillary break	Wentachee, WA	X	X	X	X	X	1992	1995	0 test sections			
Greater Wenatchee Regional Landfill ^a	Monolithic	Wentachee, WA	X	X	X	X	X	1992	1995	1 test section			
Live Oak Landfill ^a	Monolithic	Atlanta, GA	X	X	X	X	X	1993	1996	1 test section			
Sacramento site ^a	Monolithic	Sacramento, CA	X	X	X	X	X	1999	2002	2 test sections			
Beltsville facility ^a	Monolithic	MD	X	X	X	X	X	?	?	Test section			
Phelan landfill ^a	Monolithic	Sacramento, CA	X	X	X	X	X	?	?	Test section			
National Council for Air and Stream Improvement ^a	Monolithic	NM	X	X	X	X	X	?	?	Test section			
Gredé Foundries ^a	Monolithic	Reedsburg, WI	X	X	X	X	X	1995	1996	Test section			
Texas Low-Level Radioactive Waste Disposal Authority Facility ^a	Monolithic	Sierra Blanca, TX	X	X	X	X	X	1995	1997	Test section			
US Marine Corps Air and Ground Combat Center ^a	Monolithic	Twentynine Palms, CA	X	X	X	X	X	1998	1999	Test section			
Marine Corps Base ^a	Monolithic	HI	X	X	X	X	X	?	?	Test section			
US Ecology Co. Landfill ^a	Monolithic	Sheffield, IL	X	X	X	X	X	?	?	?			
Hamburg Test Site	Capillary break	Hamburg, Germany	X	X	X	X	X	1988	1992	1 test section			
Omega Hills Landfill ^a	Capillary break	Milwaukee, WI	X	X	X	X	X	?	?	1 test section			
Los Alamos National Laboratory ^a	Capillary break	Los Alamos, NM	X	X	X	X	X	1991	1995	1 test section			
Los Alamos National Laboratory ^a	Capillary break	Los Alamos, NM	X	X	X	X	X	1991	1995	2 test sections			
Arid Lands Ecology Reserve, Hanford ^a	Monolithic	Richland, WA	X	X	X	X	X	1985	1990	4 test sections			
Hanford Field Lysimeter Test Facility ^a	Capillary break	Richland, WA	X	X	X	X	X	1987	1996	8 test sections			
Hanford ^a	Capillary break	Richland, WA	X	X	X	X	X	?	?	2 test sections			
Hanford ^a	Monolithic	Richland, WA	X	X	X	X	X	1987	1989	6 test sections			
NMSU Las Cruces Site ^a	Monolithic	Las Cruces, NM	X	X	X	X	X	?	?	1 test section			
Beatty Site ^a	Monolithic	Mojave Desert, NV	X	X	X	X	X	?	?	1 test section			
Hanford Site ^a	Monolithic	Richland, WA	X	X	X	X	X	?	?	1 test section			
Hill AFB ^a	Capillary break	UT	X	X	X	X	X	1990	1993	2 test sections			

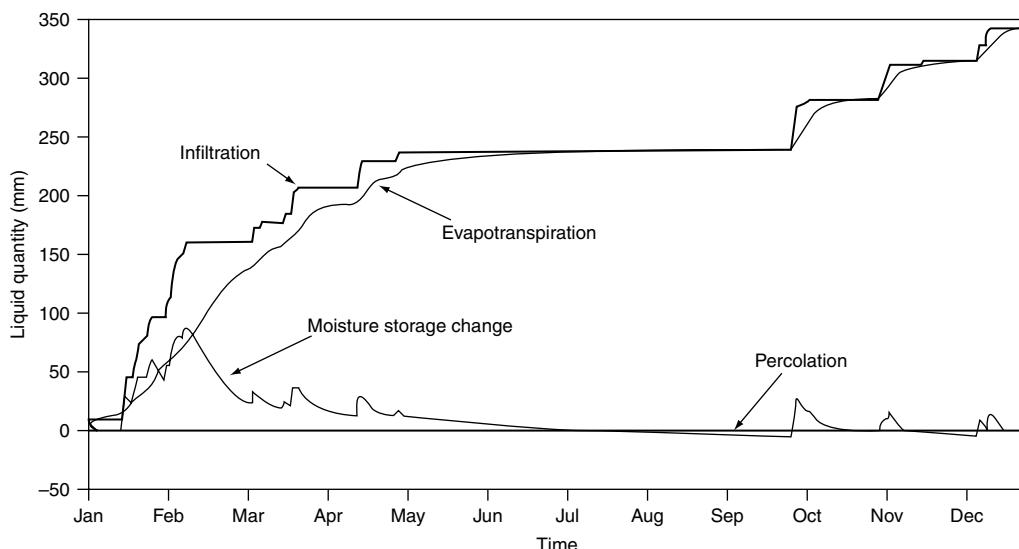
Sandia National Laboratories	Capillary break	Albuquerque, NM	X	X	X	X	1997	2002	1 test section
Sandia National Laboratories	Monolithic	Albuquerque, NM	X	X	X	X	1997	2002	1 test section
Sandia National Laboratories	Anisotropic	Albuquerque, NM	X	X	X	X	1997	2002	1 test section
Rocky Mountain Arsenal ^a	Monolithic	Denver, CO	X	X	X	X	1992	2003	4 test sections
Nevada test site	Monolithic	NV	X	a	X	X	1996	2002	2 test sections
Idaho National Engineering and Environmental Laboratory ^a	Monolithic	ID	X	X	X	X	1984	1995	1 test section
El Paso Site	Capillary break	El Paso, TX	X	X	X	X	2000	2001	1 test section
Questa Mine	Monolithic/ Oxygen cover	NM	X	X	X	X	2000	2001	1 test section
Equity Silver Mine	Monolithic/ Oxygen cover	BC, Canada	X	X	X	X	1992	1995	1 test section
Vereeniging Ash Dump	Ash Dump cover	South Africa	X	X	X	X	1995	1995	1 test section
F.R. Bowerman Landfill	Monolithic	CA	X	X	X	X	?	?	Full scale
Operating Industries, Inc. (OI)	Monolithic	Monterey Park, CA	X	X	X	X	?	?	Full scale
Puente Hills	Monolithic	CA	X	X	X	X	?	?	Full scale
Yucaipa	Monolithic	Orange County, CA	X	X	X	X	?	?	Full scale
Lopez Canyon	Monolithic	Los Angeles, CA	X	X	X	X	?	?	Full scale
Yeremo	Monolithic	Los Angeles, CA	X	X	X	X	?	?	Full scale
Riverside Co.	Monolithic	Riverside Country, CA	X	X	X	X	?	?	Full scale
29 Palms Marine Base	Monolithic	CA	X	X	X	X	?	?	Full scale
Potrero Hills	Monolithic	CA	X	X	X	X	?	?	Full scale
Chiquita Canyon	Monolithic	CA	X	X	X	X	?	?	Full scale
Needles Landfill	Monolithic	CA	X	X	X	X	?	?	Full scale
Fairmead Landfill	Monolithic	CA	X	X	X	X	?	?	Full scale
Rocketyne Site	Monolithic	Chatsworth, CA	X	X	X	X	?	?	Full scale
China Grade Landfill	Monolithic	Kern Country, CA	X	X	X	X	?	?	Full scale
McPherson Area Solid Waste Utility	Monolithic	McPherson, KS	X	X	X	X	?	?	Full scale
Ft. Carson Lakeside Reclamation Landfill	Monolithic	CO	X	X	X	X	?	?	Full scale
MSW Landfill Duvall Custodial Landfill	Monolithic	Beaverton, OR	NE	WA	X	X	?	?	Full scale

^a ACAP Sites.

TABLE 34.2 Hydraulic and Geotechnical Properties for OII Landfill Cover Soils

Series	USCS classification (ASTM D2487)	Average % passing #200 sieve (% fines)	Atterberg limits		Saturated hydraulic conductivity (m/s)	Campbell model parameters		Relative compaction ^a (%)
			Average PL (%)	Average LL (%)		a	b	
T1	CL	66	43	18	2.80E-06	-4.89	7.028	93.9
T2	CL	66	43	18	1.10E-05	-4.89	6.328	87.2
T3	CL	66	43	18	3.70E-05	-4.89	5.495	83.1
T4	CL	66	43	18	3.30E-06	-4.89	7.278	88.5
T5	CL	66	43	18	1.70E-05	-4.89	6.463	87.8
T6	CL	66	43	18	1.90E-04	-4.86	6.678	77.7

^a In relation to standard Proctor test (ASTM 698).

**FIGURE 34.12** Seasonal variation in the water balance variables (Zornberg et al. 2003).

The code LEACHM (Hutson and Wagener, 1992) was used to predict the basal percolation through both the prescriptive and the evapotranspirative covers to compare their performance using site-specific soil and meteorological conditions. LEACHM uses Campbell's (1974) equation to describe the relationship between suction and volumetric moisture content for soil:

$$\psi = a(\theta/\theta_s)^b \quad (34.12)$$

where a and b are constants obtained from curve fitting. The a and b values as well as the saturated hydraulic conductivity for one of the candidate evapotranspirative cover soils are listed in Table 34.2. The simulated water balance variables for the OII site are shown in Figure 34.12. This figure indicates that the moisture storage of the cover increases during periods of infiltration and subsequently decreases during period of evapotranspiration. The infiltration is typically higher in the early part of the year, while the evapotranspiration is higher in the late spring and summer months. Figure 34.13a shows equivalence results, as quantified by the percolation ratio (basal percolation through the ET cover divided by the basal percolation through the prescriptive cover) with time. This equivalence demonstration shown in the figure corresponds to an evapotranspirative cover system constructed using soils placed under compaction conditions defined by series T1 (Table 34.2). The percolation ratio is below 0.1 (and well below the MCPR

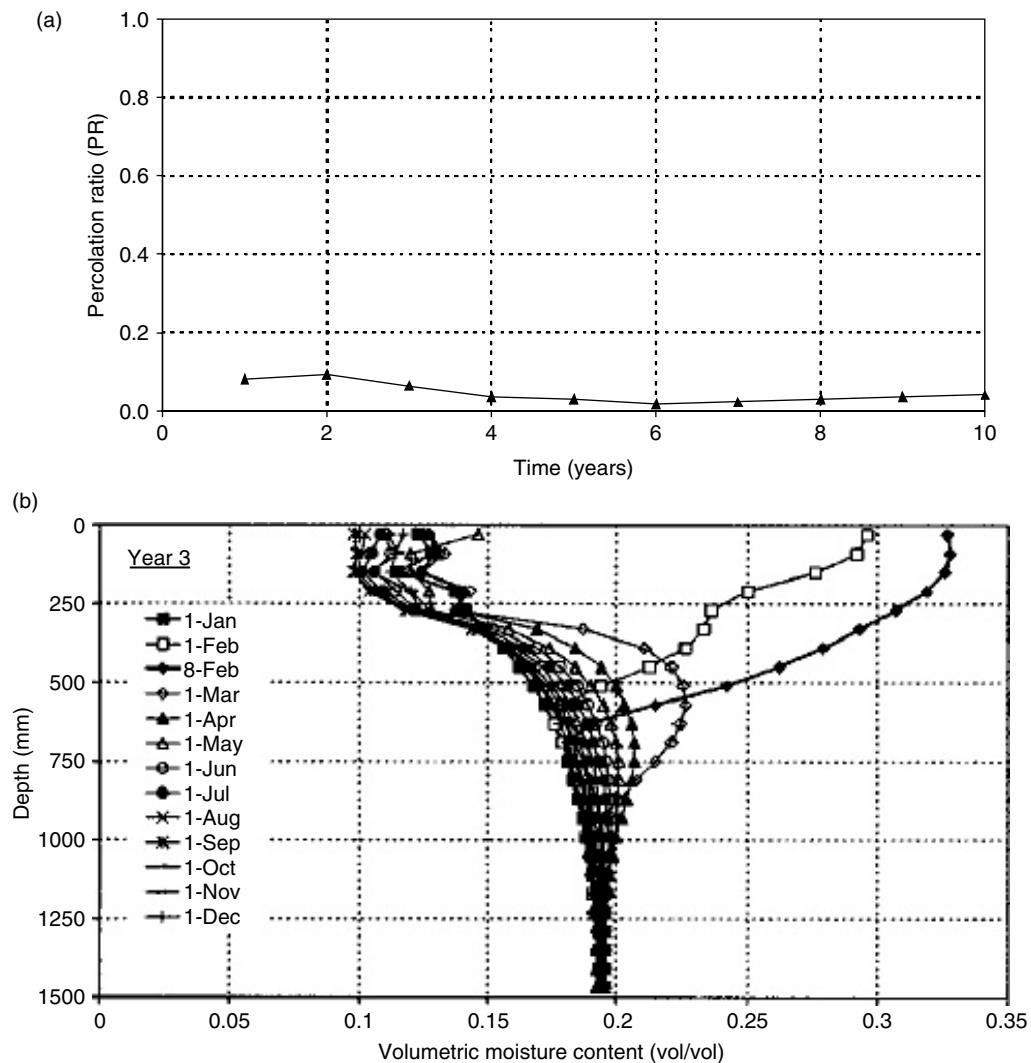


FIGURE 34.13 Performance variables for the O11 landfill evapotranspirative cover estimated using LEACHM: (a) comparative percolation ratio and (b) volumetric moisture content profiles during wettest simulation year (Zornberg et al. 2003).

of 1.0) for each year of the soil-specific, 10-year simulation. The engineered evapotranspirative cover constructed using the local soils, and placed under conditions defined by the T1 series, was then deemed to satisfy compliance with the prescriptive cover according to this demonstration. Figure 34.13b shows a prediction of the typical moisture content profiles expected for the cover. This figure indicates that the moisture fluctuations take place within the top meter of the soil cover, which reflects proper monolithic cover performance as discussed in Section 34.3.1.

34.6.3 Rocky Mountain Arsenal

Almost 200 acres of RCRA-Equivalent evapotranspirative covers have been recently designed to contain contaminated materials at the Rocky Mountain Arsenal (RMA), located near Denver, Colorado, USA. The climate in Denver is semiarid, with an average annual precipitation of 396 mm and an average pan evaporation of 1394 mm (as quantified for the 1948 to 1998 period). The wettest months of the year

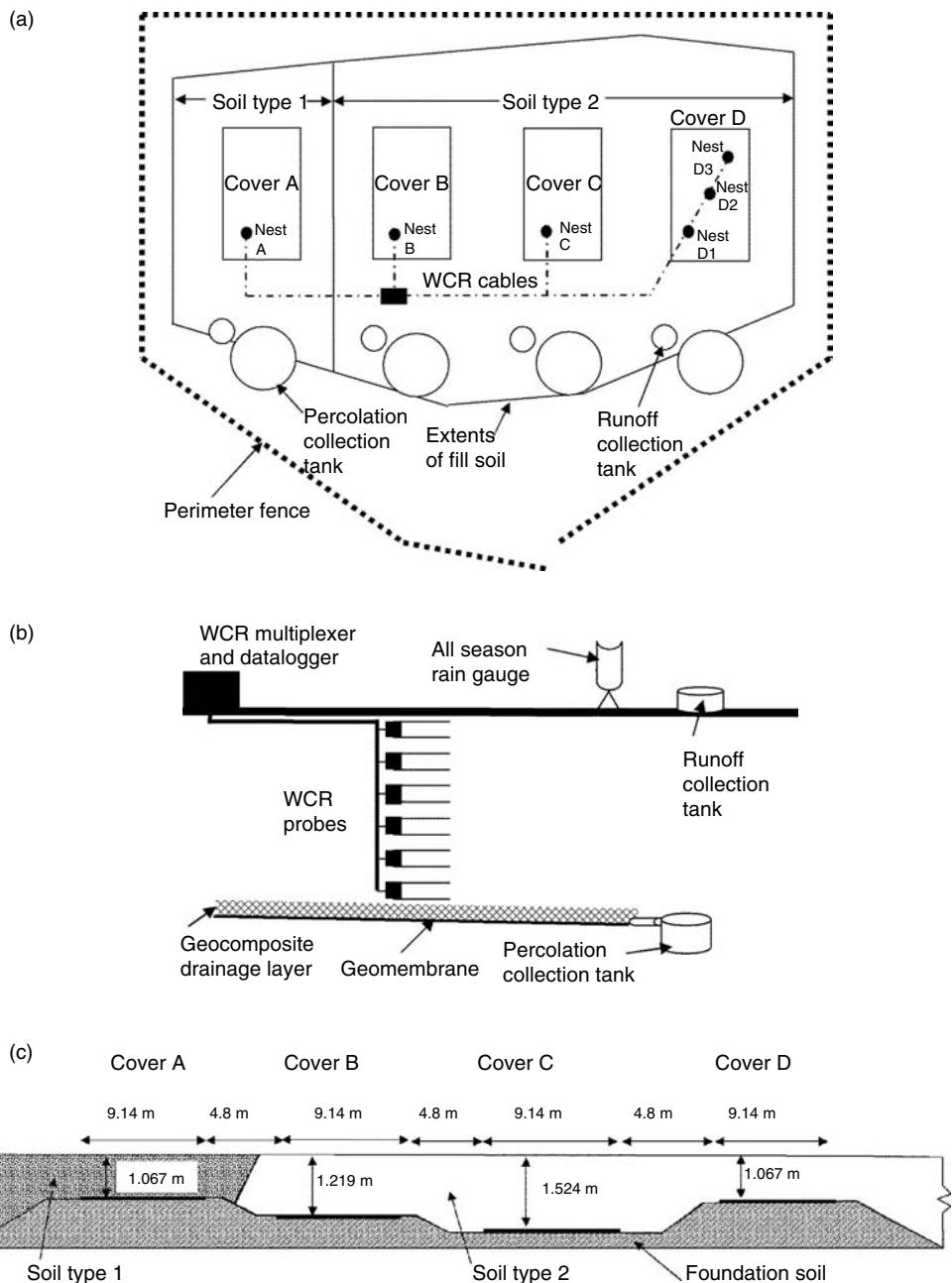


FIGURE 34.14 Layout for the monitoring program at the Rocky Mountain Arsenal: (a) plan view; (b) schematic view of instrumentation; and (c) elevation view.

(April to October) are also the months with the highest pan evaporation, which is appropriate for an evapotranspirative cover.

The Record of Decision (ROD) for the site requires a compliance demonstration to show equivalence of the alternative design with a prescriptive cover before construction of the final evapotranspirative covers. The design and compliance of the covers at the RMA site is governed by a quantitative percolation criterion. A MQPV threshold of 1.3 mm/year was selected at this site, which was based on eight years of leachate data collected from two landfill covers built to RCRA Subtitle C standards in Hamburg, Germany, according

TABLE 34.3 Hydraulic and Geotechnical Properties for Rocky Mountain Arsenal Cover Soils

Cover	Soil cover thickness (mm)	USCS classification (ASTM D2487)	Average % passing #200 sieve (% fines)		Atterberg limits		Saturated hydraulic conductivity (m/s)	van Genuchten model parameters		Relative compaction ^a (%)
			Average PL (%)	Average LL (%)	Average PL (%)	Average LL (%)		α (kPa^{-1})	N	
A	1143	SC	43.4	9	24.4	24.4	1.60E-05	0.6118	1.39	0.03
B	1270	CL	60.2	12.8	27.6	27.6	4.70E-06	0.3386	1.335	0.025
C	1676	CL	59.2	11.7	26.7	26.7	4.70E-06	0.3386	1.335	0.025
D	1168	CL	61.5	12	26.8	26.8	4.70E-06	0.3386	1.335	0.025

^a In relation to standard Proctor test (ASTM 698).

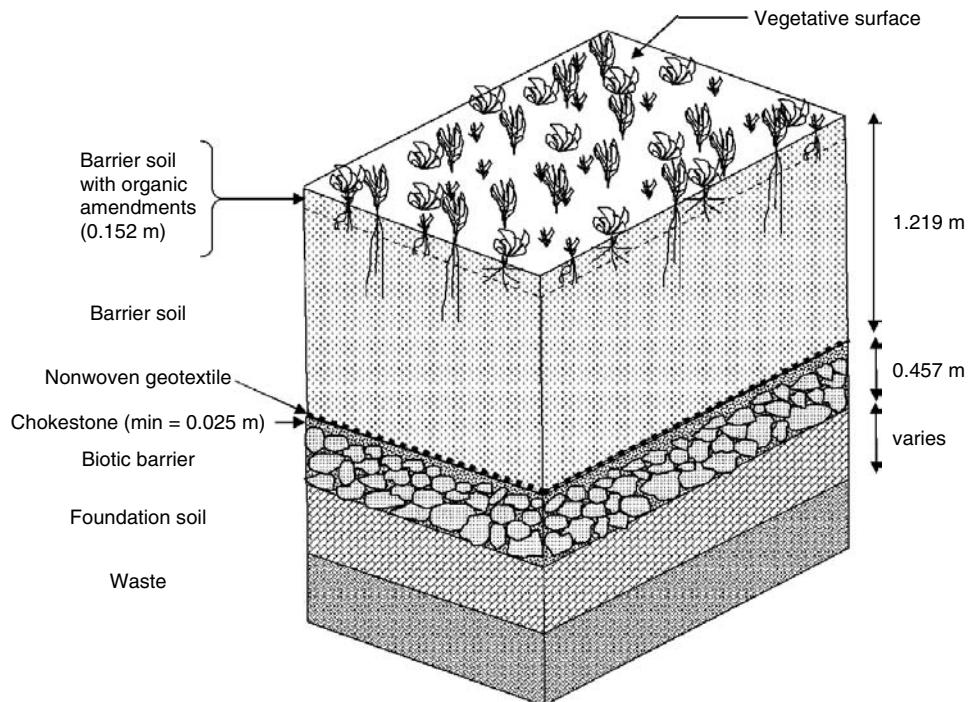


FIGURE 34.15 Final cover design at the Rocky Mountain Arsenal

to analyses described by Melchior (1997). This type of criterion was selected for its simplicity, as it sets a benchmark to be used in post-closure monitoring to demonstrate compliance, and is representative of the basal percolation for resistive covers.

The compliance demonstration at the Rocky Mountain Arsenal involved a field demonstration, which was complemented by comparative numerical analyses and a field demonstration (Kiel et al., 2002). Four evapotranspirative test covers were constructed on a rolling plain at the site in the summer of 1998. A plan view of the four test covers, referred to as covers A, B, C, and D, is shown in Figure 34.14a. The test covers were constructed using site-specific clays of low plasticity (CL), compacted atop large pan lysimeters (9.1 m by 15.2 m) placed on a 3% grade to allow gravity drainage to a collection tank. The geotechnical and hydraulic properties of the cover soils are summarized in Table 34.3. Figure 34.14b shows a schematic view of the monitoring layout used in the test covers. The instrumentation program involved monitoring of the basal percolation, precipitation, soil volumetric moisture content profile, and overland runoff in the four test covers. Basal percolation was collected in a gravity lysimeter, which consists of a geocomposite underlain by geomembrane. The lysimeters were constructed without sidewalls. Rain and snow were monitored using an all-season rain gauge. Surface water was collected in polyethylene geomembrane swales constructed around the cover perimeters. WCR probes were used to measure volumetric moisture content. Specifically, the covers were instrumented with nests of eight WCR probes. This included six WCR probes placed in a vertical nests of WCR probes and spaced evenly with depth. In addition, redundant WCR probes were placed at the same depth as the top and the bottom probes, approximately 1 foot aside from the vertical profile of WCR probes. Cover D was instrumented using three vertical profiles. Figure 34.14c shows an elevation view of the covers, indicating the depth of each cover. The covers are separated from each other by 2.4 m-wide buffer zones, and the entire area is vegetated with local grasses and shrubs.

The four test covers at the RMA were constructed to verify that the moisture flux through site-specific soils under local weather conditions remains below the MQPV of 1.3 mm/year (Kiel et al., 2002). The test plots at the Rocky Mountain Arsenal satisfied the quantitative percolation criterion over the period 1998 to 2003 of operation. That is, all four of the covers showed a yearly basal percolation rate below the MQPV despite having complemented the natural precipitation with irrigation. Although Cover D showed surface

depression, possibly due to the installation of moisture probes or burrowing animals, the collected basal percolation over this cover was still below the MPQV percolation threshold.

The monitoring results of the field demonstration program were used to develop the final cover design. As shown in Figure 34.15, the final cover design is a capillary barrier with a 1.067 m barrier layer, similar to that used in covers A and D. An additional 0.152 m of topsoil is used for vegetation. The specifications for the final cover soil are based on those for soil type I used in cover A. The final design includes a nonwoven geotextile over a chokestone layer (coarse gravel) to form a capillary break at the bottom interface of the barrier soil. The geotextile also helps prevent barrier soil particles from migrating into the chokestone layer. The chokestone is underlain by a biotic barrier consisting of crushed concrete from a demolition site. The biotic barrier is designed to prevent plants and burrowing animals from reaching the waste. The final cover will be instrumented with gravity lysimeters placed within the final cover to measure basal percolation and WCR probes to measure moisture profiles within the barrier soil.

Glossary

Anisotropic barrier Evapotranspirative cover with a system of barrier and drainage layers placed on a slope so that water is removed from the cover by lateral drainage.

Basal percolation Water that passes through a cover system into the underlying waste.

Biota barrier Layer of gravel or crushed concrete beneath the cover used to prevent entry of plants or animals into the waste.

Capillary break cover Evapotranspirative cover that exploits the increased moisture storage arising from placing a fine-grained soil over a coarse-grained soil.

Degree of saturation Ratio of the volumetric moisture content and the porosity.

Evapotranspirative cover systems Class of alternative landfill cover systems that functions by storing water from precipitation until it may be removed by evapotranspiration.

Field capacity Moisture content at which water will drain from a soil by gravity.

Geocomposite Drainage material consisting of a geonet with high transmissivity sandwiched between two nonwoven geotextile filters.

Geomembrane Low permeability geosynthetic used in hydraulic barriers.

HDU Heat dissipation unit, a device used to measure soil suction.

Lysimeter Device used to measure the basal percolation through a saturated soil layer.

MCPR Maximum comparative percolation ratio, a percolation criterion relying on comparison of the performance of an evapotranspirative cover and a prescriptive cover.

Monolithic cover Evapotranspirative cover with a single layer of soil acting as a hydraulic barrier and vegetation substrate.

MPQV Maximum percolation quantitative value, a percolation criterion relying on comparison of the performance of an evapotranspirative cover with a selected basal percolation value.

Prescriptive cover Landfill cover prescribed by government regulations that limits percolation by maximizing overland runoff.

Richards' equation Governing equation for water flow through unsaturated porous media.

Suction Difference between the pore air pressure and pore water pressure.

TDR Time domain reflectometry, a device used to infer the soil volumetric moisture content.

Volumetric moisture content Volume of water in a soil divided by the total volume of the soil.

Porosity Ratio of the volume of voids and the total volume.

WCR Water content reflectometer, a device used to measure volumetric moisture content in the field based on TDR technology.

References

- Albright, W.H. and Benson, C.H. (2002). *Alternative Cover Assessment Program 2002 Annual Report*. US EPA Publication No. 41182. 48 p.

- Albright, W.H., Gee, G.W., Wilson, G.V., and Fayer, M. (2002). *Alternative Cover Assessment Program Phase 1 Report*. US EPA Publication 41183. 203 p.
- Albright, W.H., Benson, C.H., Gee, G.W., Roesler, A.C., Abichou, T., Apiwantragoon, P., Lyles, B.F., and Rock, S.A. (2004). "Field water balance of landfill final covers." *Journal of Environmental Quality*. 33, 2317–2332.
- Anderson, J.E., Nowak, R.S., Ratzlaff, T.D., and Markham, O.D. (1993). "Managing soil moisture on waste burial sites in arid regions." *Journal of Environmental Quality*. 22, 62–69.
- Benson, C. and Gribb, M. (1997). "Measuring unsaturated hydraulic conductivity in the laboratory and field." In *Unsaturated Soil Engineering Practice*. Houston, S. and Wray, W. (eds). ASCE. Reston, VA. 113–168.
- Benson, C., Abichou, T., Wang, X., Gee, G., and Albright, W. (1999). *Test section installation instructions — Alternative Cover Assessment Program*. Environmental Geotechnics Report 99-3, Department of Civil & Environmental Engineering, University of Wisconsin-Madison.
- Benson, C., Abichou, T., Albright, W., Gee, G., and Roesler, A. (2001). "Field evaluation of alternative earthen final covers." *International Journal of Phytoremediation*. 3, 1–21.
- Brooks, R.H. and Corey, A.T. (1964). "Hydraulic properties of porous medium." Colorado State University (Fort Collins). Hydrology Paper No. 3. March.
- Burdine, N.T. (1953). "Relative permeability calculations from pore-size distribution data." *Petroleum Transactions of the American Institute of Mining and Metallurgical Engineering*. 198, 71–77.
- Bussiere, B., Aubertin, M., and Chapuis, R. (2003). "The behavior of inclined covers used as oxygen barriers." *Canadian Geotechnical Journal*. 40, 512–535.
- Cabral, A.R., Planchet, L., Marinho, F.A., and Lefebvre, G. (2004). "Determination of the soil water characteristic curve of compressible materials: Case study of de-inking residues." *ASTM Geotechnical Testing Journal*. 27(2): 154–162.
- Campbell, G. (1974). "A simple method for determining unsaturated conductivity from moisture retention data." *Soil Science*. 117(6): 311–314.
- Chandler, D.G., Seyfried, M., Murdock, M., and McNamara, J.P. (2004). "Field calibration of water content reflectometer." *Soil Science Society American Journal*. 68: 1501–1507.
- Conca, J. and Wright, J. (1992). "Diffusion and flow in gravel, soil, and whole rock." *Applied Hydrogeology*. 1, 5–24.
- Dell'Avanzi, E., Zornberg, J.G., and Cabral, A.R. (2004). "Suction profiles and scale factors for unsaturated flow under increased gravitational field." *Soils and Foundations*. 44, 1–11.
- Dwyer, S.F. (1998). "Alternative landfill covers pass the test." *Civil Engineering*. ASCE. 68(9): 50–52.
- Dwyer, S.F. (2003). *Water balance and computer simulations of landfill covers*. Doctoral thesis. The University of New Mexico. Albuquerque.
- Fayer, M. and Jones, T.L. (1990). "UNSAT-H version 2.0: Unsaturated soil water and heat flow model." PNL-679, Pacific Northwest Laboratory, Richland, Washington.
- Flint, A.L., Campbell, G.S., Ellet, K.M., and Calissendorff, C. (2002). "Calibration and temperature correction of heat dissipation matric potential densors." *Soil Science Society of America Journal*. 66, 1439–1445.
- Forbes, P.L. (1994). "Simple and accurate methods for converting centrifuge data into drainage and imbibition capillary pressure curves." *The Log Analyst*. 35, 31–53.
- Fredlund, D.G. and Xing, A. (1994). "Equations for the soil-water characteristic curve." *Canadian Geotechnical Journal*. 31, 521–532.
- Gardner, R.A. (1937). "The method of measuring the capillary pressures in small core samples." *Soil Science*. 43, 277–283.
- GeoSlope International Ltd. (2004). *Vadose/W Technical Overview*. Calgary, Canada.
- Hassler, G.L. and Brunner, E. (1945). "Measurements of capillary pressure in small core samples." *Transactions of AIME*. 160, 114–123.
- Hillel, D. (1998). *Environmental Soil Physics*. Academic Press. San Diego, CA.

- Hutson, J.L. and Wagenet, R.J. (1992). "Leaching Estimation and Chemistry Model, LEACHM." New York State College of Agriculture and Life Sciences, Cornell University.
- Kampf, M. and Holfelder, T. (1999). "Designing capillary barriers." *Seventh International Waste Management and Landfill Symposium*. CISA. Environmental Sanitary Engineering Centre, Sardinia. pp. 381–388.
- Khire, M.V., Meerdink, J.S., Benson, C.H., and Bosscher, P.J. (1995). "Unsaturated hydraulic conductivity and water balance predictions for earthen landfill final covers." *Soil Suction Applications in Geotechnical Engineering Practice*, Geotechnical Special Publication No. 48. ASCE, 38–57.
- Khire, M.V., Benson, C.H., and Bosscher, P.J. (1999). "Field data from a capillary barrier and model predictions with UNSAT-H." *Journal of Geotechnical and Geoenvironmental Engineering*, ASCE, 125, 518–527.
- Khire, M., Benson, C., and Bosscher, P. (2000). "Capillary Barriers in Semi-Arid and Arid Climates: Design Variables and the Water Balance." *Journal of Geotechnical and Geoenvironmental Engineering*, ASCE, Vol. 126, No. 8, pp. 695–708.
- Kiel, R.E., Chadwick, D.G., Lowrey, J., Mackey, C.V., and Greer, L.M. (2002). "Design of evapotranspirative (ET) covers at the Rocky Mountain Arsenal." *Proceedings: SWANA 6th Annual Landfill Symposium*.
- Kim, K.C. and Benson, C.H. (2002), *Moisture Content Calibrations for Final Cover Soils*, University of Wisconsin-Madison Geotech. Eng. Report 02-12, Madison, WI, 122 p.
- Klute, A. (1986). Water Retention: Laboratory Methods. *Methods of Soil Analysis, Part 1: Physical and Mineralogical Methods* SSSA. Madison, WI. 635–662.
- Kool, J.B. and Parker, J.C. (1987). "Development and evaluation of closed-form expression for hysteretic soil hydraulic properties." *Water Resources Research*, 23, 105–114.
- Lu, N. and Likos, W.J. (2005). *Unsaturated Soil Mechanics*. Wiley. New York. 584 p.
- McCartney, J.S. and Zornberg, J.G. (2002). "Design and performance criteria for evapotranspirative cover systems." *Fourth International Congress on Environmental Geotechnics*. Rio de Janeiro, Brazil.
- McCartney, J.S. and Zornberg, J.G. (2005a). "The Centrifuge Permeameter for Unsaturated Soils (CPUS)." *Proceedings of the International Symposium on Advanced Experimental Unsaturated Soil Mechanics, Experus 2005*. Trento, Italy, June 27–29, A.A. Balkema, pp. 299–304.
- McCartney, J.S., Kuhn, J.A., and Zornberg, J.G. (2005b). "Geosynthetic drainage layers in contact with unsaturated soils." *16th ISSMGE Conference: Geotechnical Engineering in Harmony with the Global Environment*. 12–16 September 2005. Osaka, Japan.
- Meerdink, J.S., Benson, C.H., and Khire, M.V. (1996). "Unsaturated hydraulic conductivity of two compacted barrier soils." *Journal of Geotechnical and Geoenvironmental Engineering*, ASCE, 122, 565–576.
- Melchior, S. (1997). "In-situ studies of the performance of landfill caps (Compacted clay liners, geomembranes, geosynthetic clay liner, capillary barriers)." *Land Contamination and Reclamation*, 5, 209–216.
- Meyer, P.D. and Gee, G.W. (1999). "Flux-based estimation of field capacity." *Journal of Geotechnical and Geoenvironmental Engineering*, 125, 595–599.
- Monteith, J.L. (1965) "Evaporation and environment." *Symposia of the Society for Experimental Biology*.
- Morris, C.E. and Stormont, J.C. (1997). "Capillary barriers and Subtitle D covers: Estimating equivalency." *Journal of Environmental Engineering*, ASCE, 123, 3–10.
- Nachabe, M.H. (1998). "Refining the definition of field capacity in the literature." *Journal of Irrigation and Drainage Engineering*, 124, 230–232.
- Nichol, C., Smith, L., and Beckie, R. (2003). "Long-term measurement of matric suction using thermal conductivity sensors." *Can. Geotech. J.* 40: 587–597.
- Nimmo, J.R., Rubin, J., and Hammermeister, D.P. (1987). "Unsaturated flow in a centrifugal field: Measurement of hydraulic conductivity and testing of Darcy's law." *Water Resources Research*, 23, 124–134.
- O'Kane, M., Wilson, G.W., and Barbour, S.L. (1998). "Instrumentation and monitoring of an engineered soil cover system for mine waste rock." *Canadian Journal of Geotechnical Engineering*, 35, 828–846.

- Olson, R.E. and Langfelder, L.J. (1965). "Pore water pressures in unsaturated soils." *Journal of the Soil Mechanical and Foundation Division ASCE*, 91, 127–151.
- Parent, S.E. and Cabral, A.R. (2004). Procedure for the design of inclined covers with capillary barrier effect. *Proc. 57th Canadian Geotechnical Conference*, Québec, October, 24–28.
- Ridley, A.M. and Burland, J.B. (1995). "A pore pressure probe for the in-situ measurement of soil suction". *Proceedings of Conference on Advances in Site Investigation Practice*. I.C.E.:London.
- Scanlon, B.R. and Milly, P.C.D. (1994). "Water and heat fluxes in desert soils 2. Numerical simulations." *Water Resources Research*. 30, 721–733.
- Schroeder, P.R., Dozier, T.S., Zappi, P.A., McEnroe, B.M., Sjostrom, J.W., and Peyton, R.L. (1994). "The Hydrologic Evaluation of Landfill Performance (HELP) Model: Engineering documentation for Version 3, EPA/600/R-94/168b." US EPA Risk Reduction Engineering Laboratory, Cincinnati, OH.
- Shackelford, C.D., Chang, C.-K., and Chiu, T.-F. (1994). "The capillary barrier effect in unsaturated flow through soil barriers." *1st ICEG Conference*. Edmonton, CA, pp. 789–793.
- Siddiqui, S.I., Drnevich, V.I., and Deschamps, R.J. (2000). "Time domain reflectometry development for use in geotechnical engineering," *Geotechnical Testing Journal*. 23, 9–20.
- Simunek, J., Sejna, M., and van Genuchten, M. 1998. *HYDRUS-1D: Code for Simulating the One-Dimensional of Water, Heat, and Multiple Solutes in Variably Saturated Porous Media*. Version 2.02. International Groundwater Modeling Center. Colorado School of Mines. Golden, CO.
- Sisson, J.B., Gee, G., Hubbell, J.M., Bratton, W.L., Ritter, J.C., Ward, A.L., and Caldwell, A.C. (2002). "Advances in tensiometry for long-term monitoring of soil water pressures." *Vadose Zone Hydrology*. 1, 310–315.
- Stormont, J. (1995). "The effect of constant anisotropy on capillary barrier performance." *Water Resources Res.* 32, 783–785.
- Stormont, J.C. and Anderson, C.E. (1999). "Capillary barrier effect from underlying coarser soil layer." *Journal of Geotechnical and Geoenvironmental Engineering*. 125, 641–648.
- Stormont, J.C. and Morris, C.E. (2000). "Characterization of unsaturated nonwoven geotextiles." In *Advances in Unsaturated Geotechnics: Proceedings of Sessions of Geo-Denver 2000*. Chang, N.-Y., Houston, S.L., and Shackelford, C.D. (eds) August 5–8, 2000. Denver, Colorado. pp. 153–164.
- Suwansuwat, S. and Benson C.H. (1999). "Cell size for water content-dielectric constant calibrations for time domain reflectometry." *Geotechnical Testing Journal*. 22, 3–12.
- Tarantino, A. and Mongiovì, L. (2003). "Calibration of tensiometer for direct measurement of matric suction." *Géotechnique*. 53, 137–141.
- Topp, G., Davis, J., and Annan, A. (1980). "Electromagnetic determination of soil water content: Measurement in coaxial transmission lines." *Water Resources Research*. 16, 574–582.
- Topp, G.C. and Miller, E.E. (1966). "Hysteretic moisture characteristics and hydraulic conductivities for glass-bead media." *Soil Science Society of American Proceedings*. 30, 156–162.
- van Genuchten, M. (1980). "A closed-form equation for predicting the hydraulic conductivity of unsaturated soils." *SSSA*. 44, 892–898.
- Wang, X. and Benson, C.H. (2004). "Leak-free pressure plate extractor for the soil water characteristic curve." *Geotechnical Testing Journal*. 27, 1–10.
- Watson, K.K. (1996). "An instantaneous profile method for determining the hydraulic conductivity of unsaturated porous materials." *Water Resour. Res.*, 2, 709–715.
- Waugh, W.J., Petersen, K.L., Link, S.O., Bjornstad, B.N., and Gee, G.W. (1994). "Natural analogs of the long term performance of engineered covers." In *situ Remediation: Scientific Basis for Current and Future Technologies*, Gee, G.W. and Wing, N.R. (eds), Battelle Press, Richland Washington. pp. 379–409.
- Williams, J.R., Jones, C.A., and Dyke, P.T. (1984). "The EPIC model and its application." *Proceeding of ICRISAT-IBSNAT-SYSS Symposium on Minimum Data Sets for Agrotechnology Transfer*. March 1993.

- Zornberg, J.G. and McCartney, J.S. (2003). *Analysis of monitoring data from the evapotranspirative test covers at the Rocky Mountain Arsenal*. Geotechnical Research Report, US Environmental Protection Agency, Region 8, December 2003, 227 p.
- Zornberg, J.G., Jernigan, B.L., Sanglerat, T.R., and Cooley, B.H. (1999). "Retention of free liquids in landfills undergoing vertical expansion." *Journal of Geotechnical and Geoenvironmental Engineering*. ASCE, 125, 583–594.
- Zornberg, J.G., LaFountain, L., and Caldwell, J.C. (2003). "Analysis and design of evapotranspirative cover for hazardous waste landfill." *Journal of Geotechnical and Geoenvironmental Engineering*. ASCE, 129, 427–438.

Further Information

Further information on the hydraulic properties of unsaturated soils may be found in Lu and Likos (2005). Descriptions of common techniques used to determine the SWRC can be found in Wang and Benson (2004) and the proceedings of *Experus* (2005). Descriptions of common techniques used to determine the *K*-function can be found in Benson and Gribb (1997). The ACAP Phase 1 report provides a comparative analysis of available numerical models for evapotranspirative cover performance evaluation, while the 2002 ACAP Annual Report provides a summary of field monitoring programs for evapotranspirative covers. An overview of current research topics in monitoring and analysis of evapotranspirative covers can be found in ASCE Geotechnical Special Publication 142 "Waste Containment and Remediation," held at GeoFrontiers 2005 in Austin, TX.

35

Groundwater Monitoring

35.1	Fundamentals	35-1
	Introduction • Purpose of Groundwater Monitoring •	
	Regulatory Requirements for Groundwater Monitoring •	
	Groundwater Monitoring Strategy Considerations	
35.2	Groundwater Monitoring Plans	35-7
	Introduction • Developing a Groundwater Monitoring Plan	
35.3	Design and Installation of Groundwater Monitoring Devices.....	35-13
	Overview of Groundwater Monitoring Devices •	
	Groundwater Monitoring Wells • Piezometers •	
	Groundwater Discharges • Alternative Groundwater Monitoring Devices	
35.4	Groundwater Sampling.....	35-24
	General Approach to Sampling Groundwater Monitoring Wells • Overview of Groundwater Sampling Techniques •	
	<i>In Situ</i> or Alternative Monitoring Techniques • Quality Assurance and Quality Control Samples • Sample Handling and Preservation • Documentation of Sampling Events	
35.5	Analysis of Groundwater Samples	35-34
	Field Chemical Analyses • Laboratory Chemical Analyses	
35.6	Evaluation of Groundwater Monitoring Data	35-37
	Data Validation • Evaluations of Groundwater Elevation Measurements or Depths to Groundwater • Groundwater Quality Data Analysis • What to Do if a Statistical Analysis Indicates the Presence of Contaminants	
	Glossary.....	35-43
	References	35-45
	Further Information	35-47

Michael F. Houlahan and
Paul J. Botek
GeoSyntec Consultants

35.1 Fundamentals

35.1.1 Introduction

In this chapter on groundwater monitoring, a summary of approaches and techniques for sampling, analyzing, and evaluating the quality of groundwater is provided. In the remaining portion of this subsection on fundamentals, the purpose of groundwater monitoring is addressed and an overview of the regulatory

requirements for groundwater monitoring is provided; also, strategies for developing groundwater monitoring programs are provided. In other sections of this chapter, descriptions are provided for the design and installation of groundwater monitoring systems, sampling and analysis of groundwater, and evaluation of groundwater monitoring data. The primary focus of this chapter is on groundwater monitoring at sites where contaminant releases, or potential releases, have impacted, or could impact, groundwater quality.

35.1.2 Purpose of Groundwater Monitoring

Groundwater is used for a variety of purposes, including irrigation, drinking, and manufacturing. Groundwater is also the source of a large percentage of surface water. To verify that groundwater is suited for its purpose, its quality can be evaluated (i.e., monitored) by collecting samples and analyzing them. In simplest terms, the purpose of groundwater monitoring is to define the physical, chemical, and biological characteristics of groundwater. If the characteristics of the groundwater do not meet the requirements of its intended use, or if the groundwater could be harmful to human health or the environment, then it may need to be remediated. Groundwater remediation is discussed in Chapter 36 of this handbook.

The purpose of a groundwater monitoring program should be defined before monitoring begins so that appropriate procedures, techniques, and analyses can be planned that will meet the specific needs of the project. *Detection* monitoring programs are used to detect the impact to groundwater quality. *Assessment* monitoring programs are used: (1) to assess the nature and extent of contaminants that have been detected in groundwater; and (2) to collect data that may be needed to perform a design for remediation of the groundwater. Corrective action monitoring programs are used to assess the impact of a groundwater remedy on contaminant concentrations as a tool in evaluating the success of a remedy. Performance monitoring programs are used to evaluate the effectiveness of an element of a groundwater remediation system in meeting the design criteria for the element. These four types of monitoring programs, which are commonly required at facilities regulated under the Resource Conservation and Recovery Act (RCRA), are the focus of this chapter. More detailed descriptions of the requirements of these programs are provided by the United States Environmental Protection Agency (USEPA) (1986a, 1992a, 1993c) and may also be contained in other applicable state, local, or federal guidance documents. Although the monitoring programs described in this chapter are based on the requirements of RCRA, the elements of these monitoring programs may also be useful and applicable to non-RCRA groundwater monitoring programs, such as groundwater monitoring programs at Superfund sites, at sites that are being voluntarily remediated, and at sites where a legally binding agreement has been made (e.g., a “consent agreement”) requiring monitoring of the site. A flow diagram showing the typical sequence of groundwater monitoring activities is presented in Figure 35.1. A description of the characteristics of the groundwater monitoring program types identified above is provided in the following subsections.

35.1.2.1 Detection Monitoring Programs

For sites where regulated materials or contaminants of concern are used, stored, or contained a detection monitoring program is often implemented to evaluate whether or not a release has occurred. A detection monitoring program typically involves obtaining samples from wells located proximal to the location where regulated substances are present or in a downgradient area where these constituents would be expected to migrate if released. The samples are then analyzed for a constituent or series of constituents that are indicative of a release (i.e., indicator parameters). A sampling frequency is also typically established to ensure that data are collected often enough to identify potential releases before the contaminants migrate beyond the monitoring well network. The groundwater quality data from a detection monitoring program are typically compared to background concentrations to identify changes in groundwater quality that would confirm a release. Background concentrations are usually established by collecting four initial samples ($N = 4$) from a newly installed monitoring well. The sample interval for the $N = 4$ sampling event is normally within the first three months after well installation. These comparisons may include inter-well comparisons against data collected at wells outside of the presumed area of interest (e.g., upgradient wells) or through intrawell comparisons to evaluated temporal trends in the data that may indicate an

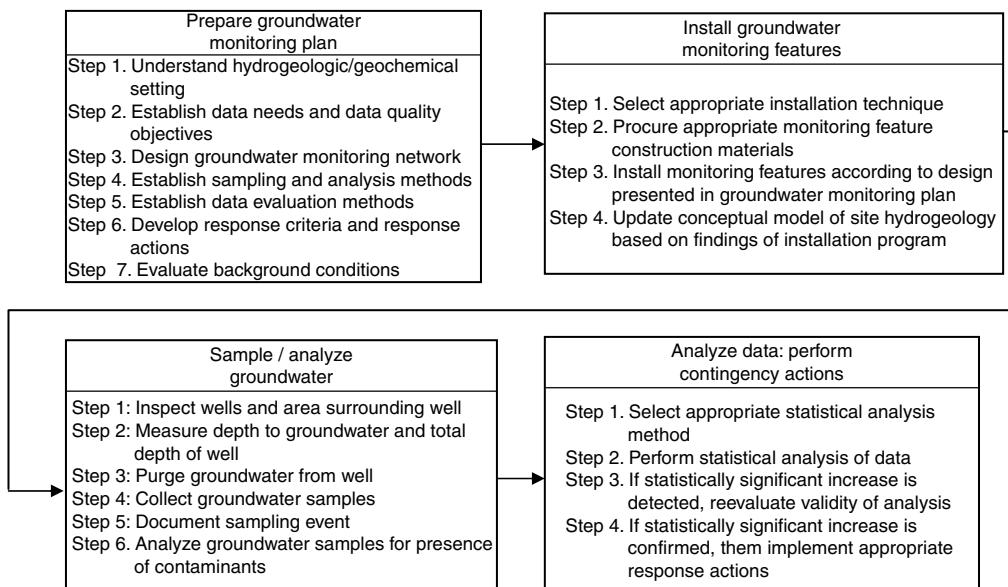


FIGURE 35.1 Sequence of activities for groundwater monitoring.

increase or decrease in a particular constituent. These comparisons typically require statistical analyses of the background and monitoring data to verify that a change in groundwater quality is verifiable. Statistical analysis of groundwater quality data is described in Section 35.6.3.2.

35.1.2.2 Assessment Monitoring Programs

For sites where a release to groundwater is known to have occurred, a monitoring program should be implemented to assess the nature and extent of the release. The purpose of assessment monitoring is to evaluate the nature and extent of the release and, if it is found that corrective action is needed, to identify potentially applicable corrective actions. An assessment monitoring plan establishes the thresholds for implementing a corrective action. Many times these thresholds would include establishing a groundwater protection standard, or a concentration that would trigger a corrective action. Typically, the groundwater protection standard is defined as the constituent's maximum contaminant level (MCL) (which is defined by the USEPA) or the background concentration, whichever is greater, or an alternate concentration limit (ACL). An alternative concentration limit is a concentration for a particular contaminant that is often times higher than background or MCL value but that provides continued protection of human health and the environment. An ACL must be established during the permitting process and included in the Facility permit. A detailed discussion of ACL development and permitting is provided in USEPA (1987). If the groundwater is discharging to surface water, concentrations protective of aquatic life may also need to be considered.

A point of compliance is also another common aspect of an assessment monitoring program that should be established prior to implementing an assessment monitoring program. The point of compliance is a point (or line) at which an exceedance of the groundwater protection would trigger a corrective action. The point of compliance is often established by the regulatory requirements and typically includes features such as the site property boundary or groundwater discharges where potential receptors may be exposed to groundwater contaminants. As the purpose of the assessment monitoring program is also to evaluate the nature and extent of contamination, the assessment monitoring field exploration should be focused primarily on providing data to fill gaps in the detection monitoring database or to provide data related to suspected contaminants. Also, if the results of assessment monitoring indicate that corrective action is required, a list of technically feasible corrective measures should be developed. It is useful to develop this

list during assessment monitoring so that data can be collected during the assessment monitoring field exploration and subsequently used to perform a preliminary assessment of the feasibility of the candidate remedial alternatives.

The assessment monitoring plan should contain criteria for either returning to detection monitoring or proceeding to corrective action. For example, corrective action may be warranted if the presence of contaminants is confirmed and the contaminants are present at concentrations that are greater than either background concentrations or the groundwater protection standard. A return to detection monitoring may be warranted for any of the following reasons: (1) through additional sampling and analysis, the detection monitoring results are found to have given a false positive indication of a release from the potential source; (2) none of the parameters on the expanded list of analytes is detected at a statistically significant level above background concentrations; or (3) the concentrations of contaminants do not exceed the groundwater protection standard.

35.1.2.3 Corrective Action Monitoring Programs

If the assessment monitoring data confirm that contaminants are present at concentrations that exceed the groundwater protection standard, then corrective action may be required. Corrective action of contaminated groundwater is described in Chapter 36 of this handbook. During corrective action, a groundwater monitoring program can be implemented to evaluate the effectiveness of the remedy. As an example, for a remedy involving extraction of groundwater to reduce contaminant concentrations in a plume of contaminated groundwater, corrective action monitoring could involve sampling and analysis of groundwater at wells (i.e., “compliance” wells) located within the plume of contamination to evaluate whether or not contaminant concentrations are decreasing at these locations. If it is found that contaminant concentrations in groundwater are not decreasing, then either the remediation system could be adjusted so that it will more likely achieve the remediation goal or it could be demonstrated that groundwater remediation at that site is technically impracticable (as described in the next chapter in Section 36.2.2). Finally, the conditions that trigger the end of corrective action (e.g., attainment of the groundwater protection standard at the compliance wells) or the need for alternative corrective action (e.g., continued exceedance of the groundwater protection standard or significant increase in the concentrations of detected contaminants) should be defined.

35.1.2.4 Performance Monitoring Programs

Monitoring the performance of a corrective action system is a useful approach for verifying that the features of the system are fulfilling their respective design criteria. As described in Chapter 36, groundwater remediation system features may include groundwater extraction wells, groundwater extraction trenches, or low-permeability subsurface barriers. A performance monitoring program typically involves monitoring to evaluate whether a feature of the remediation system is fulfilling its intended function. An example of performance monitoring for a groundwater extraction system is presented in Figure 35.2. If monitoring results indicate that the design criteria is not being satisfied by the feature, then the design should be reviewed and improvements should be identified that will more likely provide the desired result. Performance monitoring is described in more detail by USEPA (1994) and by Rumer and Mitchell (1995).

35.1.3 Regulatory Requirements for Groundwater Monitoring

Federal and state regulations have been established to provide for protection of groundwater resources. An extensive discussion of these regulations is provided in Chapter 32 of this handbook. As described in Chapter 32, these regulations require monitoring of groundwater quality near operations where materials are managed that could have an adverse impact on groundwater quality. The following questions are provided as a guide to identifying applicable groundwater monitoring regulations for a particular site or operation.

- What is the activity (e.g., underground storage of petroleum hydrocarbon products) that could have an adverse impact on groundwater quality? Are there related activities (e.g., transmission

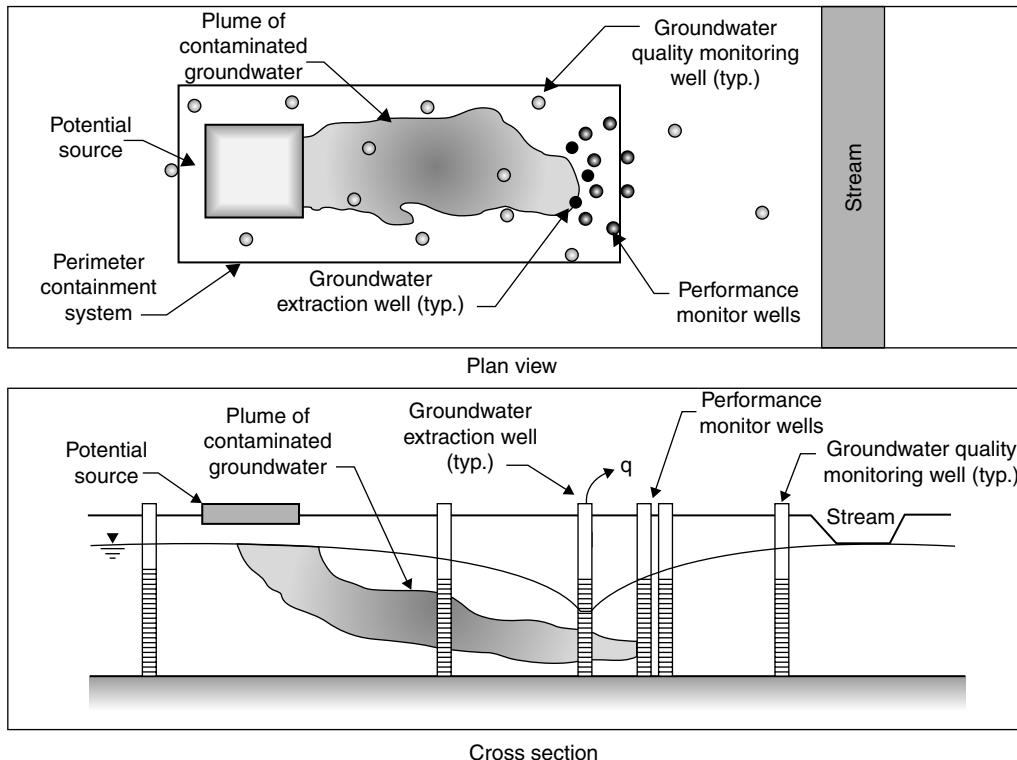


FIGURE 35.2 Performance monitoring of groundwater extraction system. (From USEPA, 1990. *Handbook — Ground Water, Volume I: Ground Water and Contamination*. EPA/625/6-90/016a. Washington, DC.)

of petroleum products to the storage tank through underground pipes) that could also have an adverse impact on groundwater quality?

- What regulations govern the activity or activities, and what portions of those regulations address either groundwater monitoring or potential impacts to groundwater quality?
- What specific requirements do the regulations contain for planning groundwater monitoring events, for sampling groundwater, for analyzing groundwater samples, and for evaluating data from groundwater sampling events?
- Do the regulations contain specific requirements for action in the event that impacts to groundwater quality are identified during the monitoring program?

Once applicable regulatory requirements have been identified, a program can be developed for addressing each applicable regulatory requirement. A useful tool for addressing the requirements is the groundwater monitoring plan. Guidelines for developing a groundwater monitoring plan are presented in Section 35.2.2.

35.1.4 Groundwater Monitoring Strategy Considerations

After identifying applicable regulatory requirements and before preparing a groundwater monitoring plan, the overall strategy of the groundwater monitoring program should be defined to guide the development of the plan. In this sense, “strategy” refers to the manner in which a hypothetical release from a regulated unit will be detected or measured. Examples of issues that should be addressed when developing a monitoring strategy include: (1) the type of monitoring data needed; (2) the locations (both horizontal and vertical) from which the samples are to be collected (i.e., definition of “target monitoring zones”); (3) the manner

in which the samples will be obtained; and (4) the ability of the monitoring features to rapidly detect a change in groundwater quality. For detection monitoring programs, the types of data needed are usually defined by regulation; for other types of monitoring programs, the types of data needed are typically based on site-specific considerations. Selection of target monitoring zones is described in Section 35.2; design of groundwater monitoring devices is described in Section 35.3; and the ability of a monitoring system to rapidly detect an impact on groundwater quality is discussed in Section 35.2.2 (see “Phase II” of Step 3 of a groundwater monitoring system design). See Chapter 2 of Nielsen (1991) for additional discussion of groundwater monitoring strategy.

Development of a groundwater monitoring strategy is illustrated in Figure 35.3 to Figure 35.5. As shown in these figures, the potential sources of contamination and the aquifers of concern should be characterized before developing a groundwater monitoring strategy because selection of target monitoring zones cannot be made until the source and the aquifer of concern have been evaluated, usually through a detailed hydrogeologic evaluation of the site. When evaluating the ability of a monitoring system to rapidly detect a release from the potential source, the impact of preferential flow paths and vertical gradients should be carefully evaluated; a two-dimensional analysis of groundwater elevation may not reveal actual upgradient

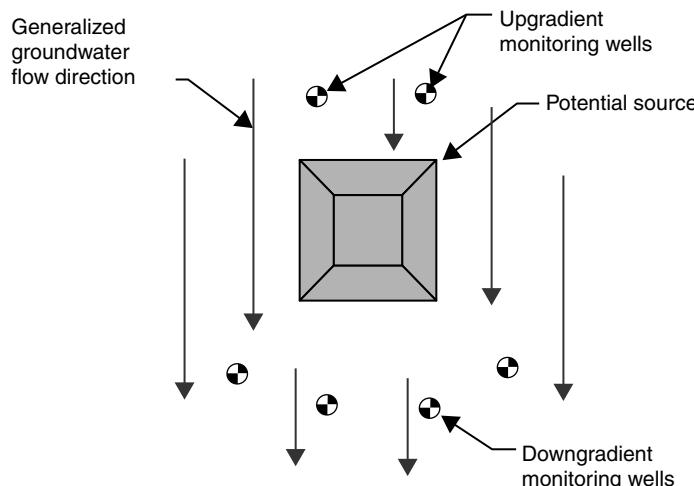


FIGURE 35.3 Plan view of typical unconfined aquifer groundwater monitoring system.

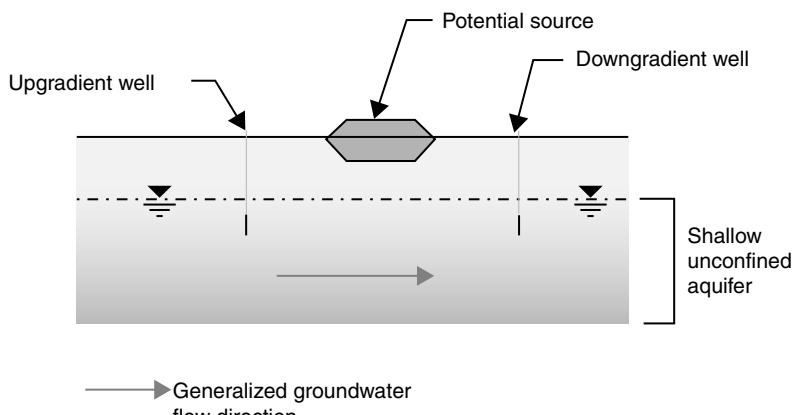


FIGURE 35.4 Section of typical unconfined aquifer groundwater monitoring system.

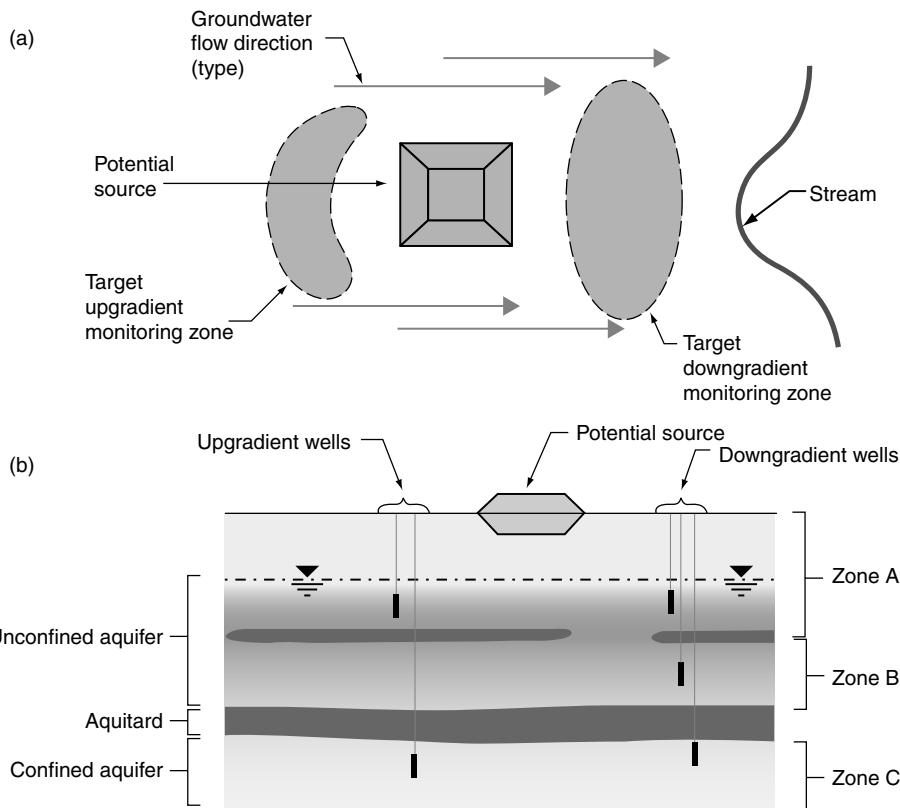


FIGURE 35.5 (a) Horizontal target monitoring zones; (b) vertical target monitoring zones.

or downgradient locations of groundwater flow. The presence of vertical gradients may significantly affect the selection of monitoring locations.

35.2 Groundwater Monitoring Plans

35.2.1 Introduction

Successful implementation of a groundwater monitoring strategy requires a well-planned approach and an understanding of the factors that could affect the quality, validity, or representativeness of groundwater samples. A groundwater monitoring plan is an effective tool for systematically addressing these factors.

Groundwater monitoring plans are required on a site-by-site basis for facilities that are regulated under RCRA and the Comprehensive Emergency Response and Compensation Liability Act (CERCLA), and may also be required for facilities that are regulated under various state and local programs. Groundwater monitoring plans are usually prepared to describe each aspect of groundwater monitoring and to control monitoring activities so that they fulfill the overall goals of the groundwater monitoring strategy. A comprehensive groundwater monitoring plan addresses each activity that will occur during sampling, analysis, and data interpretation, as well as response actions to be taken based on the results of monitoring.

35.2.2 Developing a Groundwater Monitoring Plan

Developing a groundwater monitoring plan requires an in-depth understanding of site conditions, contaminant properties, regulatory requirements, and other technical considerations. When developing a groundwater monitoring plan, a stepwise approach is recommended to minimize omissions and to

TABLE 35.1 Sample Outline — Groundwater Monitoring Plan

1.	Introduction
2.	Site Conditions
2.1	Site Topography and Surface-Water Hydrology
2.2	Site Geology
2.3	Site Hydrogeology
3.	Groundwater Monitoring System
3.1	Introduction
3.2	Existing Groundwater Monitoring Wells
3.3	Construction of New Groundwater Monitoring Wells
3.4	Development of New Groundwater Monitoring Wells
4.	Groundwater Sampling Procedures
4.1	Introduction
4.2	Health and Safety Procedures
4.3	Recordkeeping
4.4	Equipment Use, Cleaning, Decontamination
4.5	Preparation for Sampling
4.6	Groundwater Sample Collection Procedures
4.7	Groundwater Sample Handling and Preservation
4.8	Shipment of Groundwater Samples
5.	Laboratory Groundwater Analytical Program
6.	Quality Assurance and Quality Control Procedures
6.1	Data Quality Objectives
6.2	Field Sampling QA/QC Requirements
6.3	Laboratory QA/QC Requirements
7.	Evaluation of Groundwater Data
7.1	Introduction
7.2	Groundwater Quality Data Validation
7.3	Evaluation of Static Groundwater Elevations
7.4	Statistical Analysis Procedures for Groundwater Quality Data
7.5	Response Actions
7.6	Reporting Requirements for Groundwater Quality Data
8.	References

provide a clear and concise strategy for identifying the goals, requirements, and limitations of the groundwater monitoring program. In this section, a procedure is presented for developing a thorough groundwater monitoring plan. Additional guidance on addressing the elements described in this section is presented by USEPA (1992a) and Nielsen (1991). An example outline for a groundwater monitoring plan is presented in Table 35.1.

35.2.2.1 Step 1: Understand Hydrogeologic/Geochemical Setting

A comprehensive understanding of the hydrogeologic and geochemical setting of the site is an essential prerequisite to designing a groundwater monitoring system that can reliably detect a release to groundwater. This understanding should be based on an analysis of hydrogeologic and geochemical data from site explorations, literature reviews, and experience with similar hydrogeologic and geochemical settings. The types of hydrogeologic and geochemical information that should be collected and the general data needs for each type of information are described below, along with recommended means for obtaining the data.

Aquifer Hydrogeologic Parameters: The transmissivity, porosity, and storage properties (i.e., specific yield or specific capacity) of the aquifer should be evaluated. Transmissivity and storage properties are typically calculated based on the results of aquifer tests, as described in Chapter 10 of this handbook and by Freeze and Cherry (1979). Porosity is typically determined in the laboratory or estimated from the literature. Effective porosities can be estimated for bedrock aquifers by assuming that fractured rock acts as an equivalent porous medium.

Geologic Information: Characterization of the geologic units beneath the site is important to developing an understanding of the pathways that contaminants could take if introduced into groundwater (fate and transport of contaminants is described in detail in Chapter 18 and Chapter 19 of this handbook and in USEPA [1989c]). Examples of information that should be obtained before installing groundwater monitoring features include the horizontal and vertical limits of the aquifers, horizontal and vertical limits of *confining units* or *aquitards*, presence of interconnections between aquifers, *anisotropies* in aquifer material, presence of discontinuities (e.g., fractures, solution cavities, channel deposits, etc.) within or between stratigraphic units, gradual variations of the stratigraphic units with depth or with horizontal location, and the nature of the earth materials within the aquifers (e.g., particle size, angularity, dispersion-related properties, etc.).

Topographic Information: The surface topography of a site may provide important clues to the presence or extent of geologic/stratigraphic units at the site and, accordingly, may provide an indication of the potential impact of a release of contaminants to groundwater. An understanding of site topography may provide valuable insight into: (1) the manner in which contaminants may have reached groundwater after being released to the ground surface; (2) past waste management practices at the site; (3) the interconnection between surface water and groundwater; and (4) aspects of the depositional history of the subsurface soils, which may give clues to the subsurface lithology and the manner in which it might control groundwater flow and contaminant transport. A discussion of terrain evaluation approaches and their use in site investigations is provided by Lueder (1959).

Potentiometric Surface Information: Knowledge of the potentiometric surface elevation is essential to understanding the direction of groundwater flow and the fate of contaminants. Potentiometric surface information is also important for evaluating whether an aquifer is unconfined, confined, or perched. Potentiometric surface information should be obtained for the uppermost aquifer and for any underlying aquifer that could be interconnected with the uppermost aquifer, as well as unconnected aquifers that could be impacted by a release from the potential source. Subsurface monitoring data can sometimes be supplemented with surface-water elevation data (e.g., springs, streams, etc.) to provide a better understanding of groundwater flow patterns.

Geochemical Conditions: Contaminants can chemically react with geologic media. Several types of reactions are common in aquifers (e.g., sorption, chelation, complexation, ion exchange, precipitation/dissolution, and hydrolysis, which are described in Chapter 18 and Chapter 19 of this handbook). Characterizing the geochemical properties of the aquifer materials is important when evaluating the effect of these reactions on fate and transport of contaminants. Detailed discussions of groundwater geochemistry are provided in Chapter 17 of this handbook and by Chappelle (1993).

35.2.2.2 Step 2: Establish Data Needs and Data Quality Objectives

When developing a groundwater monitoring plan, a decision should be made regarding the types and quality of data needed to support decisions regarding the presence of groundwater contamination. To guide such decisions, data quality objectives (DQOs) can be established in the groundwater monitoring plan. DQOs can be specified as either qualitative (i.e., as a specific set of procedures to follow for collecting data) or quantitative (i.e., as the amount of imprecision or bias error that may be tolerated in data without incurring an unacceptable probability of making incorrect or inappropriate decisions). A more detailed discussion of DQOs is presented by USEPA (2000, 2002). Figure 35.6 depicts the DQO process.

An important step in this process is determining an appropriate sampling schedule or frequency. Many regulatory programs, such as RCRA, prescribe the minimum sampling frequencies and these requirements should be verified prior to implementing a groundwater monitoring program. If a prescribed sampling frequency is not provided, the groundwater professional must establish a sampling frequency that provides sufficient data to evaluate changes in groundwater quality over time at an interval that will not potentially miss contaminants as they migrate through the monitoring network. When developing the sampling schedule and frequency, one must consider temporal variations in the hydrogeologic conditions, such as seasonal variations in flow or timing of dewatering activities, if any, in nearby locations, and the fate and transport conditions for the contaminants of concern including groundwater, and the affect of sampling frequency on data evaluations such as statistical analyses.

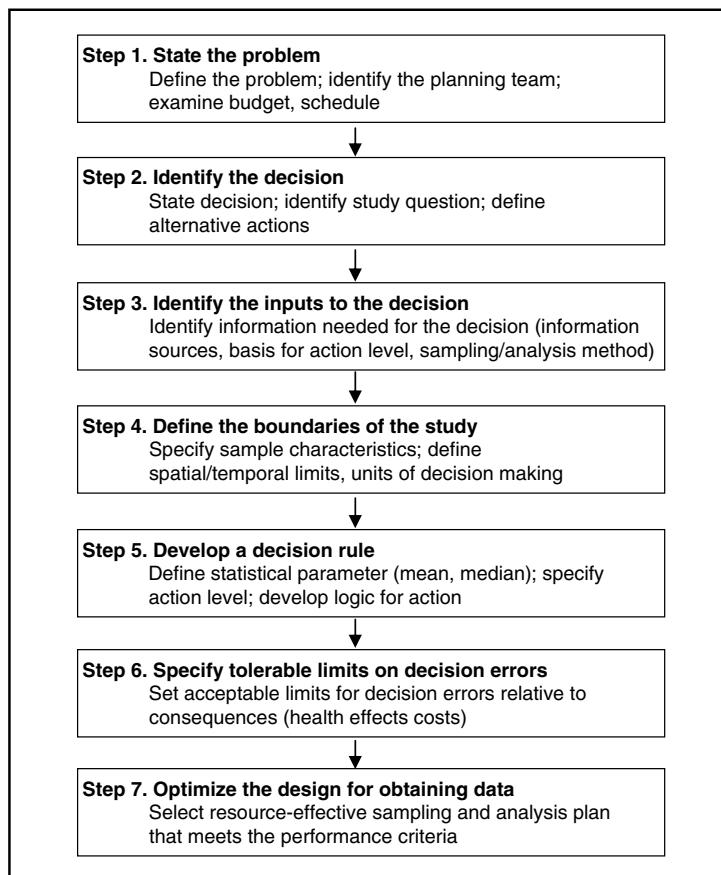


FIGURE 35.6 Data quality objectives process. *Source:* USEPA. 2000. *Guidance for the Data Quality Objectives Process—EPA QA/G-4 EPA/600/R-96/055.* Washington, DC, 100 p. August.

35.2.2.3 Step 3: Design Groundwater Monitoring Network

After hydrogeologic conditions have been evaluated and the DQOs have been established, the groundwater monitoring system can be designed. A three-phased procedure for designing a groundwater monitoring system is described below.

Phase I: Select Monitoring Locations: Locating the appropriate monitoring point locations is essential in designing a monitoring network capable of providing data of adequate quality to achieve the program objectives. At times, monitoring well locations may be prescribed by the regulations under which the groundwater monitoring program is being developed. For example, some regulations require monitoring locations be placed at a designated “point of compliance,” which is often at the property boundary or a groundwater discharge location. For other groundwater monitoring programs, the groundwater professional should select monitoring locations that provide the most reliable data needed to detect or assess a groundwater contaminant plume. To verify that the monitoring network can accomplish this goal, target monitoring zones must be selected based on the site hydrogeologic conditions and anticipated contaminant pathways, which are discussed in detail in Chapter 18 and Chapter 19.

Examples of monitoring well location layouts for a detection monitoring program in both a typical unconfined and layered aquifer system are provided in Figure 35.4 and Figure 35.6, respectively. As shown in Figure 35.5, the locations and orientation of confining units have a significant effect on potential contaminant migration paths and therefore the vertical spacing of monitoring wells. Additionally, the physical and chemical characteristics contaminant must also be considered when identifying target

monitoring zones and selecting monitoring locations. For example, the presence of LNAPLs (i.e., floaters) or DNAPLs (i.e., sinkers) in groundwater would require that monitoring locations be installed at the water table or at the top of confining units, respectively.

To facilitate the selection of monitoring locations, numerical groundwater flow models with particle-tracking capabilities (e.g., MEMO [Wilson et al. 1992], MODFLOW [McDonald and Harbaugh 1988]) can be used to predict contaminant migration pathways and identify potential target monitoring zones. Additionally, geostatistical analyses can also be used to optimize monitoring well layout both prior to implementing the groundwater monitoring program or during periodic reassessments of the groundwater monitoring program. Additional guidance for monitoring well optimization is provided by the U.S. Air Force Center for Environmental Excellence (1997).

Phase II: Select Monitoring Devices: Appropriate monitoring devices should be selected for obtaining the required samples or data from the target monitoring zones. Groundwater monitoring programs most often incorporate monitoring wells, piezometers, and groundwater discharge features as monitoring points. However, in some instances, alternative monitoring devices such as lysimeters, soil gas probes, direct push technology (DPT) samplers or other devices may be incorporated into the monitoring program. A description of alternative monitoring devices is provided in Section 35.3.5.

Phase III: Design the Monitoring Features: Finally, after the monitoring features have been identified, they should be designed as described in Section 35.3 to meet the specific goals of the monitoring program and to provide accurate, representative samples of groundwater.

35.2.2.4 Step 4: Establish Sampling and Analysis Methods

Each aspect of groundwater sampling and analysis, from extraction of the sample from the aquifer to laboratory analysis, can affect the results of the monitoring event. Therefore, it is important to specify sampling and analysis procedures that promote collection of samples representative of groundwater quality. When establishing the sampling procedures the groundwater monitoring professional must consider several factors including the monitoring point condition and design (e.g., depth, diameter, etc.), the type and properties of contaminants of concern (COCs) that will be monitored, access restrictions, availability of resources such as electric power, and any special sample or purge water handling requirements. The correct sampling procedure is an essential element of the groundwater monitoring planning process, as many of the errors or data quality issues associated with groundwater data collection occur as a result of bias introduced by the sampling procedure. Guidelines for sampling are presented in Section 35.4.

Some regulatory programs, such as RCRA Subtitle C and Subtitle D, prescribe a specific list of analytes that must be monitored at a particular site. At other sites, the groundwater professional needs to identify the COCs that are the basis of the monitoring program. COCs typically include the regulated substances contained in the source, as well as any daughter products that might be produced by reactions as the contaminants migrate. To minimize the number of tests required to characterize groundwater quality, the presence of chemical constituents is sometimes evaluated by measuring “indicator parameters,” including pH, total organic carbon (TOC), total organic halides (TOX), and specific conductance; these parameters may be indicative of the presence of elevated concentrations of organic, chloride-based, or metallic constituents, respectively, in groundwater. Use of appropriate sampling and analysis methods (as described in Section 35.4 and Section 35.5) promotes the development of consistent, high-quality data and minimizes the possibility of invalid or incomplete data. Guidelines for analysis of groundwater samples are presented in Section 35.5.

35.2.2.5 Step 5: Establish Data Evaluation Methods

After the groundwater samples have been analyzed, the data should be evaluated to provide a basis for concluding the presence or absence of groundwater contaminants. Such evaluations typically involve statistical analysis of the data to identify significant differences in groundwater quality. To guide the data evaluation process, the groundwater monitoring plan should contain descriptions of the methods that will be used to evaluate the monitoring data and the criteria for initiating response actions. Data evaluation methods are described in Section 35.6 of this chapter.

TABLE 35.2 Example Response Criteria and Corresponding Response Actions for Detection Monitoring Program

Response criteria	Response action
<ul style="list-style-type: none"> • Significant change in groundwater elevation • Change in total depth of well 	<ul style="list-style-type: none"> • Consider seasonal effects on water level, or impacts from groundwater withdrawals (e.g., water supply or irrigation wells) • If necessary, remeasure groundwater levels • Redevelop well • If necessary, reconstruct well • Evidence of intrusion in well • Evaluate possibility of vandalism or tampering • Protective cover damage • If necessary, reconstruct well • Reevaluate validity of monitoring data • Resample wells to confirm statistically significant increase • Continue detection monitoring
<ul style="list-style-type: none"> • Statistically significant increase in monitored parameters • Statistically significant increase not attributable to source or found to be invalid • Statistically significant increase confirmed and attributed to the potential source being monitored 	<ul style="list-style-type: none"> • Revise groundwater monitoring plan to prevent recurrence of problem during future monitoring events. • Consider validity of statistical analysis • Evaluate possibility of release from monitored source • Consider possibility of impact by other sources • Notify appropriate regulatory authorities • Develop plan of action with regulators; possible plans could require the following: <ul style="list-style-type: none"> • Prepare assessment monitoring plan • Submit assessment monitoring plan to regulatory agency for review (if necessary) • Implement assessment monitoring plan

35.2.2.6 Step 6: Develop Response Criteria and Response Actions

Response criteria are results of groundwater monitoring events for which specific responses must be implemented. For example, in a detection monitoring program, an appropriate response action for the response criteria “*if a statistically significant increase in a contaminant concentration is observed*” could be “*begin assessment monitoring*.” Examples of response criteria for a detection monitoring program and potentially appropriate response actions are provided in Table 35.2. The response actions for a detection monitoring program should be designed to confirm the presence of the detected contaminant or the statistically significant increase; detailed assessment or remediation of the detected contaminants may be delayed until the start of an assessment monitoring program (see Section 35.1.2.2) or a corrective action monitoring program (see Section 35.1.2.3). The groundwater monitoring plan should also identify follow-on activities for each possible outcome of a response action (e.g., “*prepare, submit, and implement an assessment monitoring program*” if the presence of contaminants is confirmed, or “*continue detection monitoring program*” if the presence of contaminants is not confirmed).

35.2.2.7 Step 7: Evaluate Background Conditions

To evaluate the impact of a potential source on groundwater quality, the characteristics of nonimpacted (i.e., upgradient) or baseline (i.e., existing conditions in the study area) groundwater must first be evaluated. This step is usually performed concurrently with the collection and analysis of geochemical conditions and includes an assessment of the ambient concentrations of target constituents. Background water quality data provide a baseline against which the results of future groundwater monitoring events can be compared. When establishing background conditions, a careful examination of the site should be made to verify that the location selected for the background sample is unaffected by either the potential source or by other sources; for example, locations that could be impacted by incidental spills (such as waste material storage areas or parking areas) should not be selected as background sampling locations. In

TABLE 35.3 Typical Groundwater Monitoring Features

	Primary use	Advantages and limitations
Wells	Groundwater sampling	Extensive record of successful use; can be targeted to specific site limitations and contaminants of concern; flexible design.
Piezometers	Groundwater potentiometric surface measurement	Allows measurement of potentiometric surface at a point; can be outfitted to provide continuous monitoring of groundwater elevation; not well suited for groundwater sampling.
Lysimeters	Sampling liquid from vadose zone	Allows characterization of liquid that could eventually impact groundwater quality; may allow loss of VOCs in samples; construction materials (usually ceramic) may affect chemical composition of samples.
Seeps	Surface sampling of groundwater seepage	Low cost for establishing sampling point; gives good indication of quality of water with which humans or organisms may come into contact.
Geophysical techniques	Plume delineation	Large areas can be quickly delineated at the “survey” level, inexpensively; can be used on ground surface or in boreholes; cannot be used to detect presence of specific constituents.
Sampling probes (e.g., HydroPunch [®] , Geoprobe [®])	Plume delineation	Allows quick sample collection without incurring the cost of installing a monitoring well; limited to one-time sampling of groundwater.

addition, sufficient baseline data should be collected to evaluate temporary variations such as tidal effects or seasonal effects. Following collection of sufficient baseline data, the monitoring program can begin.

35.3 Design and Installation of Groundwater Monitoring Devices

35.3.1 Overview of Groundwater Monitoring Devices

The purpose of this section is to identify and describe installation methods for typical groundwater monitoring devices. In Table 35.3, some typical groundwater monitoring features are identified. Each feature has a purpose and a situation for which it is best suited. The most common device used to monitor groundwater quality is the monitoring well. Installation of groundwater monitoring wells is described in Section 35.3.2.3, and other types of monitoring features are described in the remainder of this section.

35.3.2 Groundwater Monitoring Wells

35.3.2.1 General Design of Groundwater Monitoring Wells

The purpose of a groundwater monitoring well is to provide access to the target monitoring zone for collection of a representative sample of groundwater. The representativeness of the sample may be affected by installation of the well or by the materials used to construct the well; the design of the well must account

for these factors. In this section, groundwater monitoring wells and their applicability are described. The discussion presented in this section should be considered to be a general guide; site-specific conditions and applicable regulatory requirements should be considered over these guidelines when designing a groundwater monitoring well.

Standard approaches for design of groundwater monitoring wells are presented by a number of agencies and organizations, including the USEPA (1991a, 1992a) and the American Society for Testing Materials (ASTM, 1995). Examples of typical groundwater monitoring well designs are presented in Figure 35.6. The design shown in Figure 35.5(a) incorporates several features that minimize the possibility of introducing contaminants into the well (e.g., the protective cover, the bentonite seal, and the well apron). The monitoring well design shown in Figure 35.5(b) is typically used in situations where the well must be installed either: (1) beneath a confining feature (e.g., clay aquitard), where the additional casing prevents leakage from the upper aquifer to the lower aquifer; or (2) through a heavily contaminated soil layer, where the additional casing prevents mixing of drill cuttings from the target monitoring zone with the contaminants in the overlying or underlying portion of the aquifer. These designs can be modified as needed to meet site-specific conditions or regulatory requirements. For example, the number of monitoring points can be increased by installing multiple, discrete sampling points within a well; also, uncased, open boreholes can be used to monitor bedrock aquifers where migration of soil particles into the well is not expected to occur.

Some of the key features of the groundwater monitoring well, shown in Figure 35.7, are the wellscreen, filter pack, bentonite seal, cement grout backfill, concrete apron, and protective cover. The most important aspect of monitoring well design is the proper sizing and placement of the wellscreen or open-interval. When sizing the wellscreen, one must consider both the screen-interval size (i.e., slot size and screen type) and screen length for the proposed monitoring well. The screen-interval of the monitoring wellscreen should be sized based on the geologic materials outside of the screened interval and the proposed filter materials. Various guides have been established for well slot size selection including those described in USEPA (1991) and ASTM (1998). The wellscreen or open-interval length should be limited to the target monitoring zone. Monitoring wellscreen lengths typically range from 2 to 10 ft, with rare occurrences of up to 20 ft (EPA, 1991). To the extent possible, the screen length should be minimized to avoid dilution in the screened zone and to minimize interactions with, and potential contaminant migration to other zones within the aquifer. Also, as previously discussed, some regulatory agencies prescribe well design requirements particularly screen or open interval dimensions. The filter pack is intended to promote formation of a graded filter outside of the well to prevent migration of fine-grained soils into the well (because soil particles are composed of minerals that may be constituents of concern, the presence of fine-grained soils in a well can cause inaccurate groundwater monitoring results, as well as clog the well). The filter pack material should also have a characteristic particle size (i.e., the diameter greater than 85%, by weight, of the soil particles) that is bigger than the wellscreen slot size to prevent clogging of the wellscreen by the filter pack material. Similarly, the filter pack material should be capable of retaining the coarsest 15% of materials in the adjacent geologic formation. The bentonite seal is intended to prevent the cement grout backfill from migrating into the filter pack; the presence of grout in the filter pack could permanently compromise the validity of groundwater samples from the well. The concrete apron is intended to route surface water away from the well and to prevent downward migration of surface water into the wellscreen. The protective cover is intended to prevent unauthorized access to the well and to protect the exposed portion of the riser pipe from damage due to incidental contact.

When installing a groundwater monitoring well, the following potential problems should be anticipated and avoided to the extent possible (after Nielsen 1991; USEPA, 1993c):

- Use of well construction materials that are physically or chemically incompatible with either the surrounding natural earth materials or contaminants in the target monitoring zone and strong enough to prevent collapse under the stress applied by the soil.
- Improper selection of wellscreen sizes (screen sizes that are too large may allow siltation of the well, and screen sizes that are too small may prevent proper development of a graded filter around the well).

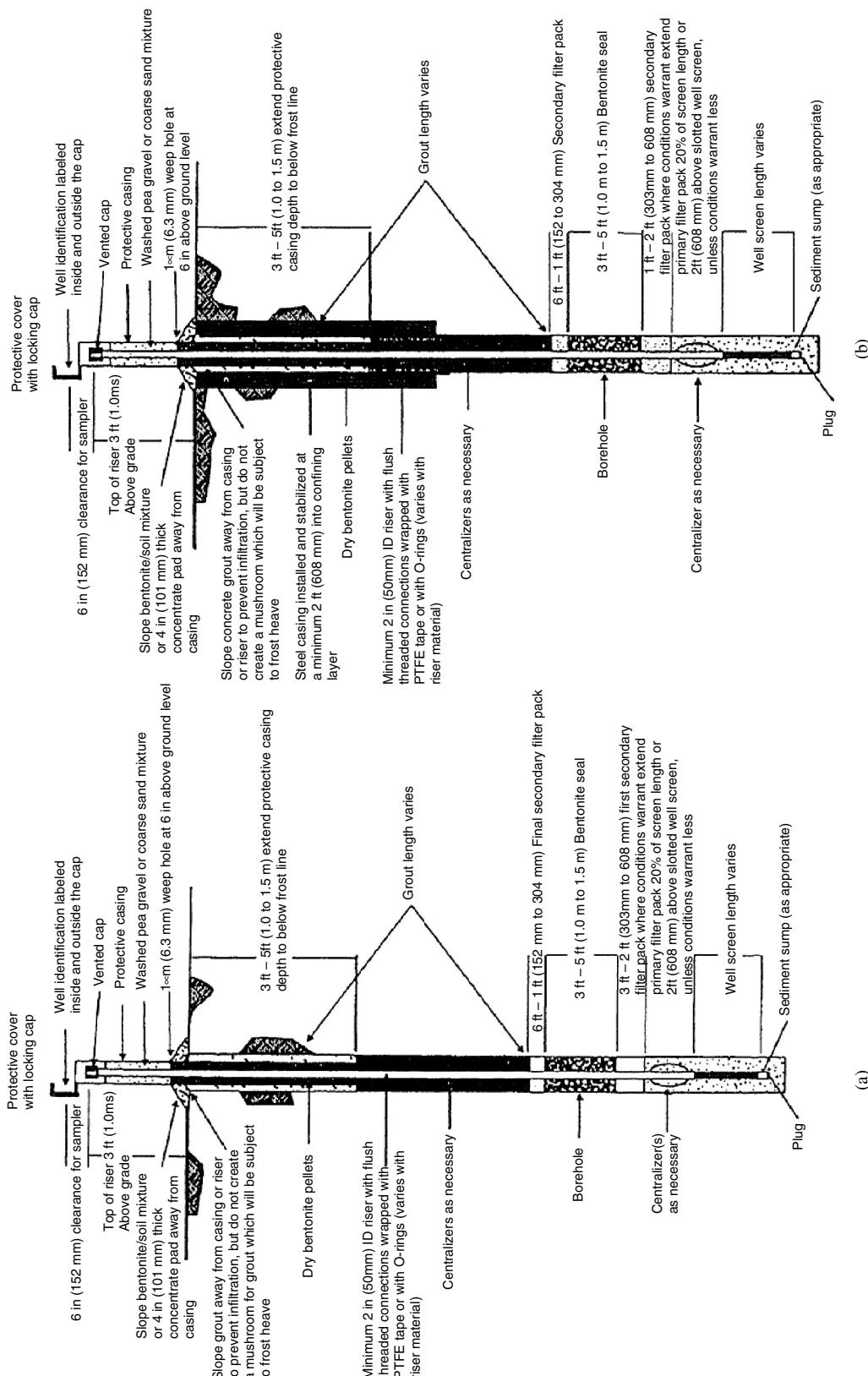


FIGURE 35.7 Monitoring well design.

- Placement of the screened interval of the well across stratigraphic zones, if the intent of monitoring is to sample discrete zones of the aquifer (this problem could also limit the use of the well for hydraulic conductivity testing of the aquifer).
- Improper selection or placement of filter pack material that could cause either siltation of the well or plugging of the wellscreen.
- Improper selection or placement of annular seal materials that can allow plugging of the filter pack, cross-linking of discrete water-bearing units, or migration of grout into the filter pack.
- Poor surface-protection measures that can allow damage to well casing materials or introduction of surface water into the well at the ground surface.
- Poor well development and evacuation techniques that may alter the aquifer formation around the wellscreen, cause excessive siltation of the filter pack and groundwater samples, or compromise well yield.

35.3.2.2 Drilling Methods

There are numerous drilling methods available for installing groundwater monitoring wells, each of which has attributes that may make it suited for a particular drilling location and well type. A summary of drilling methods that are commonly used to install groundwater monitoring wells is presented in Table 35.4. Also, the situations where each method is best suited are identified. More comprehensive discussions of drilling methods are presented by Hvorslev (1949), Driscoll (1986), ASTM (1998), and USEPA (1991).

35.3.2.3 Monitoring Well Installation

Careful installation of groundwater monitoring wells is an important step in producing representative samples of groundwater from the target monitoring zone. In this section, a generalized approach to installation of monitoring wells is presented. Further information and guidance on design and installation of monitoring wells is presented by USEPA (1975, 1989d) and ASTM (1990).

Prepare for Field Work: Before beginning monitoring well installation activities in the field, the groundwater field professional should develop as complete an understanding as possible of the conditions that will be encountered. This can be accomplished by reviewing the groundwater monitoring plan, regional and site-specific hydrogeologic information, the site health and safety plan (which must be prepared if hazardous conditions might be encountered), and the logs for installation of existing wells or other borings near the proposed well locations. The equipment that will be needed for the field work (e.g., water-level measuring instrument, personal protective equipment, sample jars, field analytical equipment, etc.) should be assembled well in advance of the field work. Monitoring equipment (e.g., water-level measuring equipment, field analytical equipment, etc.) should be calibrated. Well construction materials should be specified in advance of field work; recommendations for selection of well construction materials are provided in Table 35.5. Finally, the groundwater field professional should verify that the well construction materials that are procured by others meet the specific requirements of the groundwater monitoring plan.

Lay Out Location for Monitoring Well: The well location specified in the monitoring plan should be laid out before drilling begins. After the location is marked, the groundwater field professional should verify that the location is suitable for the well (e.g., not in a location where it is likely to be damaged, or in standing water, or where conditions are encountered that were not anticipated in the groundwater monitoring plan). The well location should also be sited in a location that will not encounter or interfere with underground utilities or structures. Most states require that public utility companies be notified prior to drilling or excavating activities so that underground utility lines are marked prior to performing the work. In addition, many industrial, commercial, and military facilities contain private underground utilities that may also need to be marked prior to well installation. Note that, after the well has been installed, the location should be surveyed by a licensed professional surveyor to provide an exact record of the location and top of casing elevation of the well.

Decontaminate Drilling Equipment and Well Construction Materials: Equipment or construction materials should be decontaminated before being introduced into the borehole using a high-pressure steam wash or other appropriate decontamination methods (see Moberly [1985] for a discussion of

TABLE 35.4 Well Drilling Selection Guide

Drilling method	Drilling fluid	Casing advance	Type of material drilled	Typical drilling depth, in ft ^a	Typical borehole sizes, in. borehole sizes, in.	Samples obtainable ^b	Coring possible	Reference section
Power auger (Hollow-stem)	None, water, mud	Yes	Soil, weathered rock	<150	5–22	S, F	Yes	6.2
Power auger (Solid-stem)	Water, mud	No	Soil, weathered rock	<150	2–10	S	Yes	6.3
Power bucket auger	None, water (below water table)	No	Soil, weathered rock	<150	18–48	S	Yes	6.4
Hand auger	None	No	Soil	<70 (above water table only) >1000	2–6	S	Yes	6.5
Direct fluid rotary	Water, mud	Yes	Soil, rock		2–36	S, R	Yes	7.3
Direct air rotary	Air, water, foam	Yes	Soil, rock	>1500	2–36	S, R, F	Yes	7.4
DTH hammer	Air, water, foam	Yes	Rock, boulders	<2000	4–16	R	Yes	7.5, 1
Wireline	Air, water, foam	Yes	Soil, rock	>1000	3–6	S, R, F	Yes	7.6
Reverse fluid rotary	Water, mud	Yes	Soil, rock	>2000	12–36	S, R, F	Yes	7.8
Reverse air rotary	Air, water, foam	Yes	Soil, rock	>1000	12–36	S, R, F	Yes	7.7
Cable tool	Water	Yes	Soil, rock	>5000	4–35	S, R, F (F-below water table)	Yes	8
Casing-advancer	Air, water, mud	Yes	Soil, rock, boulders	<2000	2–16	S, R, F	Yes	9
Direct-push technology	None	Yes	Soil	<100	1.5–3	S, F	Yes	10
Sonic (vibratory)	None, water, mud, air	Yes	Soil, rock, boulders	<500	4–12	S, R, F	Yes	11
Jet percussion	Water	No	Soil	<50	2–4	S	No	12
Jetting	Water	Yes	Soil	<50	4	S	No	12

^aActual achievable drilled depths will vary depending on the ambient geohydrologic conditions existing at the site and size of drilling equipment used. For example, large high-torque rigs can drill to greater depths than their smaller counterparts under favorable site conditions. Boreholes drilled using air/foam can reach greater depths more efficiently using two-stage positive-displacement compressors having the capability of developing working pressures of 250 to 250 psi and 500 to 750 cfm, particularly when submergence requires higher pressures. The smaller rotary-type compressors only are capable of producing a maximum working pressure of 125 psi and produce 500 to 1200 cfm. Likewise, the rig mast must be constructed to safely carry the anticipated working loads expected. To allow for contingencies, it is recommended that the rated capacity of the mast be at least twice the anticipated weight load or normal pulling load.

^bSoil = S (Cuttings), Rock = R (Cuttings), Fluid = F (some samples might require accessory sampling devices to obtain).

Source: From American Society for Testing and Materials (ASTM), Standard D 6286 -98 "Standard Guide for Selection of Drilling Methods for Environmental Site Characterization." 1998, West Conshohocken, Pennsylvania.

TABLE 35.5 Recommendations for Selection of Well Casing Materials

If	Use	Do not use
Well depth in range of 225–375 ft (68.6–114 m)	PVC, ABS, SS	PTFE
Well depth in range of 1200–2000 ft (366–610 m)	SS	PVC or ABS
pH < 7.0, DO > 2 ppm, H ₂ S ≥ 1 ppm, TDS > 1000 ppm, CO ₂ > 50 ppm, or Cl ⁻ > 500 ppm	PVC, ABS, or PTFE	SS
A neat PVC solvent/softening agent is present or if the aqueous concentration of the PVC solvent/softening agent exceeds 0.25 times its solubility in water	SS or PTFE	PVC
In All Cases		
Do not use	Use	
Solvent bonded joints for PVC casings	Threaded PVC casings	
Welded stainless joints	Threaded SS casings	
Any PVC well casing that is not NSF-ASTM approved — D-1785 and F-480	ASTM-NSF approved PVC well casings — D-1785 and F-480	
Any stainless steel casing that is not ASTM approved — A312	ASTM approved SS 304 and SS 316 casings — A312	
Any ABS well casing that is not ASTM approved	ASTM approved ABS casings — F-480	

DO = dissolved oxygen; TDS = total dissolved solids; PVC = polyvinyl chloride; SS = stainless steel; ABS = acrylonitrile-butadiene-styrene; PTFE = polytetrafluoroethylene

Source: From USEPA. 1992a. *RCRA Groundwater Monitoring: Draft Technical Guidance*, Office of Solid Waste, EPA/530-4-93-001, PB93-139-350, Washington, DC.

decontamination procedures). Decontamination minimizes the possibility of unintentionally introducing contaminants into the borehole. Drilling fluids should be inert; the use of oil and grease on drilling equipment should be restricted.

Drill Borehole: The borehole should be drilled using a technique that minimizes disturbance to the aquifer formation and does not impact the ability of the well to produce a representative groundwater sample. The potential impacts of drilling on groundwater samples are illustrated in Table 35.6, and candidate drilling techniques are identified in Table 35.6. During drilling of the borehole, the cuttings should be examined by the driller and the groundwater field professional, and a stratigraphic log should be made of the materials that are encountered; a sample boring log form is shown in Figure 35.8. The drill cuttings, drill fluid, or cores should be examined for unexpected materials (e.g., soils that are not described on logs for nearby wells, contaminants, indications of stratigraphic units at unexpected locations, etc.). The groundwater field professional should verify that the borehole is open throughout its depth, is reasonably straight, and is terminated within the target monitoring zone so that the well can be properly constructed; these can be verified by measuring the total depth of the boring after the drilling equipment has been removed from the boring. If downhole geophysical logging is required by the groundwater monitoring plan, then it may be performed at this time. (Note that for some geophysical logging techniques, logging must be performed while the borehole is open, prior to well construction.)

Inspect Well Construction Materials: Well construction materials should be inspected to verify that they meet the requirements of the groundwater monitoring plan. Wellscreen and riser pipes should be clean and undamaged; screen slots should be uniform, open, and continuous throughout the screened interval;

TABLE 35.6 Effects of Drilling Methods on Water Samples

Methods	Advantages	Disadvantages
Air rotary	<ul style="list-style-type: none"> Drilling fluid is not always used, minimizing contamination and dilution problems 	<ul style="list-style-type: none"> When more than one water-bearing zone is encountered and hydrostatic pressures are different, flow between zones occurs after drilling is completed but before the hole is cased and grouted Oil from compressor may be introduced to geologic formation Use of foam additives containing organic materials can interfere with both organic and inorganic analyses
Mud rotary	<ul style="list-style-type: none"> Borehole can be kept open inexpensively Sample collection is relatively easy 	<ul style="list-style-type: none"> Drilling fluid that mixes with formation is difficult to remove and can cause contamination Fluid circulation can cause vertical mixing of contaminants Drilling fluids and additives can interfere with subsequent water quality analyses Lubricants may cause contamination Large-diameter hole makes it difficult to assure adequate grouting
Bucket auger	<ul style="list-style-type: none"> No drilling fluid is used, minimizing contamination and dilution problems 	
Solid stem auger	<ul style="list-style-type: none"> No drilling fluid is used, minimizing contamination and dilution problems Can avoid use of lubricants 	<ul style="list-style-type: none"> Because auger must be removed before well can be set, vertical mixing can occur between water-bearing zones Must continuously add water in soft or loose formations Can cause vertical mixing of both formation water and geologic materials Can cause vertical mixing of geologic materials
Hollow stem auger	<ul style="list-style-type: none"> No drilling fluid is used, minimizing contamination and dilution problems Can avoid use of lubricants Formation waters can be sampled during drilling Well can be installed as augers are removed, decreasing interaction with water from higher water-bearing zones 	<ul style="list-style-type: none"> Can cause vertical mixing of formation waters if augers are removed before well is installed
Cable tool	<ul style="list-style-type: none"> Little or no drilling fluid required 	<ul style="list-style-type: none"> Contamination of aquifer is possible if drilling fluid is used Slight potential for vertical mixing as casing is driven
Jetting	<ul style="list-style-type: none"> May be only alternative where drill rig access is poor 	<ul style="list-style-type: none"> Large quantities of water or drilling fluid are introduced into and above sampled formation Cannot isolate target monitoring zone with a grout seal

Source: From Driscoll, F.G. 1986. *Groundwater and Wells*. Johnson Division, St. Paul, MN. With permission.

FIGURE 35.8 Typical boring/well installation form. (From Nielsen, D.M., ed. 1991. *Practical Handbook of Ground-Water Monitoring*. Lewis Publishers, Chelsea, MI. With permission.)

threaded connections at ends of pipes should be properly formed and should provide a tight connection. Bentonite seal materials should be delivered to the site in sealed containers and should not be allowed to hydrate before installation in the borehole. Filter pack material should be tested for particle-size distribution to verify that it meets the gradation requirements of the groundwater monitoring plan. Before installation, the length of each section of the riser pipe should be measured to the nearest hundredth of a foot or to the nearest centimeter so that, once installed, the depth to the top and bottom of the wellscreen is known.

Install Wellscreen and Riser Pipe: The riser pipes should be installed carefully to prevent damage during installation. The completed riser pipe should extend about 3 ft (1m) above the ground surface, as shown in Figure 35.7. Alternatively, if the presence of an above-ground riser causes an unacceptable obstacle at the site, the riser pipe can be terminated at ground level in a vault that is specially designed to prevent leakage into the well or damage to the well. The wellscreen should be in the borehole, using a centralizer if necessary; if the wellscreen is not centered in the borehole, it might not be possible to install the filter pack material evenly around the pipe, which might prevent proper development of a graded filter around the wellscreen.

Install Filter Pack, Seal, and Backfill Material: The filter pack material should be installed around the wellscreen slowly, to prevent bridging of the material between the wellscreen and the soils in the adjacent aquifer. Bridging, which leaves a portion of the wellscreen unsupported, could result in crushing of the riser pipe or migration of fine-grained soil particles into the well. For deep boreholes, the filter pack material may need to be installed using a tremie pipe, which is used to deposit the material from the bottom up (see Driscoll [1986, p. 477] for additional discussion). The bentonite seal should be installed evenly, above and on all sides of the top of the filter pack using either dry bentonite pellets or bentonite slurry. If dry bentonite is used, then it should be allowed to hydrate in the well for a period of at least one hour (preferably overnight) before the grout backfill material is installed over the seal. The backfill material should be installed carefully to prevent damage to the bentonite seal; backfill materials may include cement grout, drill cuttings (if uncontaminated), or stabilized soil, depending on the requirements of applicable regulations. The depth to the top of the filter pack, seal, and backfill material should be measured and recorded before the overlying component of the well is installed; when making these measurements, the groundwater field professional should confirm that the wellscreen section is located completely within the filter pack. Samples of well construction materials should be retained by the contractor for a period of time designated in the well construction specifications.

Complete Above-Ground Portion of Well: The above-ground portions of the well should be constructed in a manner that accommodates sampling of the well and protects the well from incidental damage. The actual configuration of the aboveground portion of the well may depend on applicable regulations; a recommended configuration is shown in Figure 35.7. For above-ground risers, the riser should be located at short depth (i.e., about 3 to 6 in., (75 to 150 mm)) beneath the top of the protective casing, and the reference point (for measuring the depth to groundwater and total depth of the well) should be clearly and permanently marked on the riser pipe. The annular space between the protective casing and the riser pipe can be filled with sand or gravel (so that items dropped in the casing can be easily retrieved) and a hole should be drilled in the casing about 6 in. (150 mm) above the ground surface to allow drainage of water that may collect in the casing during sampling. A concrete pad should be constructed around the protective casing at the ground surface and sloped downward, away from the protective casing, to prevent accumulation of standing water around the well and infiltration of surface water into the well. The volume of grout placed above ground should be minimized to prevent frost heave of the grout pad during freezing weather.

35.3.2.4 Well Development Techniques

The purposes of well development are to remove sediment from the well and surrounding aquifer materials impacted by well installation, and to form a graded filter at the interface between the aquifer materials and the filter pack material. Well development involves removing water from the well after construction is completed. Development prior to initial well sampling is essential if a representative sample is to be obtained. Common well development techniques are described by Driscoll (1986) and ASTM (1994). Hand-bailing techniques are widely used but are time consuming and may alter the geologic formation

that surrounds the filter pack, resulting in high values of suspended solids in groundwater samples or even failure of the filter pack. Mechanical development techniques (e.g., swabbing, surging, or low-flow pumping) require less physical effort and can be tailored to the specific limitations of the aquifer and well construction features. Wells should be developed until the pH, specific conductivity, and color (i.e., cloudiness) of the water stabilize; note that, if pH and specific conductivity remain consistently higher in the well than in the aquifer, there may be leakage into the filter pack from the overlying cement grout (grout typically has a pH of about 10 and a specific conductivity of about 1500 ($\mu\text{mhos}/\text{cm}$)). Cloudy development of water may be an indication that the filter pack is not filtering the aquifer formation material effectively. Such problems indicate that groundwater samples from the well may not be representative of groundwater in the aquifer near the well. Problems encountered during development should be examined carefully and, if necessary, the well should be redeveloped or decommissioned (using techniques described in Section 35.3.2.6) and then reinstalled. Also, a water level inconsistent with surrounding water levels can indicate a problem with well completion.

35.3.2.5 QC/QA of Drilling and Well Installation

To verify that the monitoring well is installed properly, the groundwater professional should implement a program of Quality Control and Quality Assurance (QC and QA) during drilling and installation. In the context of monitoring well installation, QC refers to the procedures used to confirm that the materials and methods used meet the requirements of the groundwater monitoring well design, and QA refers to the activities that are performed to verify that QC procedures were properly implemented. QC and QA procedures are typically specified in the groundwater monitoring plan. QC activities could include testing well construction materials, verifying that appropriate drilling methods are used, observing the performance of the drilling equipment, and documenting drilling and well construction details. QA is typically provided by the groundwater field professional (i.e., a geologist, an engineer, or an engineering technician) and includes oversight of the drill crew's activities, inspection of the materials of construction, and review of submittals from the driller (e.g., material specification sheets, results of filter pack laboratory gradation tests, etc.) Cooperation between the driller and the groundwater field professional is essential for quick resolution of problems that may arise during well installation.

An essential element of QC/QA is documentation of field activities. Well installation activities should be recorded in a boring log; a sample boring log form is provided in Figure 35.8. The details of well development should also be documented (e.g., pH and specific conductivity measurements for each purged well volume, cloudiness of purge water, time required to purge the well, etc.)

35.3.2.6 Decommissioning of Groundwater Monitoring Wells

In many jurisdictions, groundwater monitoring wells that are not actively used must be decommissioned (also, the terms “destroyed” or “abandoned” are used in some jurisdictions to denote decommissioning). Decommissioning involves removing the well from use so that the well does not act as either a possible conduit for contaminant migration or uncontrolled access to groundwater resources. Decommissioning can be performed either by: (1) installing cement grout within the wellscreen and filter pack and then removing the casing and riser pipe (if possible); or (2) over coring the well and removing all of the well construction materials from the ground, then backfilling the borehole with appropriate material (e.g., a bentonite seal in low-permeability zones, soil in unconsolidated soil zones, or cement grout in rock formations).

35.3.3 Piezometers

A piezometer is used to measure the hydraulic head at a specific location within the aquifer. As piezometers are not designed for the collection of groundwater samples, their design may be different from the design of a groundwater monitoring well. Attributes typical of a piezometer include a small diameter borehole, casing, and screen (to minimize the impact of the piezometer on groundwater flow in the aquifer), small screened interval (because piezometers having long screened intervals provide the average head over a

large depth, instead of the head at a specific location), small or no filter pack (because the filter pack may impact the hydraulic head near the screened interval of the piezometer), and a very effective seal at the top of the screened interval (to prevent intrusion of water from above the measuring point).

35.3.4 Groundwater Discharges

Groundwater quality samples and data can also be obtained from groundwater discharge to the surface at locations such as springs, seeps, or drains or from other groundwater discharges such as production/extraction wells. These discharge locations are often incorporated into groundwater monitoring programs because they typically provide a measure of groundwater quality at a location where a receptor may be present (i.e., drinking water well, discharge to a surface water body) and are therefore often used as a point of compliance. When using discharge locations as monitoring points, the DQOs of the groundwater monitoring program should be carefully considered to verify that the groundwater quality obtained at the discharge locations are comparable to the data collected at other monitoring points in the program. Variability in groundwater quality and sampling error at discharge locations can occur through interaction with the atmosphere and dilution with surface water bodies or groundwater from outside of the study area (depending upon the capture zone of the well or discharge feature). This variability can affect contaminant concentrations observed in discharge samples and bias the results of the analysis.

35.3.5 Alternative Groundwater Monitoring Devices

At times, groundwater monitoring programs may incorporate alternative monitoring devices that either provide short-term data to supplement the conventional monitoring program, provide indirect data that can be inferred to assess potential releases (i.e., geophysical variations); or provide data regarding conditions outside of the target monitoring zone (e.g., soil gas, vadose zone, etc.) that can be extrapolated to assess groundwater quality within the target monitoring zone. These techniques are often incorporated into monitoring programs at sites that may be small, dry, or remote, where the use of conventional groundwater monitoring devices and sampling techniques may be impractical or cost-prohibitive. A detailed discussion of alternative monitoring techniques for these types of sites is provided by USEPA (1995).

Alternative monitoring devices that are used to collect and extract samples of groundwater include direct push technology (DPT) samplers such as GeoProbe[®], Hydropunch[®] (Edge and Cordry 1989), or other specialized grab samplers equipped on DPT or cone penetrometer testing (CPT) devices. These devices allow discrete groundwater samples to be collected directly from the formation. Some DPT equipment can also be fitted with specialized probes to monitor constituent concentrations or other parameters (e.g., specific conductance) vertically through saturated and unsaturated zone. Some examples of specialized probes include membrane interface probes (MIP), electrochemical sensors, and Site Characterization and Analysis Penetrometer (SCAP) or ROSTTM probes, which analyze the groundwater *in situ* rather than extracting a sample to the surface. Descriptions of selected monitoring probes and sensors are provided in Section 35.4.3. DPT techniques are typically used during site characterization and assessment because they provide only temporary access to the formation and therefore allow only one-time sampling. However, they can be incorporated into monitoring programs if data is needed to supplement long-term monitoring results. The advantages of DPT techniques are that they are low cost, allow rapid collection of samples, and are applicable to a wide range of soil types. Additional detail regarding DPT groundwater sampling methodologies is provided in ASTM (1996).

Various surface and borehole geophysical methods including electrical induction, electrical resistivity, electromagnetic induction and induced polarization, and complex resistivity can be used to monitor variations in aquifer and groundwater properties that can be indicative of releases. Geophysical evaluations are useful in plume delineation and monitoring as they can provide a gross measure of changes in an aquifer system, but most are not sensitive enough to detect minor changes in concentrations or provide precise concentrations values for specific contaminants that may be of interest.

Soil gas probes and soil gas survey data can be used to infer groundwater quality impacts from some volatile compounds in shallow groundwater regimes. Soil gas surveys involve inserting a temporary or permanent probe or tube into the vadose zone and extracting gas from the subsurface to collect a sample. The soil gas data do not provide a direct measurement of groundwater quality and are limited to assessing volatile compounds. In addition, soil gas samples require special handling and shipping considerations. More information on soil gas survey procedures, requirements, and limitations is provided by ASTM (2001) and USEPA (1995).

Lysimeters are used to collect water samples from the vadose zone either passively (by gravity collection) or through applying a vacuum or pressure to extract water from the unsaturated soils. The use of lysimeter in a groundwater monitoring program typically occurs in recharge areas near source zones. Lysimeter data from these areas can be used to obtain concentration data for waters recharging the aquifer through the source zone that may be used to evaluate contaminant flux and fate and transport. However, lysimeter samples are often difficult to obtain and pressure changes induced by some lysimeters may affect the solution chemistry of some contaminants thereby affecting data quality.

35.4 Groundwater Sampling

35.4.1 General Approach to Sampling Groundwater Monitoring Wells

A general approach to the sampling of groundwater monitoring wells is presented in a stepwise fashion in this section. The following procedures are a compilation of groundwater sampling techniques recommended by the USEPA Yeskis and Zavala (2002), Puls and Barcelona (1996), Nielsen (1991), Driscoll (1986), and ASTM (2001). In contaminated areas, sampling should proceed from the least contaminated well to the most contaminated well if any sampling equipment is to be used at more than one well.

35.4.1.1 Step 1: Inspect Wells and Area Surrounding Wells

The well should be examined for evidence of tampering, damage, or other activity that could compromise the ability of the well to produce a representative groundwater sample. The ground surface near the well should be examined for signs of recent activity (including accidental spills of constituents that could impact groundwater quality) and the area being monitored should be examined for changes since the last sampling event (e.g., new construction, new storage activities for waste materials, agricultural activity, etc.); such changes may require reexamination of the adequacy of the monitoring well network layout. Observations should be recorded by the field sampling professional for future reference.

35.4.1.2 Step 2: Measure Depth to Groundwater and Total Depth of Well

The first activity to be performed after the well is inspected should be measurement of the depth to groundwater and the total depth of the well. There are many techniques available for measuring the depth to groundwater in a well, as shown in Table 35.7, including hand measurement using a wetted tape (which is simple and inexpensive, but somewhat time consuming) and electronic measurement (which is more expensive but less time consuming, and can be configured to provide a continuous record of data, if needed). All measurement equipment should be decontaminated (as described by USEPA [1992a]) before introducing it into the well. Measurements of the depth to groundwater that are made shortly after sampling the well may not be representative of the actual depth to groundwater. Measurement of the total depth of a well is important for evaluating changes in the total depth; for example, a decrease in the total depth could be indicative of siltation in the bottom of the well, or a crushed wellscreen or riser pipe.

35.4.1.3 Step 3: Purge Groundwater from Well

The purpose of purging the well is to remove stagnant water from the well and to cause groundwater from the aquifer formation around the well to flow into the well, allowing collection of a representative groundwater sample from the target monitoring zone. Depending upon the hydrogeologic conditions and monitoring device design, purging may not be necessary. However, most commonly used groundwater

TABLE 35.7 Summary of Methods for Manual Measurement of Well Water Levels in Nonartesian and Artesian Wells

Measurement method	Measurement accuracy (ft)	Major interference or disadvantage
Nonartesian Wells		
Wetted-tape	0.01	Cascading water or casing wall water
Air-line	0.25	Air line or fitting leaks; gauge inaccuracies
Electrical	0.02 to 0.1	Cable wear; presence of hydrocarbons on water surface
Transducer	0.01 to 0.1	Temperature changes; electronic drift; blocked capillary
Acoustic probe	0.02	Cascading water; hydrocarbon on well water surface
Ultrasonics	0.02 to 0.1	Temperature changes; well pipes and pumps; casing joints
Artesian Wells		
Casing extensions	0.1	Limited range; awkward to implement
Manometer/pressure gauge	0.1 to 0.5	Gauge inaccuracies; calibration required
Transducers	0.02	Temperature changes; electronic drift

Source: From Nielsen, D.M., ed. 1991. *Practical Handbook of Groundwater Monitoring*. Lewis Publishers, Chelsea, MI. With permission.

sampling techniques require purging to ensure that a sample representative of the formation water is obtained. Purge water may need to be collected and treated if there is a chance that it could be contaminated.

35.4.1.4 Step 4: Collect Groundwater Samples

Groundwater in the well should be sampled using the techniques described in Section 35.4.2 or Section 35.4.3. Extreme care should be used when collecting, preserving, storing, and shipping groundwater samples (as described in Section 35.4.5) to provide samples to the laboratory that are representative of groundwater quality. Sample quality is much more likely to be compromised in the field than in the controlled environment of a laboratory.

35.4.1.5 Step 5: Document Sampling Event

During the sample event, each step of the sampling process should be documented to allow detailed examination of the sampling procedures in the future. Recommended documentation procedures are described in Section 35.4.6. Good documentation is essential because it provides the basis for verifying the validity of the monitoring event. Also, if monitoring problems cannot be traced to laboratory error or resolved using field logs, then the notes of the field sampling professional may be valuable in identifying the source of the problem.

35.4.2 Overview of Groundwater Sampling Techniques

Groundwater sampling typically requires extracting and collecting a volume of groundwater from a monitoring device for subsequent analysis. A variety of sampling devices are used to extract an aliquot of groundwater from the target monitoring zone for use as a sample. Table 35.8 provides a groundwater sampling device matrix that details the characteristics of most common sample extraction devices (*get rid of in situ* and portable determination, most portables can be dedicated). In addition to the characteristics provided in the matrix in Table 35.8, the groundwater professional should consider factors such as the sampling technique, access to the monitoring locations, availability of power, and ease of operations when identifying the appropriate monitoring devices. Additionally, the groundwater professional must examine the potential for the sampling device to affect data quality. For example, although bailers have historically

TABLE 35.8 Generalized Groundwater Sampling Device Matrix^a

Device	Approx. maximum sample depth	Minimum well diameter	Delivery sample rate or volume ^a	Groundwater parameters									
				Inorganic					Organic				
				EC	pH	Redox	Major ions	Trace metals	Nitrate fluoride	Dissolved gases	Non-volatile	Volatile	TOC
Portable Sampling Devices													
<i>GRAB</i>													
Open bailer	No limit	1/2 in	Variable	•	□	•	•	•	□	□	□	•	•
Point Source	No limit	1/2 in	Variable	•	•	•	•	•	□	•	•	•	•
Bailer													
Syringe	No limit	1–1/2 in	0.01–0.2 gal	•	•	•	•	•	□	•	•	•	•
<i>Submersible</i>													
Gear-drive pump	200 ft	2 in	0–0.5 gpm	•	•	•	•	•	•	•	•	•	□
Bladder pump	400 ft	1–1/2 in	0–2 gpm	•	•	•	•	•	•	•	•	•	•
Helical rotor pump	160 ft	2 in	0–1.2 gpm	•	•	•	•	•	•	•	•	•	□
Pump													
Gas-driven piston pump	500 ft	1–1/2 in	0–0.5 gpm	•	□	•	•	•	□	□	•	•	□
Centrifugal (low-rate) pump	Variable	2 in	Variable	•	•	•	•	•	•	•	•	•	•
<i>Suction</i>													
Peristaltic pump	26 ft	1/2 in	0.01–0.3 gpm	•	□	•	•	•	□	□	•	•	•
<i>Gas Contact</i>													
Gas-lift pump	Variable	1 in	Variable	□	□	□	□	□	□	□	□	□	□
Gas-drive pump	150 ft	1 in	0.2 gpm	•	□	□	•	•	□	□	•	•	□
<i>In Situ^b Sampling Devices</i>													
Pneumatic pump	No limit	No limit	0.01–0.13 gpm	•	•	•	•	•	□	□	•	•	•

^a Sample delivery rates and volumes are average ranges based on typical field conditions. Actual delivery rates are a function of diameter of monitoring well, size, and capacity of sampling device, hydrogeologic conditions, and depth to sampling point. For all devices, delivery rate should be carefully controlled to prevent aeration and degassing of the sample.

^b Sampling devices on this chart are divided into two categories: (1) portable devices for sampling existing monitoring wells, and (2) *in situ* monitoring devices (often multilevel) that are permanently installed. Sampling device construction materials (including tubing, haul lines, etc.) should be evaluated for suitability in analyzing specific groundwater parameters. It is assumed on this chart that existing monitoring wells are properly installed and constructed of materials suitable for detection of the parameters of interest.

- Indicates device is generally suitable for application (assuming device is cleaned and operated properly and is constructed of suitable materials).
- Indicates device may be unsuitable or is untested for application.

Source: From USEPA, 1992a, RCRA Groundwater Monitoring: Draft Technical Guidance, Office of Solid Waste, EPA/530-4-93-001, PB93-139-350, Washington, D.C.

been commonly used in groundwater sampling activities, they are prone to disturbing the water column and sediments in the well that can significantly affect sample quality. In addition, the use of peristaltic or suction pumps for collection of samples for volatile organic compounds is often not recommended due to potential off-gassing resulting from the negative pressure exerted on the sample as it is being extracted. The groundwater professional should also determine whether portable sampling equipment or dedicated sampling equipment should be used in the groundwater monitoring program. This decision should be based on factors such as the proposed frequency of sampling, duration of the monitoring program, decontamination requirements, overall data quality, and cost analyses.

The groundwater sampling techniques most often used to collect samples from monitoring wells include: (1) the low-stress approach (or low-flow sampling, minimal drawdown sampling); (2) the well volume approach; (3) minimal purge approach; (4) discrete interval sampling; and (5) passive sampling. Selecting the appropriate sampling technique is based on several factors including well construction, hydrogeologic characteristics of the formation, and the proposed analytical requirements. Guidelines for selection of monitoring well purging and sampling approaches are provided in Table 35.9 and ASTM (2001).

35.4.2.1 Low-Stress (Minimal Drawdown) Approach

The low-stress approach is based on the assumption that purging and sampling at a low-flow rate that minimizes stress (drawdown) in the well allows a sample to be obtained from the aquifer formation at a discrete interval (i.e., the pump intake elevation) with minimal interaction with stagnant water in the well column. The low stress approach is performed using a variable speed sampling pump to extract water from the screen interval at a low-flow rate, typically between 0.1 and 0.5 l/min. During well purging, drawdown within the well is monitored and the flow rate is adjusted to minimize drawdown to less than 0.1 m (0.33 ft) from the static water level in the well. The pump intake should be positioned in the center, or just above the center of the wellscreen, to minimize the potential to entrain solids that may have collected at the bottom of the well. In wells where the wellscreen spans the water table, the pump intake should be positioned near the center of the water column rather than the wellscreen.

During well purging, the discharge is typically routed through a flow through cell and selected water quality parameters are monitored using portable field monitoring probes. These parameters typically include pH, ORP, specific conductance, dissolved oxygen, and turbidity. Purging is continued until the values for these parameters stabilize to established criteria. Table 35.10 provides stabilization criteria for common water quality indicator parameters. Following stabilization, samples are collected directly from the discharge of the pump at the same flow rate that was used during well purging.

This approach has become widely accepted in recent years and is often the preferred technique under many regulatory programs. This approach has become popular because it provides samples that are representative of the mobile load of contaminants present (i.e., dissolved and colloidal phases); minimizes sample disturbance and artifacts; and reduces the volume of purge water that oftentimes is subject to special treatment or disposal requirements. The disadvantages of this approach include: additional capital costs for specialized equipment, additional personnel training requirements, greater set up times in the field.

35.4.2.2 Well Volume Approach

Another sampling approach that is commonly used is the well volume approach. In this approach, the volume of the standing water column in the monitoring well is calculated based on water level measurements and well construction details. The well is then purged until a minimum number of well volumes (typically between three and five volumes) have been evacuated. This approach is typically performed using either dedicated or portable pumping systems. When using this approach the pump intake is placed in the water column to a depth that will not result in drawdown below the pump intake that could allow air to enter the pump. The pump should also be lowered slowly and set sufficiently above the bottom of the well to minimize disturbance of sediments in the well that could affect sample integrity.

TABLE 35.9 Applicability of Different Approaches for Purging and Sampling Monitoring Wells

	Low-stress approach	Well-volume approach	Others (such as passive diffusion samplers, <i>in situ</i> samplers, and other non-traditional groundwater sampling pumps)
Applicable Geologic Materials ^a	Materials with moderate to high hydraulic conductivities. May be applicable to some low hydraulic conductivities, if can meet minimal drawdown criteria.	Materials with low to high hydraulic conductivities	Materials with very low to high hydraulic conductivities
Aquifer/Plume Characterization Data Needs prior to Choosing Sampling Method	High definition of vertical hydraulic conductivity distribution and vertical containment distribution.	Plume and hydraulic conductivity distributions are less critical	May need to consider the degree of hydraulic and contaminant vertical distribution definition dependent on Data Quality Objectives and sampler type.
Constituent Types Method is Applicable	Mainly recommended for constituents that can be biased by turbidity in wells. Applicable for most other contaminants	Applicable for all sampling parameters. However, if turbidity values are elevated, low-stress approach may be more applicable if constituents of concern are turbidity sensitive.	Constituents of concern will be dependent on the type of sampler.
Data Quality Objectives	1) High resolution of plume definition both vertically and horizontally. 2) Reduce bias from other sampling methods if turbidity is of concern. 3) Target narrow sections of aquifer.	1) Basic characterization 2) Moderate to high resolution of plume definition (will be dependent on screen length). 3) Target sample composition to represent entire screened/open interval.	1) Can be applicable to basic site characterization, depending on sampler and methodology used. 2) Can reduce bias from other sampling methods. 3) May yield high resolution of plume definition.

^aHydraulic conductivities of aquifer materials vary from low hydraulic conductivities (clays, silts, very fine sands) to high conductivities (gravel, sands, weathered bedrock zones). This term for the use on this table is subjective, and is more dependent on the drawdown induced in a monitoring well when sampled with a groundwater sampling pump. For instance, in a well being pumped at 4 liters per minute (l/min) with less than 0.1 ft of drawdown, can be considered to have high hydraulic conductivity. A well that can sustain a 0.2 to 0.4 l/min pumping rate, but has more than 0.5 ft of drawdown, can be considered to have low hydraulic conductivity. To assign absolute values of hydraulic conductivities to well performance and sustainable pumping rate cannot be completed because of the many factors in monitoring well construction, such as well diameter, screen open area, and length of screen.

Source: Yeskis, D. and Zavala, B., 2002. Groundwater Issue Paper: Groundwater Sampling Guidelines for Superfund and RCRA Project Managers, U.S. Environmental Protection Agency, EPA/542/S-02-001. 53 p.

TABLE 35.10 Stabilization Criteria with References for Water-Quality Indicator Parameters

Parameter	Stabilization Criteria	Reference
pH	± 0.1	Puls and Barcelona 1996; Wilde et al., 1998
Specific electrical conductance (SEC)	$\pm 3\%$	Puls and Barcelona 1996
Oxidation-reduction potential (ORP)	$\pm 10 \text{ mV}$	Puls and Barcelona 1996
Turbidity	$\pm 10\%$ (when turbidity is greater than 10 NTUs)	Puls and Barcelona 1996; Wilde et al., 1998
Dissolved oxygen (DO)	$\pm 0.3 \text{ mg/l}$	Wilde et al., 1998

Source: Yeskis, D. and Zavala, B., 2002. Groundwater Issue Paper: Groundwater Sampling Guidelines for Superfund and RCRA Project Managers, U.S. Environmental Protection Agency, EPA/542/S-02-001. 53 p.

To minimize potential effects to the quality of the sample obtained, the well should be purged at a flow rate that does not produce turbulence within the well. Minimizing turbulent flow in the well will reduce the potential for volatilization or excessive turbidity. Care should also be taken to minimize drawdown below the top of the screened interval of the monitoring well that could expose the aquifer to the atmosphere. When the minimum number of well volumes is purged, samples containers are filled directly from the pump discharge and prepared for analysis.

Although the well volume approach has been replaced with the low-stress approach in many groundwater monitoring programs, regulations, and guidance, some regulatory agencies continue to require the well volume purge approach. The advantages to this approach are ease of implementation, cost and availability of sampling equipment, comparability with historical results, and minimal training requirements for field personnel. However, the representativeness of the samples obtained using this approach are sometimes compromised due to increased stress on the formation that may allow preferential flow from areas outside of the discrete zone. In addition, increased flow rates often create turbulence in the well that can result in loss of volatile compounds, oxidation of samples, and entrainment of sediments into the sample; all of which can affect data quality.

35.4.2.3 Minimal Purge Approach

The minimal purge approach is used in low-yielding zones where extensive well purging prior to sample collection may not be feasible. The minimal purge approach is typically performed with dedicated, downhole sampling equipment set at a discrete depth within the screened interval. Purging is performed to the extent that only the volume of water contained within the sampling system (tubing, sample intake, bladder, etc.) is evacuated with no attempt to remove water from the well casing or formation. The sample is then collected directly from the discharge of the sampling equipment. When minimal purging is used attempts should be made to minimize the purge volume by using small diameter tubing and the smallest feasible sampling chamber. The minimal purge approach may also be used in high-yielding aquifers where sufficient flow across the open interval is occurring such that formation water mixing with the stagnant water in the well column is minimal.

35.4.2.4 Discrete Interval Sampling

Discrete interval sampling requires no purging and is performed with down-hole, typically non-dedicated, samplers that are slowly lowered into the well to a discrete depth within the screen interval. A sample is then drawn into the sampling chamber and extracted. Several discrete interval samplers are commercially available and most common samplers operate using either a mechanical trigger mechanism to release a check valve at the desired depth that allows a sample to fill the sample chamber or pressurized systems that are lowered to the desired depth where the pressure is released or hydrostatic pressure forces water into the sampling chamber. Discrete interval sampling is often used in extremely low yielding wells where even minimal purging efforts are not possible. This method could also be used in higher yielding wells where flow across the screen interval is sufficient to minimize mixing with stagnant water from above the screen interval.

35.4.2.5 Passive Sampling

Passive sampling techniques include sampling protocols that are used to collect a sample of environmental media for analysis, but the sample is not an aliquot of groundwater. Passive samplers also rely on formation water moving through the wellscreen or open interval, rather than forcing water through the well by evacuating groundwater from the wellscreen or applying a vacuum. Passive samplers typically utilize specialized sampling media that interact with the groundwater in the well. The contaminants in the well are then transferred to the media via physical processes such as diffusion or sorption and the sampling media is extracted and analyzed.

The most common passive samplers are polyethylene diffusive bag (PDBs) samplers. A PDB consists of a low density polyethylene (LDPE) tube filled with distilled water that is lowered into the well. The PDB operates by allowing compounds to diffuse from the groundwater through the semi-permeable LDPE

membrane and into bag. When the concentration of compounds within the sampler reaches equilibrium with the surrounding groundwater, the bag is removed and the sample is extracted from the bag and analyzed. PDBs are not useful for all contaminants such as metals and other inorganics; however, many VOCs have been demonstrated to be reliably monitored using this technology. A detailed discussion of PDBs is provided by ITRC (2004).

35.4.3 *In Situ* or Alternative Monitoring Techniques

As an alternative to extracting and analyzing groundwater samples from a conventional groundwater monitoring device, some groundwater monitoring programs may include the use of direct monitoring devices, such as downhole sensors or probes that can analyze groundwater quality *in situ*. Direct monitoring devices are commonly used in groundwater monitoring programs because they provide capabilities such as long-term deployment, automated data collection functions, and data logging that can provide low-cost data collection with minimal effort. The following is a description of several direct monitoring devices that are commonly deployed at groundwater monitoring sites.

Multi-parameter probes: Various commercially available multi-parameter probes or sondes are available that can be used to collect *in situ* groundwater quality data for various analytes. Typically, these devices include single or multiple electric sensors that are capable of analyzing one or more of a variety of parameters, such as head pressure, pH, specific conductance, dissolved oxygen, or ORP.

Ion-selective probes: These probes can be used to detect specific ions in groundwater. The probes are designed to detect specific ions by producing an electric signal that can be compared to a reference signal for a specific ionic constituent. This method is particularly useful for preliminary groundwater characterization and tracer studies.

Fiber optic chemical sensors: These are also designed to detect a specific chemical in groundwater. The sensor is made of a reagent that is physically confined or chemically immobilized at the end of a fiber optic cable. The cable is inserted into a monitoring well, piezometer, or borehole, and then a signal is transmitted through the cable, which detects the constituent, if it is present. This method, which is also in the developmental stages, can be used on an extremely wide variety of constituents, uses portable equipment, and produces results at very low cost.

Specialized DPT or CPT probes: As described in Section 35.3.5, various types of CPT and DPT probes are available to obtain *in situ* measurement of contaminant concentrations in groundwater as boreholes are being advanced. Some DPT or CPT probes, such as the MIP, utilize a semi-permeable membrane to collect vapor samples that are extracted to the surface and analyzed using a variety of instruments. Other DPT probes utilize technologies such as laser-induced fluorescence (e.g., the ROST™ system or SCAP) or x-ray fluorescence to assess organic or metal contaminant concentrations in soils or groundwater within the borehole. These systems are generally used for site characterization rather than long-term monitoring events because they are deployed during DPT or CPT borehole installation.

35.4.4 Quality Assurance and Quality Control Samples

Most groundwater monitoring programs will incorporate the collection and analysis of quality assurance and quality control (QA/QC) samples to assess the precision, accuracy, representativeness, completeness and comparability (PARCC) of the data. The type and frequency of QA/QC data collection necessary for any groundwater monitoring program depends upon the DQOs and the checks that need to be performed to verify that the data collected achieve the DQOs. QA/QC samples can generally be categorized as blanks, spikes, calibration checks, and duplicate or replicate samples.

Blanks are samples that are intended to contain none of the analytes of interest and are therefore analyzed to assess bias that may be introduced to the field samples during field or laboratory processes.

Spikes are samples in which a known quantity of a substance is added to the sample and the sample is analyzed for that substance to assess the accuracy of the analytical method.

TABLE 35.11 Project Quality Control Checks

QC check	Information provided
Blanks	
bottle blank	cleanliness of sample bottles
field blank	transport, storage, and field
reagent blank	handling bias
rinsate or equipment blank	contaminated reagent
method blank	contaminated equipment response of an entire analytical system
Spikes	
matrix spike	analytical (preparation + analysis) bias
matrix spike replicate	analytical bias
analysis matrix	instrument bias
surrogate spike	analytical bias
Calibration check samples	
zero check	calibration drift and memory effect
span check	calibration drift and memory effect
mid-range check	calibration drift and memory effect
Replicate, splits, etc.	
field collocated samples	sampling + measurement precision
field replicates	precision of all steps after acquisition
field splits	shipping + interlaboratory precision
laboratory splits	interlaboratory precision
laboratory replicates	analytical precision
analysis replicates	instrument precision

Source: USEPA. 2002. *EPA Guidance for the Quality Assurance Project Plans — EPA QA/G-5* EPA/240/R-02/009. Washington, DC, 111 p. December.

Calibration Checks are standards containing a known concentration of the subject analyte that are analyzed typically at the beginning, middle, and end of the analyses of the batch of field samples. Calibration checks are performed to verify that the analytical equipment is producing accurate results within the anticipated range of the analytical method.

Replicates or duplicates are samples that are obtained at the same time from the same population of a particular media. The duplicate samples are analyzed and compared to assess the precision and reproducibility of the analytical method.

There are several specific QA/QC samples that fit into the categories described above that are intended to assess specific aspects of the sampling program. Table 35.11 provides a summary of the specific QA/QC samples that are commonly used in groundwater monitoring programs and the information that they provide. More detailed discussions of QA/QC sampling is provided by USEPA (1992a, 1998).

35.4.5 Sample Handling and Preservation

After the well has been sampled, the sample must be properly transferred to the sample container and preserved for transport to the laboratory. During sampling, every effort must be made to minimize changes in the chemistry of the sample. To minimize such changes in chemistry, samples should be collected, preserved, and stored correctly. Guidelines for each of these activities are presented by USEPA (1992a) and Nielsen (1991) and are summarized below. A generalized groundwater sampling protocol is presented in Table 35.12.

Sample collection refers to the transfer of the sample from the sampling device to the sample container. The contents of the sampling device should be transferred in a controlled manner that minimizes sample agitation and aeration, which can cause gasification of samples, allow release of volatile organics, or cause oxidation and precipitation of metals in the sample. Groundwater samples should be collected soon after the well is purged to prevent interaction of the samples with the atmosphere and well casing. In samples that will be tested for volatile organic constituents, there should be no air in the sample containers

TABLE 35.12 Generalized Groundwater Sampling Protocol Steps

	Goal	Recommendations
1. Hydrologic measurements	Establish nonpumping water level. Remove or isolate stagnant H ₂ O that otherwise bias representative sample. Collect samples at land surface or in well-bore with minimal disturbance of sample chemistry. Filtration permits evaluation of characteristics of soluble constituents. It should be performed in the field as soon as possible after sample collection.	Measure the water level to ± 1 cm (± 0.01 ft). Pump water until well purging parameters (e.g., pH, temperature, specific conductance, Eh) stabilize to ± 10 percent over at least two successive well volumes pumped. Pumping rates should be limited to ~ 100 mL/min for volatile organics and gas-sensitive parameters. Filter trace metals, inorganic anions/cations, alkalinity. Do not filter samples to be analyzed for TOC, TOX, volatile organic compounds, other organic compounds.
2. Well purging		
3. Sample collection		
4. Filtration		
5. Preservation	Retard chemical changes that could affect the chemistry of a sample after it is extracted from the well.	See Table 35.13.
6. Field determinations	Field analyses of samples will effectively avoid bias in determining	Parameters or constituents that are not preservable should be analyzed in the field if at all possible. At least one blank and one standard for each sensitive parameter should be made up in the field on each day of sampling. Spiked samples are also recommended for complete QA/QC. Observe maximum sample holding or storage periods recommended by the agency. Actual holding periods should be carefully documented.
7. Field blanks/standards	parameters/constituents that can change significantly after sample collection (e.g., temperature, alkalinity, pH). These blanks and standards will permit the correction of analytical results for changes that may occur after sample collection, preservation, storage, and transport.	
8. Sample storage/transport	Refrigerate and protect samples to minimize their chemical alteration prior to analysis.	

Source: Data from USEPA. 1990. *Handbook — Groundwater, Volume I: Groundwater and Contamination*. EPA/625/6-90/016a. Washington, D.C.; USEPA. 1992a. *RCRA Groundwater Monitoring: Draft Technical Guidance*, Office of Solid Waste, EPA/530-4-93-001, PB93-139-350, Washington, DC.

(which would allow volatilization of the compounds during shipment). Also, the samples should be transferred from the sampling device directly into the sampling container to minimize the opportunity for contamination of samples or changes in sample chemistry.

Filtration of samples is frequently performed to provide samples that do not contain suspended solids. Filtration is usually performed by draining or forcing samples through a filter having openings of 0.45 μm . The advantage of filtering samples is that the constituents adsorbed to the suspended solids can be distinguished from constituents that travel in dissolved phase in the groundwater. Potential disadvantages of filtering include changes in groundwater chemistry during filtration and the removal of some solid particles (e.g., colloids) that may actually travel with groundwater. More detailed discussions of the advantages and disadvantages of filtering groundwater are provided by Puls and Barcelona (1989) and USEPA (1995). In monitoring scenarios where significant quantities of suspended solids exist, it may be useful to analyze both filtered and unfiltered samples to allow a comprehensive evaluation of groundwater quality.

Sample preservation refers to actions taken to minimize changes to the chemistry of the sample after it is removed from the well. As there are significant changes to the environment of the sample after it is removed from the well (e.g., temperature, light, presence of air, etc.), the sample can experience geochemical changes that could render the sample unrepresentative if it is not properly preserved. Sample preservation methods are intended to retard volatile loss and chemical reactions such as oxidation, biodegradation, and sorption. Preservation methods are generally limited to pH adjustment, refrigeration, and protection from light. A list of appropriate preservation measures is presented in Table 35.13.

TABLE 35.13 Sample Preservation Measures

Parameters (type)	Volume required (ml) 1 sample ^a	Container (material)	Preservation and storage requirements	Maximum holding period
Well purging				
pH (grab)	50	T,SS,P,G	None; field det.	<1 h ^b
specific conductivity (grab)	100	T,SS,P,G	None; field det.	<1 h ^b
T (grab)	1000	T,SS,P,G	None; field det.	None
Eh (grab)	1000	T,SS,P,G	None; field det.	None
Contamination indicators				
pH, specific conductivity, (grab)	As above	As above	As above	As above
TOC	40	G,T	Dark, 4°C	35 h
TOX	500	G,T	Dark, 4°C	5 days
Water quality				
dissolved gases (O ₂ , CH ₄ , CO ₂)	10 ml minimum	G,S	Dark, 4°C	<35 h
Alkalinity/Acidity	100 (filtered)	T,G,P	4°C/None	<6 h ^b / 35 h
Fe, Mn, Na ⁺ , K ⁺ , Ca ⁺⁺ , Mg ⁺⁺	All filtered, 1000 ml	T,P	Field acidified to pH < 2 with HNO ₃	6 months ^Δ
PO ₄ , Cl, Silicate	@50	(T,P,G glass only)	4°C	35 h/7 days; 7 days
NO ₃ ⁻	100	T,P,G	4°C	35 h
SO ₄ ²⁻	50	T,P,G	4°C	7 days
NH ₄ ⁺	400	T,P,G	4°C/H ₂ SO ₄ to pH < 2	35 h/7 days
Phenols	500	T,G	4°C/H ₃ PO ₄ to pH < 4	35 h
Drinking water suitability				
As, Ba, Cd, Cr, Pb, Hg, Se, Ag	Same as above for water quality cations (Fe, Mn, etc.)	Same as above for water quality cations (Fe, Mn, etc.)	Same as above for water quality cations (Fe, Mn, etc.)	6 months
F ⁻	Same as chloride above	Same as chloride above	Same as chloride above	7 days
Remaining organic parameters				
As for TOX/TOC, except where analytical method calls for acidification of sample				35 h

Modified after Scalf, M.R., McNabb, Dunlap, W.J., Crosby, R.L., and Fryberger, J. 1981. *Manual of Groundwater Quality Sampling Procedures*. EPA/600/2-81/160 (NTISPB82-103045).

^aIt is assumed that at each site, for each sampling date, replicates, a field blank, and standards must be taken at equal volume to those of the samples.

^bTemperature correction must be made for reliable reporting. Variations greater than ± 10 percent may result from longer holding period.

^ΔIn the event that HNO₃ cannot be used because of shipping restrictions, the sample should be refrigerated to 4°C, shipped immediately, and acidified on receipt at the laboratory. Container should be rinsed with 1:1 HNO₃ and included with sample.

Note: T = Teflon, SS = stainless steel, P = PVC, polypropylene, polyethylene. G = borosilicate glass.

From USEPA. 1993a. Office of Research and Development. *Subsurface Characterization and Monitoring Techniques: A Desk Reference Guide, Volume I: Solids and Groundwater, Appendices A and B*. U.S. Environmental Protection Agency, Washington, D.C. EPA/625/R-93-003a.

Sample storage refers to measures used to maintain sample quality during transportation. Samples should be cooled to a temperature of 4°C as soon as possible after they are collected and should be maintained at that temperature until they are received at the laboratory. Samples should be shipped in containers that minimize agitation of the samples and should be accompanied by proper documentation

(see Section 35.4.6). Note that most samples must be analyzed within a specified period (i.e., holding time) after the sample is collected; data from samples that are analyzed after the holding time limit has been exceeded are considered to be unreliable. Coordination between the field sampling team and the laboratory can help to prevent exceedance of holding times. Sample storage requirements and maximum recommended holding times are listed in Table 35.13.

35.4.6 Documentation of Sampling Events

Each sampling event should be documented so that the validity of the sample collection, preservation, and storage techniques can be verified in the future. If this information cannot be verified, then (depending on the data quality objectives specified for the project) the data could be rendered invalid and resampling could be required. An example field sampling documentation log form is presented in Table 35.14. As shown in the table, each aspect of the sampling event should be recorded, including the time of sampling, ambient weather conditions, time required to purge the well and the volume of water purged, purge water characteristics, decontamination procedures, sample equipment calibration, and procedures for preservation of samples. Also, chain-of-custody of documentation should be prepared. The record of sampling should be stored in the project files for future reference.

35.5 Analysis of Groundwater Samples

35.5.1 Field Chemical Analyses

Many analyses can be performed in the field using electronic sensors or other field methods such as reagent test kits or bioassay tests; in fact, several parameters must be measured in the field to be relevant. Analyses that should be performed in the field include measurement of temperature, pH, specific conductance, redox potential, dissolved oxygen, sulfides, nitrate, and ferrous iron. Temperature measurements that are not made in the field immediately after the sample is collected are not representative of *in situ* conditions because the temperature of the sample may change rapidly after the sample is removed from the well. Measurements of pH and redox potential are strongly affected by temperature and, therefore, should also be measured in the field. Other measurements such as sulfide, dissolved oxygen, nitrates, and ferrous iron may be significantly affected by exposure to the atmosphere. These analyses can be performed quickly using readily available, relatively inexpensive equipment. A good discussion of field measurement techniques is presented by USEPA (1993a).

Many of the analyses that are typically performed in the laboratory (which are described in Section 35.5.2) can be performed in the field by setting up a field laboratory. Field laboratories have become increasingly common as manufacturers of analytical equipment have developed increasingly reliable, portable analytical devices. The advantages of performing analyses in the field include fast return of results, decreased chance of sample disturbance or changes in sample chemistry during shipment, ability to obtain an additional sample quickly if problems occur during analysis, and lower cost per sample (if large quantities of analyses are performed). The disadvantages of performing analytical tests in the field include difficulty in implementing quality assurance measures, increased potential for contamination of samples and equipment, less sophisticated instrumentation (which usually results in higher method detection limits and lower precision and accuracy), and correspondingly less reliable test results. However, in spite of these shortcomings, field analysis can be an extremely valuable and time-saving approach for site characterization studies and for monitoring of remediation activities.

35.5.2 Laboratory Chemical Analyses

Most analyses of groundwater samples are performed in an offsite, permanent laboratory to provide a high level of quality control of analyses. Laboratory analyses may be conducted using a variety of analytical techniques and methods. Analytical “technique” refers to a particular procedure and type of instrument

TABLE 35.14 Example of a Field Sampling Log (Source: USEPA, 1993)

GROUND WATER SAMPLING INFORMATION FORM*			
Sheet _____ of _____ Side 1 of 2			
General Information			
Location (Site/Facility Name)	Sampling Point (common name)		
Project Name/#	Type (mon. well, spring, etc.)		
Field Personnel	Field Sample (Event) ID#*		
Sampling Organization	Facility ID (for IGWIS data entry)		
Weather	Station ID (for IGWIS data entry)		
Sampling Station (Well) Details			
Read from left to right: <input checked="" type="checkbox"/> <input type="checkbox"/>	top - bottom		
Well Depth (ft. below MP)	Casing Diameter (inches)	Open Interval (depth below GS) <input checked="" type="checkbox"/> <input type="checkbox"/>	
Static Depth to Water (below MP) <small>(0.01 ft)</small>	Static DTW (ft below GS) <small>(0.1 in)</small>	Date	Time
Water Column Length (L) (ft)	One WC Volume (cu ft)	One WC Volume (gals)	
Condition: Securely Locked?	Station (Well) Damaged? <input checked="" type="checkbox"/> <input type="checkbox"/>	Surface Contamination (visible) <input checked="" type="checkbox"/> <input type="checkbox"/>	
Purging			
PID/FID Reading @ Wellhead*	Concentration ppm	Background Conc. ppm	
Free Product <small>LAPL, LNAPL, DNAPL</small>	Detected/Sampled? <input checked="" type="checkbox"/> <input type="checkbox"/>	Appearance	
Well Purging Equipment	Pump, bailer?	Type*	
Purging Date/Time	Start <input checked="" type="checkbox"/> /	Finish <input checked="" type="checkbox"/> /	
Pump/Bailer Intake Set at	Feet below MP	Avg. Purge Rate ppm	
Amt. Purged before Sampling	Gals/WC Volumes	Purge Protocol of WCV's met? <input checked="" type="checkbox"/> <input type="checkbox"/>	
Field Water-Quality Measurements and Observations			
Date/Time Measurements Began	/	Purge Rate for Measurements (gpm)	
Submersible Pump with direct line to Flow Cell used for all Field Water Quality Measurements?	<input checked="" type="checkbox"/> <input type="checkbox"/>		
All Field Measurement Instruments Calibrated according to Protocol?	<input checked="" type="checkbox"/> <input type="checkbox"/>		
All Field Water Quality Parameters Stabilized according to Protocol Criteria just before filling sample containers?	<input checked="" type="checkbox"/> <input type="checkbox"/>		
The Measurements below Represent: (1) stabilization, (2) sample water collected, (3) both a and 2, (4) other:			
Sample Appearance: <input checked="" type="checkbox"/> <input type="checkbox"/>	Odor:		
Field Measurement	Value	Military Time	Comments*
Temperature	°C		
Electrical Conductivity	µMhos/cm		= meter reading x magnitude x k
Specific Conductance	µMhos/cm		EC corrected to 25 °C
pH	Standard Units		
Dissolved Oxygen	mg/l		
Eh	mV		
Turbidity	NTU		
Sample Collection			
Sampling Device (type of pump/bailer)*	Sample Medium (well water, LNAPL, etc.)*		
Permanently Installed Pump? <input checked="" type="checkbox"/> <input type="checkbox"/>	Dedicated Equipment? <input checked="" type="checkbox"/> <input type="checkbox"/>	Used Same Equip. for Purge? <input checked="" type="checkbox"/> <input type="checkbox"/>	<input checked="" type="checkbox"/> <input type="checkbox"/>
Pump Intake/Bailer Set at (ft below MP)		Interval Samples Represent (ft below GS) <input checked="" type="checkbox"/> <input type="checkbox"/>	/ Bottom
Date / Time Sampling Began		Date / Time Sampling Finished	
Depth to Water (ft below MP)		Depth Water (ft below MP)	
QC Samples Collected? <input checked="" type="checkbox"/> <input type="checkbox"/> (see reverse)		Sample Withdrawal Rate	
All Field Protocols were followed with no exceptions (Y,N)		Enter Protocol Codes* 1. <input type="checkbox"/> 2. <input type="checkbox"/>	
Remarks (1)* (include protocol exceptions)			
Form Completed by .	<input type="checkbox"/>	<input type="checkbox"/>	Date _____
60 _____ *See 2 of this form contains definition of abbreviations, protocol codes, additional room for equipment specification, QC sample description and other comments			
Form GWS #2 Revised 9-2-93			

that is used in the laboratory to analyze the sample; an analytical technique may incorporate one or more analytical methods to analyze a sample. An analytical method refers to the requirements and protocols that are employed when using a particular analytical technique to evaluate a specific group or subset of constituents. Note that appropriate methods and techniques will be a function of required detection limits, regulatory requirements, and DQOs. Laboratories will often recommend an appropriate technique and method depending upon your site-specific conditions and data needs.

The most commonly used laboratory analytical techniques include chromatography (i.e., detection of constituents based on the rate that they migrate through a stationary medium), spectroscopy (i.e., identification of constituents based on the changes in light spectrum caused by irradiated light as the light passes through the groundwater sample), and photometry (i.e., the measurement of the intensity of light or the relative intensity of different lights as they pass through a groundwater sample) (USEPA, 1993b). Some of the more commonly used laboratory analytical techniques are described below (USEPA, 1993b).

Gaseous-phase analyses are performed using instruments designed to detect constituents in gases or that require conversion of the sample to the gaseous phase before analysis. Gaseous-phase techniques include photo-ionization, flame-ionization, explosimetry, gas chromatography, mass spectrometry, and atomic adsorption spectrometry.

Luminescence/spectroscopic analyses involve exciting a sample (e.g., x-rays or other means of radiation), and measuring the radiation emitted by the compound, either during excitation or as the electrons in the constituent return to their original state. Luminescence/spectroscopic techniques include x-ray fluorescence (i.e., measurement of the secondary radiation emitted when a sample is excited using x-rays), fluorometry (i.e., measurement of the radiation emitted when electrons in a molecule return to their original state after excitation), or spectrometry (i.e., measurement of the absorption spectra of narrow bandwidths of radiation from the sample).

Wet chemistry analyses include a wide range of colorimetric (e.g., titration, colorimetry, filter photometers, and spectrophotometers), immunochemical (e.g., enzyme immunoassay and fluoroimmunoassay), liquid chromatography, and electrochemical (e.g., voltammetry, polarography, pH, Eh, dissolved oxygen, and electrical conductance) techniques. These procedures are relatively straightforward and inexpensive but they require strict application of QA/QC procedures, are time consuming, require different reagents for each analyte of concern, and have limited application for some toxic chemicals.

Other analytical techniques include radiological, gravimetric, magnetic, microscopic, biological, and chemical sensor techniques. Each of these applies to particular constituents. The potential advantages of these techniques include: (i) economy of scale when analyzing for a small list of constituents from a large number of samples; (ii) better reliability of results for certain compounds; and (iii) (in some cases) simple operation of equipment. Disadvantages could include expensive equipment, difficult QA/QC procedures, and high cost of analyses. Refer to USEPA (1993b) for a discussion of these and other candidate analytical techniques.

The specific analytical method used in a groundwater monitoring program is often specified by the regulations. For example, most groundwater monitoring data collected under RCRA or other federal or state cleanup programs require that analyses be performed in accordance with USEPA methods for evaluating solid waste contained in USEPA (1996). These analytical methods specify the analytical technique, requirements, and protocols that need to be implemented to perform the required analysis and requisite quality control for a specific group of compounds or elements. Other standard analytical methods may also be applicable, for example, USEPA drinking water test methods for projects where groundwater from residential or public water supply wells is evaluated. At times, specialized methods must be developed to detect a particular contaminant for which a standard methodology does not exist or to increase the precision, accuracy, and sensitivity of a standard method to achieve DQOs. If such specialized methods are required, they should be developed by an experienced analytical chemist.

When reporting analytical results, the laboratory may present a qualifier with the data to clarify the validity of the test result. Qualifier types and definitions vary between laboratories, but some qualifiers are used widely by most laboratories. Typical qualifiers are presented below (USEPA, 1992a); the laboratory

should be required to submit a list of qualifier definitions with the reports of laboratory analyses.

J = value is estimated (typically below the method detection limit).

U = compound analyzed for but undetected (value presented is the quantitation limit).

R = data does not meet laboratory QC requirements and is rejected.

N = compound is tentatively identified (i.e., its presence is likely, based on the judgment of the laboratory, but not conclusively confirmed by the test result).

35.6 Evaluation of Groundwater Monitoring Data

35.6.1 Data Validation

Data validation refers to the process of verifying that the groundwater sampling and analysis was performed in accordance with the requirements of the groundwater monitoring plan and in accordance with the requirements of applicable QA/QC procedures. Data that does not comply with these specific requirements must be considered to be potentially invalid. Example causes of invalid data include insufficient purging affecting representativeness of the sample; improper sampling, filtration, and preservation techniques; failure to analyze the samples before the recommended holding time is exceeded; the presence of air in VOC sample containers; failure to properly implement chain-of-custody procedures; improper calibration of laboratory analytical equipment; presence of contaminants in field blank, method blank, or trip blank samples; or failure of the analytical laboratory to properly implement or document QA/QC procedures. If data has been found to be potentially invalid, then one of the following may be required: (i) if remaining portions of the samples exist, then they may be analyzed and substituted for the invalid data; (ii) additional samples can be collected and analyzed; or (iii) documentation of the sampling event can be reviewed and, if it is found that the improper procedure would not have impacted the analytical results, notations can be made in the project files regarding the occurrence and resolution of the problem, with the analytical result tentatively used in data evaluations.

35.6.2 Evaluations of Groundwater Elevation Measurements or Depths to Groundwater

The elevation of the groundwater potentiometric surface is typically evaluated by measuring the depth to groundwater at several locations, calculating the elevation of groundwater at each location, and then interpolating contours of groundwater elevation for the area. The groundwater elevation can be calculated for the well location by subtracting the depth to groundwater from the surveyed elevation of the reference measuring point on the monitoring well casing.

Measurements of the depth to groundwater can be complicated by the presence of LNAPLs in the monitoring well or piezometer, as shown in Figure 35.9 and as discussed by Blake and Hall (1984). As shown in the figure, the thickness of LNAPLs (e.g., petroleum hydrocarbon products) in an aquifer can be significantly different from the thickness of LNAPL in a well. LNAPLs typically “float” on groundwater at the top of the capillary fringe. LNAPLs flow into the well until the elevation of the top of LNAPL above the capillary fringe equals the elevation of the top of LNAPL inside the well. As groundwater in the capillary fringe does not flow into the well, the thickness of LNAPL present in the well (T_A) is typically greater than the actual thickness of LNAPL in the aquifer ($H_M + T_M - H_C$). The actual thickness of LNAPL, and a more accurate estimate of depth to groundwater, can be estimated from:

$$T_M = T_A - (D_{WT} + H_M) \quad (35.1)$$

where T_M = thickness of mobile LNAPL (ft), T_A = apparent thickness of LNAPL (ft), H_M = distance from bottom of the mobile LNAPL layer to the water table, and D_{WT} = depth of LNAPL in the well below

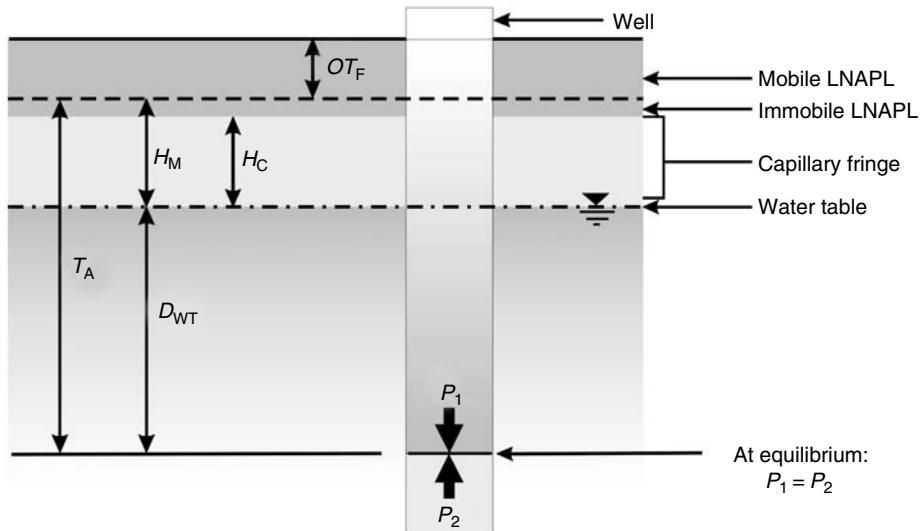


FIGURE 35.9 Distribution of LNAPL and groundwater in a well within an LNAPL-contaminated aquifer.

the actual water table (ft). D_{WT} can be calculated as

$$D_{WT} = H_M \gamma_{LNAPL} / (\gamma_w - \gamma_{LNAPL}) \quad (35.2)$$

where γ_{LNAPL} = the specific weight of LNAPL (lb/ft^3), and γ_w = the specific weight of water (lb/ft^3). H_M is approximately equal to the thickness of the capillary fringe, which can be estimated based on the grain-size distribution of the soil (see Lambe and Whitman 1969).

Once the elevation of groundwater has been calculated for the wells at a site, the configuration of the potentiometric surface can be estimated to provide an indication of the likely direction of groundwater flow and the gradient of the potentiometric surface. There are numerous techniques for interpreting the contours of the potentiometric surface, the most common and basic of which is manual contouring based on the triangular linear interpolation of hydraulic head data from a minimum of three monitoring wells. Often, more complex computer-generated contouring programs are available such as kriging, triangulation with linear interpolation, inverse distance weighted averaging, or minimum curvature. Representations of the potentiometric surface are typically generated using specialized software such as CADD or geographic information systems equipped with geostatistical packages or specialized groundwater modeling software. Additional information regarding potentiometric surface contouring is provided by Kresic (1997).

An illustration of a potentiometric surface map, as interpreted from groundwater elevations at several wells, is provided in Figure 35.10. As shown in the figure, a potentiometric surface can be produced by interpolating data between monitoring wells; the groundwater professional should carefully consider all available site data when evaluating groundwater flow direction to prevent the type of problem illustrated in Figure 35.10, which shows that a limited subset of the data would imply an incorrect groundwater flow direction. Surface-water features that are connected with groundwater can also provide valuable information regarding the configuration of the potentiometric surface. The gradient of the potentiometric surface is calculated as the ratio of vertical change of elevation in the potentiometric surface to the distance over which the change occurred.

The velocity of groundwater can then be calculated for a given area as (Freeze and Cherry 1979):

$$v = - \frac{Ki}{n_e} \quad (35.3)$$

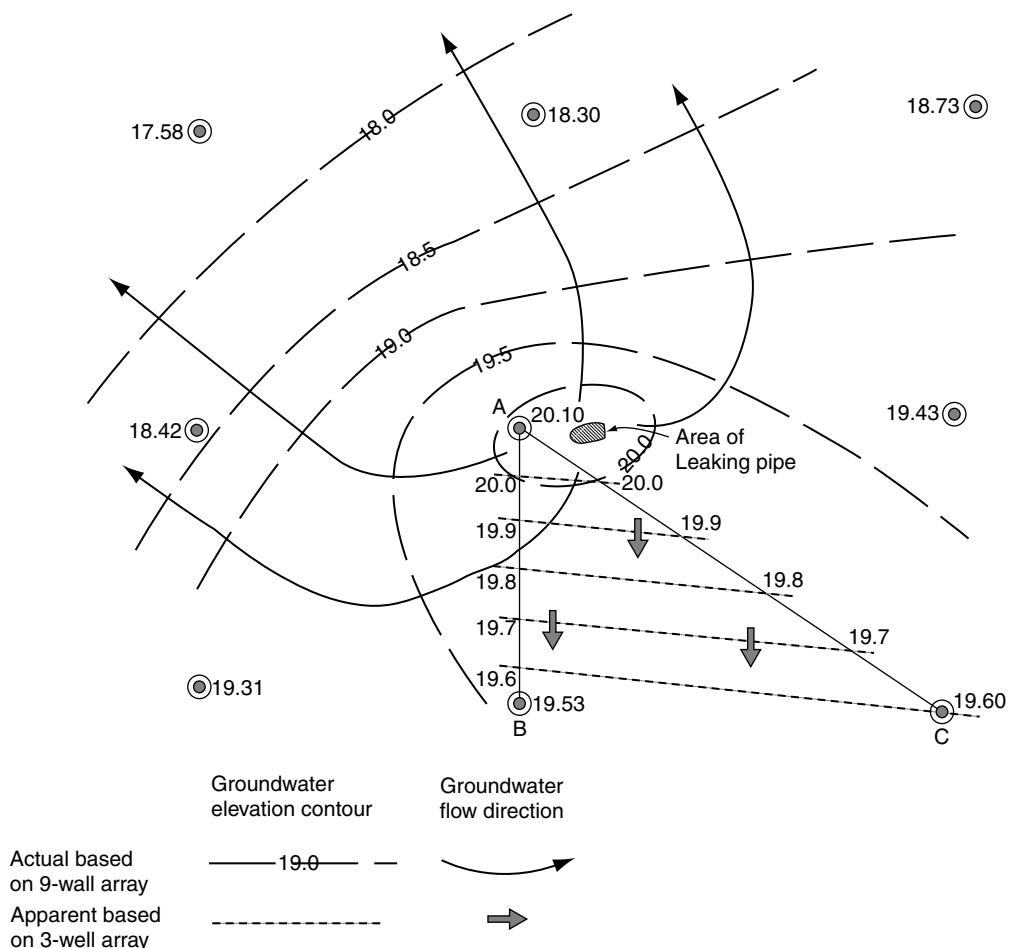


FIGURE 35.10 Potentiometric surface mapping. (From Nielsen, D.M., ed. 1991. *Practical Handbook of Ground-Water Monitoring*. Lewis Publishers, Chelsea, MI. With permission.)

where v = groundwater flow velocity [ft/sec], K = hydraulic conductivity [ft/sec], i = groundwater hydraulic gradient (which is usually a negative number) [—], and n_e = effective porosity of the aquifer material [—]. This velocity can be used as an initial estimate of the speed at which a nonreactive solute might move through the aquifer.

35.6.3 Groundwater Quality Data Analysis

35.6.3.1 Comparisons to Standards

Some simple, yet common, type groundwater quality data evaluation are comparisons to published values such as ranges of naturally occurring groundwater quality, regulatory standards, or site-specific criteria established in permits or the groundwater monitoring plans. Typical ranges of physical, chemical, and biological characteristics of uncontaminated groundwater are presented in Table 35.15. It should be noted that the ranges provided in Table 35.15 are generalities and that natural groundwater quality varies widely depending upon geologic conditions and other environmental and geographic factors. More specific information regarding natural groundwater quality for a particular geographic location and geologic formation can often be obtained from state or federal geologic survey publications or from other local organizations. The most common regulatory standards by which groundwater data are compared are

TABLE 35.15 Key Physical, Chemical, and Biological Properties of Groundwater

Category	Property	Standard range of values in natural groundwater
Physical	Temperature	10–20°C
	Elevation of potentiometric surface, or depth below land surface	Varies widely
	Presence of NAPLs	None
	Total suspended solids	100–500 ppm
	Total dissolved solids	100–1000 ppm
Chemical	pH	6.5–8.5 standard units
	Dissolved oxygen	2–5 ppm
	Total organic carbon	1–10 ppm
	Specific conductance	100–1000 $\mu\text{mhos}/\text{cm}$
	Manganese	0–0.1 ppm
	Iron	0.01–10 ppm
	Ammonium	0–2 ppm
	Chloride	2–200 ppm
	Sodium	1–100 ppm
	Volatile organics	<40 ppb
	Sodium, calcium, bicarbonate, magnesium, arsenic, cadmium, iron, lead, nickel, selenium, zinc	1000 to 1,000,000 ppb
	Beryllium, Mercury, Silver, Thallium	1 to 1,000 ppb
Biological	Coliforms, viruses	0 organisms

Source: Data from Driscoll, F.G. 1986. *Groundwater and Wells*. Johnson Division, St. Paul, MN.; Freeze, R.A. and Cherry, J.A. 1979. *Groundwater*. Prentice-Hall, Inc., Englewood Cliffs, NJ.; Hem, J.D. 1989. Study and Interpretation of the Chemical Characteristics of Natural Water. U.S. Geological Survey Water Supply Paper 2254, 3rd Edition.

drinking water standards established by USEPA, which include maximum contaminant levels (MCLs) and secondary maximum contaminant levels (SMCLs). Many states have also developed numeric standards for groundwater cleanup sites based on contaminant toxicity and risk. Although a comparison to a particular value or standard is a simple method, a statistically valid comparative approach is often required to confirm the validity of the findings, as described below.

35.6.3.2 Statistical Evaluations of Data

Another common approach used to evaluate groundwater quality data is to statistically assess the monitoring data to identify spatial or temporal variations in the data. After data have been validated, they should be evaluated to determine if groundwater quality conditions have changed since the last sampling event. As a first step in this process, the presence of the contaminants should be confirmed and an evaluation of the contribution of the contaminants by the potential source should be performed. This first step is typically made by analyzing the data using statistical analyses, which can be used as the basis for concluding whether a contaminant is present in groundwater. An adequate quantity of data is required before a statistical analysis can be performed; often, samples from multiple background wells and from multiple sampling events must be available. A discussion of statistical analysis of groundwater analytical data is presented in USEPA (1989b).

Statistical analyses are based on the probability theory that the likely range of all values (e.g., constituent concentrations at any location within the monitored area) can be estimated based on the distribution of the known values. If a groundwater quality dataset is found to be consistent with a standard distribution (e.g., normal, lognormal, etc.), then the probability that the concentration will exceed a certain value (e.g., a groundwater protection standard) can be estimated. Two key parameters that are used in nearly all statistical analyses are the mean and the standard deviation of the data set:

$$\bar{x} = \left(\sum_{i=1}^n x_i \right) / n \quad (35.4)$$

$$s = \left[\sum_{i=1}^n (x_i - \bar{x})^2 / (n - 1) \right]^{1/2} \quad (35.5)$$

where x = sample mean; i = observation number; x_i = the value for the i th observation; n = total number of observations; and s = sample standard deviation. Groundwater quality data are usually well characterized by a normal or lognormal distribution, although some data cannot be adequately characterized by any common distribution (USEPA, 1989b). As most statistical tests are based on an assumption regarding normality or lognormality, the data must be tested for conformance to either a normal distribution or other type of distribution that is assumed by the particular statistical test. One test for evaluating the fit of a data set to a normal distribution is the coefficient of variation test. To perform this test, the sample mean and standard deviation are first calculated. Then the sample standard deviation is divided by the sample mean; if the result is greater than one, then the data do not fit a normal distribution well. If the data are found not to fit a normal distribution, then a log transformation should be performed on the original data and the coefficient of variance test rerun to see if the data fit a lognormal distribution. If the data fit either a normal or a lognormal distribution, then a parametric approach (described in Section 35.6.3.2) to statistical analyses can be used; otherwise, a nonparametric analysis (also described in Section 35.6.3.2) must be used. Other more accurate but more complicated tests are available to test the fit of the data to a normal or lognormal distribution (e.g., probability plots, coefficient of skewness, the Shapiro-Wilk test, and the probability plot correlation coefficient, all described in USEPA (1989b)).

35.6.3.3 Tests for Statistical Significance of Detected Constituents

To select a statistical approach, the type of comparison being made (e.g., compliance well vs. background, introwell, etc.) must be selected, and then a method that best suits the approach. Types of comparisons and recommended statistical methods are presented in Table 35.16. As shown in the table, the recommended method of statistical analysis varies based on the type of comparison being made and the number of times the compound of concern was detected. If the constituent was not previously detected, then the analytical results should simply be compared to the quantitation limit of the test apparatus.

TABLE 35.16 Recommended Statistical Methods for Groundwater Monitoring

Compound	Type of comparison	Recommended method
Any compound in background	Background vs. compliance well Introwell	ANOVA tolerance intervals, prediction intervals, Control charts
ACL/MCL specific	Fixed standard	Confidence intervals, tolerance intervals
Many nondetects in database(synthetic)	% nondetects	Recommended method
	$0 \leq x \leq 15\%$	Replace nondetect values with one half the method detection limit or quantitation limit (as appropriate), then proceed with any of the following parametric procedures: ANOVA Tolerance intervals Prediction intervals Control charts. If residuals from transformed do not meet parametric ANOVA requirements, use nonparametric approaches.
	$15\% \leq x < 50\%$	Treat nondetect values as ties, then either proceed with nonparametric ANOVA, or use Cohen's adjustment, then proceed with: Tolerance intervals, Confidence intervals Control charts
	50–99%	Test of proportions
	100%	Compare sample best result values to sample quantitation limits

Key: ND = Nondetect QL = Quantitation Limit MDL = Method Detection Limit ACL = Alternate Concentration Limit From USEPA. 1989b. *Statistical Analysis of Groundwater Monitoring Data at RCRA Facilities, Interim Final Guidance*. Office of Solid Waste, Washington, DC.

One of the most powerful statistical tests for groundwater monitoring is the Analysis of Variances (ANOVA) test, also known as the “general linear model procedure.” This method is useful for comparing background data to compliance well data and is useful in situations where constituents are detected in many wells at differing concentrations. The method is used to compare the means of different groups of observations (i.e., from different wells on the same sampling date) and to identify significant differences among the groups; if there are significant differences, then the data can be further analyzed to identify contrasting aspects of the dataset. If the compound is not routinely detected in groundwater, then the nondetect data points are replaced with nonzero approximations of data before proceeding with the statistical test. The analysis can be performed using either a parametric or a nonparametric approach. The parametric ANOVA procedure assumes that the differences (called residuals) between the mean of the values and the values themselves are normally distributed with equal variance; this case can be checked as described in Section 35.6.3.1. If this test shows that the residuals do not meet this assumption, then the nonparametric ANOVA procedure must be performed. For a nonparametric ANOVA, it is first assumed that there is no contaminant present (i.e., all of the data comes from uncontaminated groundwater having the same continuous distribution of constituents) and that, therefore, the median concentration of contaminants must be the same at all wells. The nonparametric ANOVA is performed by comparing the median concentration of hazardous constituents for all wells and then, if the median differs significantly, by comparing the average rank of each of the compliance wells with the average rank of the background well.

35.6.4 What to Do if a Statistical Analysis Indicates the Presence of Contaminants

If the results of a statistical analysis indicate that contaminants are present in a particular groundwater monitoring well, then the data should first be reexamined to verify that the test results were not a false-positive indication of contamination. The data should be examined for the following problems that may cause a false-positive indication of groundwater contamination. See also Nielsen (1991), USEPA (1989b), and Miller and Miller (1986) for discussions of false-positive groundwater monitoring results.

Natural variation in groundwater quality: Groundwater quality sometimes varies naturally according to season, temperature, geology, or a number of other factors. If the statistically significant increase is due to a constituent that is present in uncontaminated groundwater (e.g., manganese, sodium, iron, etc.), then the historic data for the site should be reexamined to evaluate the expected range of values for the detected constituent. Constituents of concern may be naturally present in groundwater and may not be the result of the potential source; for example, arsenic and lead are naturally present at levels exceeding MCLs in uncontaminated groundwater in many parts of the United States. These false positives can be identified and resolved by obtaining adequate representative background data.

Contaminant source different from the potential source: The potential source being monitored may not be the actual source of contamination. This could be confirmed by either demonstrating that a hydraulic interconnection exists between an alternative source and the well having the statistically significant increase or by showing that the contaminants present in the well could not have been derived from the potential source, given the type of waste in the potential source. Another type of false-positive result is the contamination of groundwater samples by condensation of landfill gas in monitoring wells.

Well construction problems: The statistically significant increase may be due to a problem with well construction. For example, high pH or alkalinity values may indicate that grout from the well is “bleeding” down into the filter pack, either through the aquifer formation or through an inadequate bentonite seal. Alternatively, the statistically significant increase may be a result of poor well maintenance; an example of such a problem is infiltration of contaminated surface water ponded around the well at the ground surface.

Analytical error: The statistically significant increase may be the result of laboratory error. For example, methylene chloride and acetone (which are commonly used in the laboratory to clean test devices) are routinely detected in groundwater samples but are often a false-positive indication of contamination. Such analytical errors can usually be detected by preparing and analyzing laboratory, trip, and method

blanks. Contaminants present in laboratory or trip blanks that are also present in groundwater samples are not indicative of groundwater contamination. Good QA/QC procedures and documentation by the laboratory can usually lead to resolution of analytical errors.

Statistical error: These errors are typically mathematical errors or inappropriate application of statistical procedures (e.g., improper consideration of nondetects in sample datasets). If such a problem is expected, then the results of the statistical analysis should be reviewed by an expert in statistical analysis.

Sampling error: These errors include problems in the field that occur during sampling, such as improper filtering of samples, improper evacuation of wells before sampling, agitation of wells during sampling, agitation of samples during transfer to sample containers, and improper preservation of samples. This source of error is typically the most difficult to evaluate because, usually, very little documentation is generated during actual handling of the sample (because the sample technician's hands are occupied), making it difficult to isolate the source of the problem. The best method of preventing such problems is to use properly trained professional field technicians and to routinely implement good QA/QC procedures during sampling. The use of good field forms is an effective way of ensuring that key observations are made and recorded.

If no alternate source can be identified for the statistically significant increase, then the statistically significant increase should be attributed to the potential source being monitored and subsequent measures (such as those described in Section 35.2.2, Step 6) should be implemented.

Glossary

Anisotropy Variation in geologic formations in the vertical or horizontal direction.

Anthropogenic Resulting from, or caused by, human activities.

Aquitard A lithologic unit that impedes, but does not completely prevent, groundwater movement.

Background Sampling Evaluation of initial water quality conditions.

Bridging The development of gaps caused by obstructions in either grout or filter pack materials.

Also refers to blockage of particles in natural formation materials or artificial filter pack materials that may occur during well development.

Casing An impervious, durable pipe placed in a borehole to prevent the walls of the borehole from caving and to prevent flow of undesirable groundwater, surface water, gas, or other fluids into the well.

Chain of Custody A method for documenting the possession history of a sample from the time of its collection through its final disposition.

Confining Unit A relatively low-permeability material stratigraphically adjacent to one or more aquifers.

Contaminant A substance that is not normally present in groundwater or that is present at an unusually high concentration.

Cross-Contamination The movement of contaminants between aquifers or water-bearing zones through an unsealed or improperly sealed borehole or other conduit.

Detection Limit The lowest concentration of a chemical that can be reliably detected by an analytical device.

Development The act of removing materials introduced during drilling from a well and adjacent aquifer formation.

DNAPL A dense nonaqueous phase liquid that is relatively immiscible and has a density greater than that of water. Also known as "free product" or a "sinker."

Downdgradient In the direction of decreasing hydrostatic head.

Drilling Fluid A fluid (liquid or gas) that is used in drilling operations to remove cuttings from the borehole, to clean and cool the drill bit, and to maintain the integrity of the borehole during drilling.

Equipotential Surface A surface, in a three-dimensional groundwater flow system, for which the total hydraulic head is the same at every point on the surface.

False Negative A condition for which contamination is present, but the results of sample analyses fail to indicate its presence.

False Positive A condition for which no contamination is present, but the results of sample analyses indicate presence of contamination.

Field Blank A laboratory-prepared sample of known properties that is transported to the sampling site for use in validating field sampling procedures.

Filter Pack A clean silica-sand or sand and gravel mixture that is installed in the annular space between the borehole wall and the wellscreen for the purpose of retaining and stabilizing the particles from the adjacent strata.

Groundwater Protection Standard The acceptable quantity of contaminants in groundwater; the groundwater protection standard could be an MCL or another value.

LNAPL A light nonaqueous-phase liquid. Also known as “free product” or a “floater.”

Maximum Contaminant Level (MCL) The concentration at which the excess cancer risk of exposure to humans is estimated to be 1×10^{-6} , as defined by the USEPA.

Nonaqueous Fluids that are relatively insoluble.

Parts Per Billion (ppb), or Parts Per Million (ppm) Unit weight of solute per billion (or per million, respectively) unit weights of solution (solute, plus solvent).

Piezometer A well that is used only to measure the elevation of the water table or the potentiometric surface.

Potentiometric Surface A surface that represents the level to which water will rise in a tightly cased well. The “water table” is the potentiometric surface for an unconfined aquifer.

Preservation Measures taken during sample storage to minimize the change in concentration of a constituent of interest until analyses can be performed (e.g., storage of sample in acidic solution).

Quantitation Limit The lowest concentration at which a chemical can be accurately and reproducibly quantitated. Usually equal to the instrument detection limit times a factor of three to five.

Quality Assurance A management function used to establish and monitor quality control protocols and to evaluate their outcomes.

Quality Control Technical and operational procedures that are used to investigate and confirm the proper conduct of field, sample transportation, and laboratory activities necessary to assure accuracy and precision in the data.

Statistically Significant Exceedance of a certain level of probability, based on the results of a statistical analysis.

Target Monitoring Zone The portion of an aquifer for which there is a reasonable likelihood that a vertically placed well will intercept migrating contaminants.

Upgradient In the direction of increasing hydrostatic head.

Volatile Organic Compounds (VOCs) Compounds that will partition relatively easily into the air phase, sometimes defined as compounds with Henry's law constant greater than around 10^{-6} atm-m³/mol.

Wellscreen A filtering device used to retain the filter pack and aquifer materials while allowing groundwater to enter the well, commonly constructed of slotted or perforated casing material.

References

American Society for Testing and Materials (ASTM) 1998. Standard D 6286-98 *Standard Guide for Selection of Drilling Methods for Environmental Site Characterization* 16 p., West Conshohocken, PA.

American Society for Testing and Materials Standard D5092-90. 1990. *Standard Practice for Design and Installation of Groundwater Monitoring Wells in Aquifers*. Philadelphia.

American Society for Testing Materials (ASTM) 1992. *Standard Guide for Sampling Groundwater Monitoring Wells, Standard D4448-85a*; ASTM Standards on Environmental Sampling. 1995. Philadelphia, 220–233

- ASTM 1994. Standard D5521-94, *Standard Guide for Development of Groundwater Monitoring Wells in Granular Aquifers* 17 p., West Conshohocken, Pennsylvania.
- ASTM 1995b. Standard D5792-95, *Standard Practice for Generation of Environmental Data Related to Waste Management Activities: Development of Data Quality Objectives* 17 p.
- ASTM 1996. Standard D6001-96, *Standard Guide for Direct-Push Water Sampling for Geoenvironmental Investigations* 14 p., West Conshohocken, Pennsylvania.
- ASTM 1998. Standard D5092-90 (reapproved 1995) *Standard Practice for Design and Installation of Groundwater Monitoring Wells in Aquifers* 14 p., West Conshohocken, Pennsylvania.
- ASTM 2001. Standard D5314-92 (reapproved 2001), *Standard Guide for Soil Gas Monitoring in the Vadose Zone* 36 p., West Conshohocken, Pennsylvania.
- Barcelona, M.J., Gibb, J.P., Helfrich, J.A., and Garske, E.E. 1985. *Practical Guide for Groundwater Sampling*. EPA/600/2-85/104.
- Blake, S.B. and Hall, R.A. 1984. Monitoring petroleum spills with wells: some problems and solutions. Proceedings of the Fourth National Symposium and Exposition on Aquifer Restoration and Groundwater Monitoring. National Water Well Association. Columbus, OH. 305–310.
- Chappelle, F.H. 1993. *Groundwater Microbiology and Geochemistry*. John Wiley & Sons, New York.
- Cohen, R.M. and Mercer, J.W. 1993. *DNAPL Site Evaluation*. C. K. Smoley, Boca Raton, FL.
- Driscoll, F.G. 1986. *Groundwater and Wells*. Johnson Division, St. Paul, MN.
- Edge, R.W. and Cordry, K. 1989. The HydroPunch®: An *in situ* sampling tool for collecting groundwater from unconsolidated sediments. *Groundwater Monitoring Review*, 177–183.
- Fetter, C.W. 1993. *Contaminant Hydrogeology*. Macmillan Publishing Co., New York.
- Freeze, R.A. and Cherry, J.A. 1979. *Groundwater*. Prentice-Hall, Inc., Englewood Cliffs, NJ.
- Hem, J.D. 1989. Study and Interpretation of the Chemical Characteristics of Natural Water. U.S. Geological Survey Water Supply Paper 2254, 3rd Edition.
- Hvorslev, M.J. 1949. *Subsurface Exploration and Sampling of Soils for Civil Engineering Purposes*. U.S. Waterways Experiment Station, Vicksburg, MS. (Reprinted by Engineering Foundation, New York.)
- Kresic, Neven, 1997. Quantitative Solutions in Hydrogeology and Groundwater Modeling Law Engineering and Environmental Services, Inc. Lewis Publishers, New York.
- Lambe, T.W. and Whitman, R.W. 1969. *Soil Mechanics*. John Wiley & Sons, New York.
- Lueder, D.R. 1959. *Aerial Photographic Interpretation*. McGraw-Hill Book Company, New York.
- McDonald, M.G. and Harbaugh, A.W. 1988. *A Modular Three-Dimensional Finite-Difference Groundwater Flow Model*. USGS Techniques of Water-Resources Investigations, Book 6, Chapter A1, USGS, Reston, VA.
- Miller, J.C. and Miller, J.N. 1986. *Statistics for Analytical Chemistry*. John Wiley & Sons, New York.
- Moberly, R.L. 1985. Equipment Decontamination. *Groundwater Age*, 19:36–39.
- Nielsen, D.M., ed. 1991. *Practical Handbook of Groundwater Monitoring*. Lewis Publishers, Chelsea, MI.
- Puls, R.W. and Barcelona, M.J. 1989. Filtration of groundwater samples for metals analysis. *Hazardous Waste and Hazardous Materials*. 6.
- Puls, R.W. and Barcelona, M.J. 1996. Groundwater Issue Paper: Low-Flow (Minimal Drawdown) Groundwater Sampling Procedures; U.S. Environmental Protection Agency, EPA/540/S-95/504, 12 p.
- Rumer, R.R. and Mitchell, J.K. 1995. *Assessment of Barrier Containment Technologies — A Comprehensive Treatment for Environmental Remediation Applications*. National Technical Information Services, Springfield, VA.
- Scalf, M.R., McNabb, J. F., Dunlap, W.J., Crosby, R.L., and Fryberger, J. 1981. *Manual of Groundwater Quality Sampling Procedures*. EPA/600/2-81/160 (NTISPB82-103045).
- The Interstate Technology & Regulatory Council (ITRC), February, 2004. *Technical and Regulatory Guidance for Using Polyethylene Diffusion Bag Samplers to Monitor Volatile Organic Compounds in Groundwater* 78 p.
- U.S. Air Force Center for Environmental Excellence. 1997. *Long-Term Monitoring Optimization Guide*. Version 1.1.
- USEPA. 1975. *Manual of Water Well Construction Practices*. USEPA Office of Water Supply, Report No. EPA-5709-75-001. Washington, D.C.

- USEPA. 1986a. *RCRA Groundwater Monitoring Technical Enforcement Guidance Document*. Office of Waste Programs Enforcement and Office of Solid Waste and Emergency Response, OSWER-9950.1. Washington, D.C.
- USEPA. 1986b. *Test Methods for Evaluating Solid Wastes*. Office of Solid Waste and Emergency Response, Washington, D.C., SW-846.
- USEPA. 1987. *Alternate Concentration Limit Guidance, Interim Final EPA/530-SW-87-017*. 114 p. July.
- USEPA. 1989a. *Risk Assessment Guidance for Superfund: Interim Final Guidance*. Office of Emergency and Remedial Response (EPA/540/1-89/002). Washington, D.C.
- USEPA. 1989b. *Statistical Analysis of Groundwater Monitoring Data at RCRA Facilities, Interim Final Guidance*. Office of Solid Waste, Washington, D.C.
- USEPA. 1989c. *Seminar Publication — Transport and Fate of Contaminants in the Subsurface*. EPA/625/4-89/019. Washington, D.C.
- USEPA. 1989d. *Handbook of Suggested Practices for the Design and Installation of Groundwater Monitoring Wells*. PB90-159-807. Washington, D.C.
- USEPA. 1990. *Handbook — Groundwater, Volume I: Groundwater and Contamination*. EPA/625/6-90/016a. Washington, D.C.
- USEPA. 1991. *Handbook — Groundwater, Volume II: Methodology*. EPA/625/6-90/016b. Washington, D.C.
- USEPA. 1992a. *RCRA Groundwater Monitoring: Draft Technical Guidance*, Office of Solid Waste, EPA/530-4-93-001, PB93-139-350, Washington, D.C.
- USEPA. 1992b. *Statistical Analysis of Groundwater Monitoring Data at RCRA Facilities, Addendum to Interim Final Guidance*, Office of Solid Waste, Washington, D.C.
- USEPA. 1993a. Office of Research and Development. *Subsurface Characterization and Monitoring Techniques: A Desk Reference Guide, Volume I: Solids and Groundwater, Appendices A and B*. U.S. Environmental Protection Agency, Washington, D.C. EPA/625/R-93-003a.
- USEPA. 1993b. Office of Research and Development. *Subsurface Characterization and Monitoring Techniques: A Desk Reference Guide, Volume II: The Vadose Zone, Field Screening and Analytical Methods; Appendices C and D*. U.S. Environmental Protection Agency, Washington, D.C. EPA/625/R-93/003b.
- USEPA. 1993c. *Technical Manual — Solid Waste Disposal Facility Criteria*. EPA530-R-93-017, PB94-100-450.
- USEPA. 1994. *Handbook — Groundwater and Wellhead Protection*. EPA/625/R-94/001.
- USEPA. 1995. *Groundwater Sampling — A Workshop Summary*, November 30 to December 2, 1995, Dallas, Texas. EPA/600/R-94/205.
- USEPA. 2000. *Guidance for the Data Quality Objectives Process — EPA QA/G-4* EPA/600/R-96/055. Washington, D.C., 100 p. August.
- USEPA. 2002. *EPA Guidance for the Quality Assurance Project Plans — EPA QA/G-5* EPA/600/R-98/018. Washington, D.C., 111 p. December.
- Wilde, F.D., Radtke, D.B., Gibbs, J. and Jwatsubo, R.T. eds. 1998. National Field Manual for the Collection of Water Quality Data: US Geological Survey Techniques of Water-Resources Investigations. Book 9 Handbooks for Water-Resources Investigations, Variously Paginated.
- Wilson, C.R., Einberger, C.M., Jackson, R.L., and Mercer, R.B. 1992. Design of groundwater monitoring networks using the monitoring efficiency model (MEMO). *Groundwater Age*, 30:965-970.
- Yeskis, D. and Zavala, B., 2002. Groundwater Issue Paper: Groundwater Sampling Guidelines for Superfund and RCRA Project Managers, U.S. Environmental Protection Agency, EPA/542/S-02-001. 53 p.

Further Information

Nielsen (1991) gives a good, broad description of groundwater monitoring techniques. Driscoll (1986) gives a thorough discussion of groundwater well construction techniques and materials of construction for groundwater wells. USEPA (1993a, b) gives a detailed description of monitoring and analytical techniques for subsurface media. Complete citations for these works can be found in References.

36

Remediation of Contaminated Groundwater

36.1	Fundamentals	36-1
	Introduction • Groundwater Remediation Goals • Risks Associated with Contaminated Groundwater	
36.2	Groundwater Remediation System Design.....	36-6
	Introduction • Step 1. Define the Problem • Step 2. Define the Goal of Groundwater Remediation • Step 3. Screen Candidate Remedies • Step 4. Prepare Detailed Design • Step 5. Implement the Design • Step 6. Confirm the Effectiveness of the Design	
36.3	Hydraulic Containment of Groundwater	36-13
	Overview • Physical Barriers • Hydraulic Barriers • Other Options for Hydraulic Containment	
36.4	Design of Groundwater Extraction Systems	36-21
	Introduction • Extraction Well Systems • Extraction Trench Systems • Time Required to Extract a Plume of Contaminated Groundwater • Other Groundwater Extraction Approaches	
36.5	Treatment of Contaminated Groundwater	36-30
	Introduction • <i>In Situ</i> Treatment • <i>Ex Situ</i> Treatment	
36.6	Performance Monitoring of Groundwater Remediation Systems	36-43
	Glossary.....	36-44
	References	36-45
	Further Information	36-48

Michael F. Houlihan and
Michael H. Berman
GeoSyntec Consultants

36.1 Fundamentals

36.1.1 Introduction

This chapter presents comprehensive approaches for developing and implementing programs for the remediation of contaminated groundwater. It also presents references for further information on such approaches. In this section on fundamentals, the purpose of, and typical goals for, groundwater remediation programs are addressed, and the role of risk assessment in groundwater remediation programs is

described. Approaches to groundwater remediation are described based on the current requirements of the United States Environmental Protection Agency (USEPA) for remediating sites that are regulated under the Resource Conservation and Recovery Act (RCRA), the Comprehensive Environmental Response, Compensation and Liability Act (CERCLA), or analogous state programs. Other requirements, such as local regulations (as described in Chapter 32) and the need to mitigate human health or other environmental concerns posed by the contaminated groundwater may also define the needs of a groundwater remediation program.

36.1.2 Groundwater Remediation Goals

Groundwater is an extremely important source of water. In the United States, one-third of all drinking water is obtained from groundwater sources (Hutson et al., 2004). Further, a significant majority of Americans live near industrialized population centers, which are typically located near viable groundwater supplies. In the past several decades, instances of groundwater contamination have illustrated the adverse impacts that contaminated groundwater can have on human health. For all of these reasons, protecting the integrity of groundwater supplies is crucial to the protection of human health and the environment.

The need for remediation of contaminated groundwater is typically established based on the results of a groundwater assessment or monitoring program (see Chapter 35). In general, if the results of groundwater monitoring indicate that groundwater contains contaminants at concentrations that make it a threat to human health or the environment, then the groundwater may need to be remediated. Remediation, in this sense, is a broad term that refers to the reduction of risk caused by exposure to contaminated groundwater.

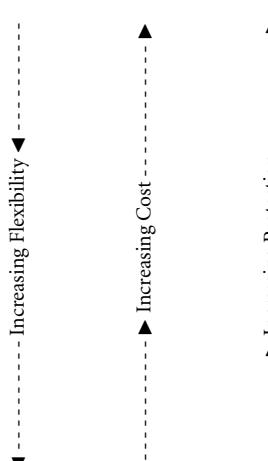
One of the most important steps in a remediation program is defining the goals of the program. There are many different goals that can be defined for a groundwater remediation program, including the following (see NRC, 1994):

- *Complete restoration*, which involves returning the aquifer to its condition prior to being contaminated
- *Nondegradation*, which involves removal of contaminants that exceed either the detection limits of available analytical equipment or background concentrations
- *Remediation to health-based standards*, which involves removal of contaminants that are present at a concentration that could cause adverse health effects (some examples of health-based standards are maximum contaminant levels (MCLs), alternate concentration limits (ACLs; see Section 36.1.3) and local or Federal drinking-water standards)
- *Remediation to risk-based standards*, which involves removal of contaminants that are present above a concentration that is determined to present an unacceptable risk to site-specific receptors based on anticipated use of the site
- *Remediation to the limits of technology-based standards*, which involves use of the best available technology to remove as much of the contaminants as possible
- *Partial-use restrictions* (or *institutional controls*), such as legal restrictions on the use of groundwater in areas where groundwater has been contaminated, or physical barriers (e.g., fences) to prevent access to contaminated media
- *Containment*, which involves the use of engineered systems for preventing migration of the contaminants to locations where receptors could be exposed to the contaminants.

A brief summary of the advantages and disadvantages of each of these remediation goals is presented in Table 36.1. When selecting a remediation goal, the degree to which groundwater can actually be remediated should be examined. For example, until the 1990s, pump-and-treat was considered one of the best technologies available for restoring groundwater quality. However, based on studies by the USEPA and others, the actual success rate for pump-and-treat remedies is extremely low. A summary of the USEPA's findings, as presented by the National Research Council (NRC) (1994), is provided in Table 36.2. Although there were various reasons for the failure of the systems to meet their goals, the data strongly suggest that, for most circumstances, groundwater pump-and-treat technologies alone are not capable of

TABLE 36.1 Advantage and Disadvantage of Cleanup Goals

Goal	Advantages	Disadvantages
Complete restoration Nondegradation	Eliminates all risk Reduction of contaminants to lowest level measurable	Likely impossible Extremely difficult, expensive, and time-consuming for many contaminants and hydrogeologic settings
Health-based standards	Designed to prevent measurable impacts to human health or environment	Health-based standards are difficult to define and may not accurately address all possible health impacts of exposure to contaminated groundwater
Risk-based standards	Designed to remove unacceptable risk to actual and anticipated site receptors based on anticipated use of site	May not reduce risk to acceptable levels if site use changes; may need to be used in conjunction with partial use restrictions
Technology-based standards	Allows treatment to the best standards capabilities of current technology	May not reduce risk to a level that is protective of human health and the environment
Partial use restrictions	Prevents contact between contaminants and receptors in a cost-effective manner	Leaves contaminants that could cause risk if partial use restrictions are ineffective
Containment	Relatively predictable and reliable; typically less costly than other remediation approaches	Leaves contamination that could migrate if containment system fails



Source: Adapted from National Research Council (NRC), Water Science and Technology Board. 1994. *Alternatives for Groundwater Cleanup*. Washington, DC. With permission.

TABLE 36.2 Summary of Pump-and-Treat System Performance Data

Category ^a	No. of sites	No. of sites containment achieved	No. of sites cleanup goal achieved	No. of sites goal not achieved
1	2	1	1	1
2	13	8	4	9
3	19	12	4	15
4	36	18	0	36

Note: ^a Indicates relative ease of cleanup, with 1 easiest and 4 hardest, as defined by NRC (1994) and as illustrated in Table 36.8.

This summary is based on information provided by NRC (1994) for sites where data were available regarding attainment of the cleanup goal.

Source: Adapted from National Research Council (NRC), Water Science and Technology Board. 1994. *Alternatives for Groundwater Cleanup*. Washington, DC. With permission.

restoring groundwater to its original quality or health-based standards. However, returning groundwater to its useful purpose, or eliminating health and environmental risks, does not always require restoration to original quality. Accordingly, a growing number of remediation systems are being designed with the goal of either containment or remediation to health-based standards. A typical approach for achieving this remediation goal is, first, to estimate the risk associated with exposure to the contaminated groundwater and then design a remediation system that will improve groundwater quality to a level that is protective of human health and the environment (i.e., health-based standards). It has become commonplace for regulators to approve remedies involving *in situ* remediation technologies (ISRTs, see Section 36.5.2) or monitored natural attenuation (MNA, see Section 36.5.2.5), to meet such goals. An approach for deriving an achievable, health-based remediation goal is presented in Section 36.1.3.

After a remediation goal has been selected, a remedy must be implemented to achieve the goal. In general, groundwater can be remediated in one or more of the following three manners.

- *Containment.* Containment of a plume of contaminated groundwater involves preventing the plume from migrating to a location where receptors can be exposed to it. Techniques for containing contaminated groundwater are described in Section 36.3.
- *Extraction.* For some remediation approaches, groundwater must be extracted from the ground, either to allow *ex situ* treatment or to provide hydraulic control of contaminated groundwater. Groundwater extraction approaches are described in Section 36.4.
- *Treatment.* Contaminated groundwater can be treated, either *in situ* or *ex situ*, using a variety of physical, chemical, or biological methods. Techniques for treatment of contaminated groundwater are described in Section 36.5.

Before beginning a program of groundwater remediation, it is recommended that the groundwater professional define a stepwise approach to developing and implementing the program. Recommended steps for developing and implementing a plan for groundwater remediation are presented in Section 36.2.

36.1.3 Risks Associated with Contaminated Groundwater

Contaminated groundwater may pose a risk to human health or the environment if it comes into contact with receptors and if the groundwater contains concentrations of constituents that could affect the health of the receptors. Risk can be defined as the probability of occurrence of human health or environmental impacts due to a release of contaminants to the environment. Risk can be evaluated by performing a baseline risk assessment, as described by USEPA (2000b) and as summarized in Figure 36.1. The goal of performing a baseline risk assessment is to estimate the current and future risks associated with contaminated groundwater and to compare the estimated risk to the acceptable threshold of risk (definition

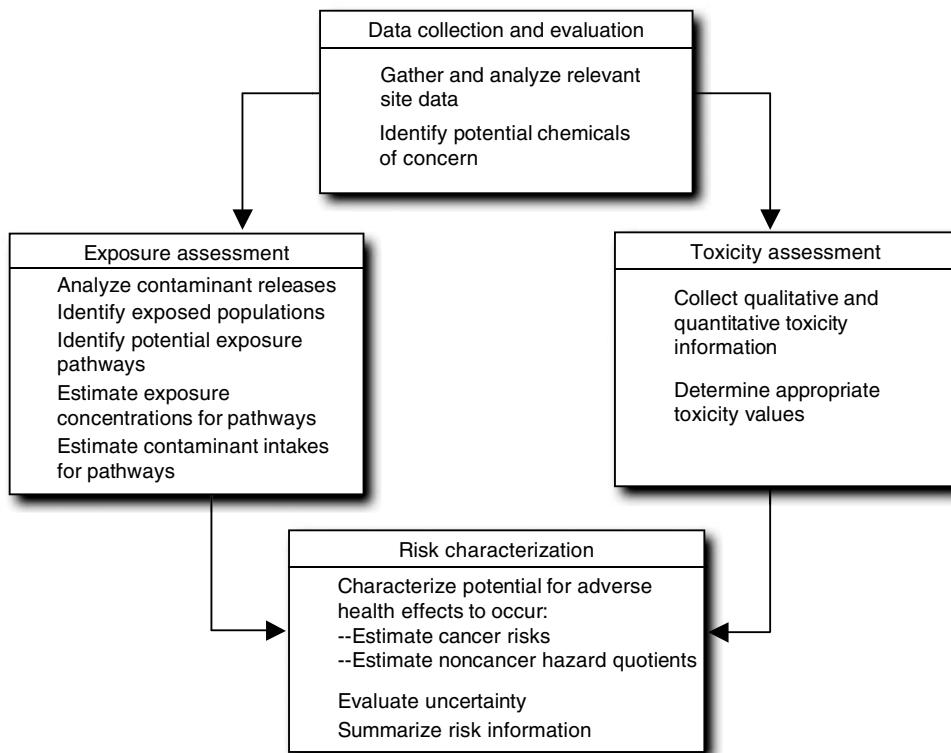


FIGURE 36.1 Baseline risk assessment. (From USEPA. 2000b. *Risk Assessment Guidance for Superfund: Interim Final Guidance*. Office of Emergency and Remedial Response. Washington, DC.)

of an “acceptable” level of risk is addressed at the end of this subsection); if the acceptable threshold level of risk is exceeded, groundwater remediation is likely warranted. Risk assessment techniques can also be used to estimate the risk that would remain after a remediation program has been implemented. Estimating the resulting future risk for several candidate remedies can provide a basis for selecting a remedy that reduces risk most effectively. Based on the results of a risk assessment, a non-zero concentration limit for contaminants in groundwater could be proposed that results in an acceptable risk to potential future receptors. For several contaminants, USEPA has defined exposure concentrations that result in a risk of 10^{-6} (i.e., or one additional cancer case in 1,000,000 people), these concentrations are defined as “Maximum Contaminant Levels” (or MCLs). USEPA also allows owners of contaminated sites to propose, on a case-by-case basis, Alternate Contaminant Levels (ACLs) that result in a site-specific exposure level (after considering the fate and transport of the contaminants) of 10^{-6} .

There are numerous sources of uncertainty in currently available risk-assessment methods (NRC, 1994). First, there is very little data to establish a conclusive link between human exposure to specific contaminants and specific health or environmental impacts. The accuracy of epidemiologic studies performed on humans to date is limited, with significant uncertainties related to: (1) exposure quantity and duration; (2) latency of observable impacts; (3) the relatively small size of the study populations; (4) inadequate control over comparison (i.e., “control”) groups; (5) inability to attribute observed health problems to only the known exposure; and (6) the synergistic effects in humans of exposure to some contaminants. Similarly, the accuracy of epidemiologic studies on animals is limited by: (1) the need to extrapolate the results of large-dose impacts on animals to estimate impacts of small doses on humans; and (2) the focus of such studies on single-contaminant exposures. Finally, because the concentrations of contaminants in groundwater may change over time due to factors such as treatment or attenuation over the exposure period (i.e., a lifetime), the quantity of contaminants to which a receptor may be exposed (as calculated in Step 2, *Exposure*

Assessment) could change during the period under consideration. Because of these uncertainties, risk assessment results have a degree of uncertainty. To account for this degree of uncertainty, very conservative assumptions are typically used when estimating toxicity and exposure quantity to minimize the possibility of underestimating the actual risk. Therefore, the estimated risk may be extremely conservative.

After the risk of exposure to contaminated groundwater has been estimated, the groundwater professional, stakeholder, or regulator must decide whether or not the risk is high enough to warrant remediation. This decision process, which is referred to as risk management (a process that is separate from risk assessment), is typically performed based on input from both regulatory agencies and the population likely to be exposed to the risk. For example, a decision has been made by the US Congress, as presented in the National Contingency Plan, that a risk level of 10^{-4} to 10^{-6} is acceptable for contaminants at Superfund sites. The USEPA typically considers a similar range of risk to be acceptable at RCRA sites. The acceptable level of risk may vary from location to location and is a balance between the benefits to the public of a clean environment and the public cost of attaining a clean environment. The benefit of a clean environment is difficult to determine, but can be estimated in terms of dollars that the public would be willing to spend to clean up the environment (i.e., the “contingent valuation” approach); however, this approach is complex because of the great number of site-specific variables and non-quantifiable factors that must be considered (NRC, 1994; USEPA, 2002). Because of difficulties in obtaining an accurate estimate of the value of a clean environment, the commonly used value for acceptable risk (e.g., 10^{-6}) has been assigned.

36.2 Groundwater Remediation System Design

36.2.1 Introduction

The design of an efficient, cost-effective groundwater remediation system requires a comprehensive understanding of the nature and extent of contamination, the remediation objectives, and a careful evaluation of remedial technologies and their abilities to meet the objectives. When preparing a remediation system design, the use of a stepwise approach can simplify the design process and minimize the chances of failure of the system to achieve the remediation goal. In this section, recommended steps are presented for preparing a design to remediate contaminated groundwater. In subsequent sections, approaches to the design of some typical remediation systems are described.

36.2.2 Step 1. Define the Problem

The first step in the design of the groundwater remediation system is to define the problem in adequate detail to allow the design of an efficient and cost-effective groundwater remediation system. The problem should be defined in terms of the extent of contamination, the risk associated with the contaminated groundwater, the regulations that apply to the problem, the degree of remediation required, and the subsurface conditions. The extent of contamination may be defined using procedures described by the USEPA (2004a) and others. If the vertical and horizontal extent of contamination have not been adequately evaluated during the site characterization or groundwater monitoring program, then additional field exploration data should be collected to complete this evaluation. In some cases, such as sites contaminated with DNAPL, it may not be economically feasible to fully define the source of contamination. The risk associated with contaminated groundwater can be estimated as described in Section 36.1.3 of this chapter. The regulations that apply may be identified based on the information presented in Chapter 32 of this handbook, and based on input from legal counsel, if necessary. Subsurface conditions may be defined as described in Chapter 2. The various manners in which groundwater can become contaminated are illustrated in Figure 36.2; hydrogeologic considerations for groundwater remediation are summarized in Table 36.3 and Table 36.4. It should be noted that the hydrogeologic properties described in Table 36.3 also are a key component in understanding how treatment agents injected in association with some ISRTs will be transported. As illustrated in Table 36.5, the hydrogeology of the site (as well as other factors) has a strong bearing on the type of remediation system that can be considered.

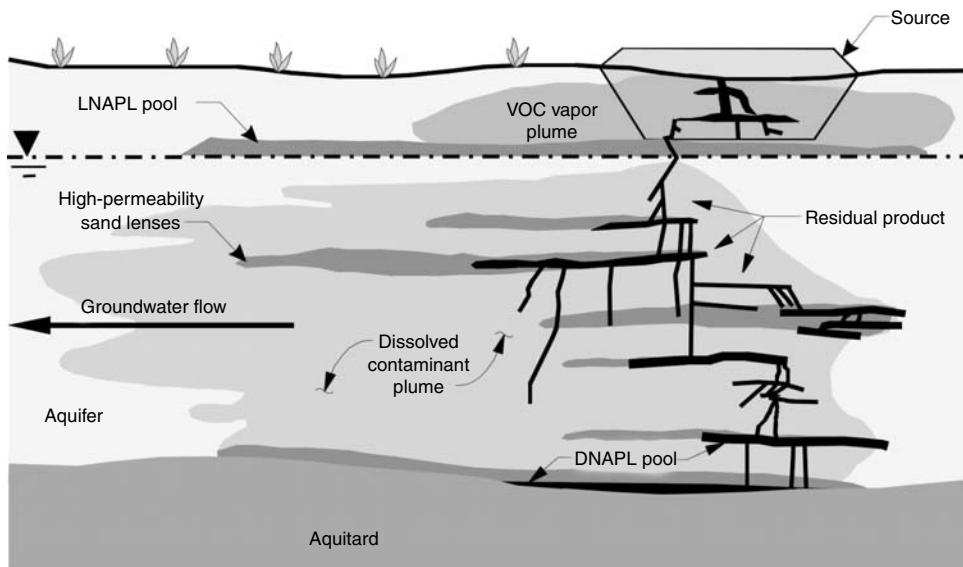


FIGURE 36.2 Groundwater contamination by dissolved vapor and nonaqueous phase constituents.

TABLE 36.3 Hydraulic Properties of Aquifers Important for Groundwater Cleanup

Property	Description	Importance for groundwater cleanup
Porosity	Volume of pore space relative to the total volume	Pores store water and contaminants
Effective porosity	Interconnected pore space that can transmit fluid	Water and contaminants flow through interconnected pores
Groundwater velocity	Rate of fluid movement	Influences the direction and velocity of dissolved contaminant movement
Hydraulic gradient	Elevation and pressure differences that cause fluids to flow	Influences the direction of contaminant movement
Hydraulic conductivity	Ease with which water can move through a geologic formation	Influences the rate at which fluid can be pumped for treatment, as well as the rate and direction of groundwater migration
Transmissivity	Product of formation thickness and hydraulic conductivity	Influences the rate at which a plume of contaminants migrates and the rate at which fluid can be pumped for treatment
Storage coefficient	Volume released by pressure changes per unit area during pumping in a confined aquifer	Influences the quantity of fluid that can be obtained by pumping
Specific yield	Fraction of total pore volume released as water by gravity drainage during pumping of an unconfined aquifer	Influences the quantity of fluid that can be obtained by pumping
Specific retention	Fraction of total aquifer volume retained as water above the water table after pumping an unconfined formation	Influences the quantity of contaminant that remains in the subsurface after pumping

Source: Adapted from National Research Council (NRC), Water Science and Technology Board. 1994. *Alternatives for Groundwater Cleanup*. Washington, DC. With permission.

36.2.3 Step 2. Define the Goal of Groundwater Remediation

The next step in designing a groundwater remediation system is to define the goal of remediation. This step involves: (1) identifying remediation technologies that could possibly be used to remediate the contaminated groundwater; then (2) screening the list to identify remediation technologies that could be

TABLE 36.4 Summary of the Mechanisms Influencing the Fate of Contaminants in the Environment

Process	Environmental conditions	Contaminant
Movement (i.e., advection)	Water flow rate	Amount of material
	Formation permeability	Physical state
	Water motion	Solubility
	Gravity	Viscosity
	Surface tension	
Retention (physical processes affecting transport)	Soil/Sediment	Aqueous solubility
	Organic matter content	Ionic character
	Sorptive capacity	
Reaction (chemical processes affecting transport)	pH	Chemical transformation
	Redox status	Biodegradability
	Microbial communities	

Source: From USEPA. September 1989. Seminar Publication — *Transport and Fate of Contaminants in the Subsurface*. EPA/636/4-89/019. Washington, DC.

successfully implemented at the site; and finally (3) selecting an appropriate remediation goal based on the remedies that can be successfully implemented. A summary of remediation technologies that could be used is presented in Table 36.6. Candidate goals for remediating contaminated groundwater are described in Section 36.1.2; candidate remediation approaches that correspond to these goals are presented in Table 36.7.

For sites where it is demonstrated that remediation is needed, the remediation goal should be developed with as complete an understanding as possible of the goals that are actually attainable. This begins with an understanding of the difficulty of remediating or restoring groundwater quality in certain hydrogeologic settings, as summarized in Table 36.5. Then, the possibility of achieving the remediation goal should be evaluated based on past experience with similar problems in similar hydrogeologic settings. For example, the likelihood of success of groundwater restoration (based on the profession's documented experience in achieving health-based remediation goals) is illustrated in Table 36.8, which was developed based on an analysis of the information presented in Table 36.2. As shown in Table 36.8, it is currently feasible to remediate some types of sites to health-based standards (e.g., homogeneous aquifers contaminated with dissolved, mobile contaminants, as shown in "Group A" in Table 36.8), but not other types of sites (e.g., heterogeneous or fractured rock aquifers contaminated with non-aqueous phase liquid contaminants "Group C" in Table 36.8). As a further consideration, restoration of groundwater to original quality represents a significant financial burden to the nation, as illustrated in Table 36.9. Regardless of the goal selected, it should be noted that, as stated by the NRC (1994), "The attainment of zero contaminant concentration as an outcome for groundwater remediation should be recognized as an unattainable goal no matter how far cleanup technologies advance in the future." The difficulty of restoring groundwater quality at some sites is also addressed by the USEPA in its document entitled, *Guidance for Evaluating the Technical Impracticability of Ground-water Restoration* (1993). In this document, the USEPA identifies several hydrogeologic settings (based on information in Table 36.5) in which, based on the USEPA's experience, groundwater quality likely cannot be restored to background conditions. Once the remediation goal has been selected and feasible remedies have been identified that could be implemented at the site, the groundwater professional should screen the list of candidate remedies to those that, based on experience, can be expected to achieve the remediation goal. These remedies should then be further evaluated during Step 3.

36.2.4 Step 3. Screen Candidate Remedies

After a list of feasible candidate remediation technologies has been prepared, the remedies should be screened to identify the remedy that can most efficiently and cost effectively achieve the remediation goal. Screening involves evaluating the likely effectiveness and cost of the feasible remediation technologies

TABLE 36.5 Examples of Factors Affecting Groundwater Remediation

Characteristic	Less difficult	Generalized remediation difficulty scale		More difficult
		→	→	
Nature of release	Small volume Short duration Slug release	→	→	Large volume Long duration Continual
Biotic/Abiotic decay potential	High	→	Chemical properties	Low
Volatility	High	→	→	Low
Contaminant retardation (sorption) potential				
Tendency for larger plume	High	→	→	Low
Difficulty removing contaminants	Low	→	→	High
Contaminant phase	Aqueous, gaseous	→	Contaminant distribution	
Volume of contaminated media	Small	→	Sorbed → LNAPLs	DNAPLs
Contaminant depth	Shallow	→	→	Large
Stratigraphy	Simple geology, (e.g., planar bedding)	→	Geology	Complex geology, (e.g., interbedded and discontinuous strata, fractured media)
Texture of unconsolidated deposits	Sand	→	→	Clay
Degree of heterogeneity	Homogeneous (e.g., well-sorted sand)	→	→	Heterogeneous (e.g., interbedded sand and silts, clays, fractured media, karst)
Hydraulic conductivity	High (10^{-2} cm/sec)	→	Hydraulic/Flow	Low (10^{-6} cm/sec)
Temporal variation	Little/None	→	→	High
Vertical flow	Little	→	→	large downward flow component

Source: From USEPA. 1993. *Guidance for Evaluating the Technical Impracticability of Ground-Water Restoration*. EPA-540-R-93-080. Washington, DC.

and selecting a remedy for detailed design based on the results of this evaluation. Three valuable steps in performing the evaluation are described below.

- *Prepare Preliminary Conceptual Design for Each Candidate Remedy.* By preparing a conceptual design, the effectiveness and cost of the remedy can be evaluated, as described below. Developing and implementing design criteria can be an effective way to focus the conceptual design effort on the keys to success for the project. Design criteria are guidelines to be used during design that, if achieved, will likely result in successful performance of the remediation system. These criteria are defined by the owner, engineer, and the operator. Design criteria should consider

TABLE 36.6 Summary of Commonly Used Groundwater Remediation Technologies

Alternative technology	Residual groundwater concentration	Residual sorbed concentration in source area	Cleanup time ^a	Number of peer-reviewed publications ^b
<i>Source remediation</i>				
Conventional pump-and-treat	Low to medium	Medium to high	Long	Some
Vacuum extraction and bioventing	NA	Low to medium	Short	Some
Air sparging (vertical or horizontal wells)	Low to medium	Low to medium	Short to medium	Limited
<i>In situ</i> bioremediation-hydrocarbons	Low to medium	Low to high	Short to medium	Some
<i>In situ</i> bioremediation-chlorinated solvents	Low to medium	Low to high	Medium to long	Some
Cosolvent and surfactant flushing	Low to medium	Low to medium	Short to medium	Some
Steam stripping	Low to medium	Low to medium	Short	Some
Resistivity heating	Low to medium	Low to medium	Short	Limited
<i>In situ</i> thermal desorption	Low to medium	Low to medium	Short	Some
<i>In situ</i> chemical oxidation	Low to medium	Medium to high	Medium	Some
<i>In situ</i> bioremediation-metals	Low to medium	Low to high	Medium to high	Limited
Monitored natural attenuation	Low to medium	Low to high	Long	Limited
<i>Plume remediation</i>				
Conventional pump and treat	Low	Medium to high	Long	Many
Air sparging (vertical or horizontal wells)	Low to medium	Low to medium	Medium to long	Some
<i>In situ</i> bioremediation- hydrocarbons	Low to medium	Low to high	Medium to long	Many
<i>In situ</i> bioremediation-chlorinated solvents	Low to medium	Low to high	Medium to long	Many
<i>In situ</i> reactive barriers	Low	NA	Long	Some
Monitored natural attenuation	Low to medium	Low to medium	Long	Many

Note: A “low” residual concentration and “medium” cleanup time reflect relatively good performance, while a “high” residual concentration and “long” cleanup time reflect much less effective performance. “NA” denotes that the technology is not applicable to this situation. Cost of the technology and feasibility for different situations are not addressed in this table but should be carefully evaluated before selecting a technology.

^a Because few cases of achieving cleanup goals have been reported, these qualitative assessments reflect the judgment of NRC.

^b “Limited” indicates that very little information about this technology is available in peer-reviewed publications, while “some” indicates a greater availability of information.

Source: Modified from National Research Council (NRC), Water Science and Technology Board. 1994. *Alternatives for Groundwater Cleanup*. Washington, DC. With permission.

constructability concerns, performance limitations of the technology, budget constraints, operation and maintenance requirements, and regulatory requirements. The candidate remedies can be compared by examining the results of these evaluations and selecting remedies that can most cost-effectively achieve the remediation goal. Conceptual designs should then be prepared that address the design criteria that are relevant at the conceptual design stage.

- *Evaluate Effectiveness of Conceptual Designs.* The effectiveness of the designs in meeting the remediation goal should be evaluated next. This evaluation can be performed objectively (e.g., by modeling the performance of the remediation system as described in Chapter 18 and Chapter 19 of this handbook) or subjectively (using the information on remedies presented in Table 36.6 and the information on remediation feasibility presented in Table 36.8). Also, the time required to achieve the remediation goal should be estimated for use in the subsequent cost evaluation.
- *Evaluate Cost of Remedies.* The cost of each remedy should be estimated. A good reference for remediation cost estimating techniques is presented in USEPA (2000a). In particular, the present-worth value of each remedy should be calculated, considering both capital and long-term operation and maintenance costs, to provide a consistent basis for comparing the costs of the candidate remedies.

TABLE 36.7 Candidate Groundwater Remediation Goals and Corresponding Remediation Approaches

Remediation goal	Examples of candidate remediation approaches
Groundwater restoration	Monitored natural attenuation Bioremediation Combinations of conventional remediation approaches
Nondegradation	Source removal Conventional pump-and-treat Bioremediation Soil vapor extraction/Air sparging
Return aquifer to health-based standards	Source removal Conventional pump-and-treat Bioremediation Soil vapor extraction/Air sparging
Apply technology-based standards	Phytoremediation Electrokinetics Solvent extraction Thermal desorption
Implement restricted-use policies	Deed restrictions Local ordinances
Containment	Physical barriers Hydraulic barriers Capping Reactive barriers

Source: Adapted from National Research Council (NRC), Water Science and Technology Board. 1994. *Alternatives for Groundwater Cleanup*. Washington, DC. With permission.

TABLE 36.8 Likelihood of Success of Groundwater Restoration

Hydrogeology	Contaminant chemistry					
	Mobile, dissolved (degrades/ volatilizes)	Mobile, dissolved	Strongly sorbed, dissolved (degrades/ volatilizes)	Strongly sorbed, dissolved	Separate phase LNAPL	Separate phase DNAPL
Homogeneous, single layer	A (1)	A (1-2)	B (2)	B (2-3)	B (2-3)	B (3)
Homogeneous, multiple layers	A (1)	A (1-2)	B (2)	B (2-3)	B (2-3)	B (3)
Heterogeneous, single layer	B (2)	B (2)	B (3)	B (3)	B (3)	C (4)
Heterogeneous, multiple layers	B (2)	B (2)	B (3)	B (3)	B (3)	C (4)
Fractured	B (3)	B (3)	B (3)	B (3)	C (4)	C (4)

Note: Group A represents types of sites for which cleanup of the full site to health-based standards should be feasible with current technology. Group C represents types of sites for which full cleanup of the source areas to health-based standards will likely be technically infeasible. Group B represents sites for which the technical feasibility of complete cleanup is likely to be uncertain. The numerical ratings indicate the relative ease of cleanup, where 1 is easiest and 4 is most difficult.

Adapted from National Research Council (NRC), Water Science and Technology Board. 1994. *Alternatives for Groundwater Cleanup*. Washington, DC. With permission.

TABLE 36.9 Costs of Various National Policies for Hazardous Waste Site Remediation

National policy	Present value of resource cost (billions of 1991 U.S. dollars)		
	Lower bound	Best guess	Upper bound
Current policy ^a	180	280	390
Less stringent policy ^b	140	180	260
More stringent policy ^c	360	440	630

Note: This table converts the figures in Russell et al. (1991) to present value by assuming that costs are prorated equally each year for a 30-year time period and that the discount rate is 4 percent.

^a According to Russell et al. (1991), current policy means “the set of principles and practices for hazardous waste remediation that are inferred to be in place in the period 1988–91 when the experience base and data for this study were collected.”

^b According to Russell et al. (1991), less stringent policy means relying more on containment and less on full cleanup.

^c According to Russell et al. (1991), more stringent policy means application of more intensive treatment technologies and reduced burden on future generations.

Source: Adapted from National Research Council (NRC), Water Science and Technology Board. 1994. *Alternatives for Groundwater Cleanup*. Washington, DC. With permission.

For remedies that require more technically complex remediation technologies, the process of screening candidate remedies tends to incorporate a more gradual scale up to a full-scale remedy through bench studies and pilot tests. Bench studies are often employed during the early stages of the screening process. For example, if bioremediation were to be considered as a remedy, a bench study would be conducted to determine whether native microbes present in the aquifer at the site are capable of degrading the contaminants. Then, pilot studies are used during the latter stages of technology screening, or even after a remediation technology has already been selected, to provide confirmation that a particular approach will be effective for the specific site and to obtain additional data that can be used during the detailed design stage (Section 36.2.5).

36.2.5 Step 4. Prepare Detailed Design

After the preferred remedy has been selected, a detailed design of the approach should be prepared that is suitable for construction of the system. Recommendations for design of specific elements of some of the more commonly used groundwater remediation systems are presented in Section 36.3, Section 36.4, and Section 36.5. USEPA (1995a) provides an overall framework for developing a remedial design. It is recommended that detailed calculations be performed to demonstrate that each of the design criteria is satisfied. Also, it is recommended that the ability of the design to mitigate adverse impacts to human health or the environment be demonstrated; ideally, this is accomplished by performing a risk assessment (as described in Section 36.1.3) for the conditions that will exist during and after operation of the remediation system. An effective remedy will achieve the remediation goal, mitigate the risk of exposure to contaminated groundwater, and provide reliable service throughout the design life of the system.

Like all engineering design activities, remedial design is an iterative process. Therefore, the design should be critiqued several times to identify improvements that can be made to enhance the performance or reduce the cost of the system. Such an evaluation could be performed by subjecting the preliminary design to a detailed review by: (1) a peer reviewer who is qualified in the technology being considered and who has not been an integral part of the design process (to provide an unbiased, critical evaluation of the design); (2) a contractor (to provide an evaluation of the constructability of the design); (3) a risk

assessment professional (to give an indication of the ability of the design to meet the remedial action objectives for mitigation of risks to human health or the environment); or (4) an operator of a similar system (for a value-engineering review or an evaluation of the potential costs or difficulties in operating the proposed system). Although such reviews may appear to be somewhat time consuming, they frequently result in valuable input that can improve the likelihood that the design criteria will be met. Also, because such reviews reduce the possibility of failure of the remediation system, they can be an effective component of a liability and loss prevention program.

36.2.6 Step 5. Implement the Design

After the design has been completed and reviewed, it must be implemented, which involves constructing and operating the system. The first step in constructing the system is deciding how and by whom the system will be constructed. There are several methods of arranging to construct a remediation system, including: (1) construction by the owner; (2) soliciting competitive bids from independent contractors qualified to perform the work; (3) negotiating a contract with an independent contractor; or (4) retaining a design/build contractor. For the last three approaches to construction, a detailed set of bid documents should be prepared that can be used by the bidder to prepare a responsive bid. The advantages of procuring a contractor through a competitive bid process include: (1) a generally lower construction cost, resulting from the competitive nature of the bid process; (2) well-defined roles of each party to the construction contract and well-defined terms of the construction contract; and (3) documentation of the information that the contractor used to prepare its bid, which can be valuable if additions or subtractions to the contract scope of work and budget are required. The disadvantages of competitively bidding work are: (1) the formality required, which involves a preparation of detailed plans and legal documents, sometimes at a significant cost; and (2) difficulty in factoring in “subjective” criteria in the bid evaluation, such as the contractor’s qualifications for performing the work and the contractor’s past performance on similar projects.

36.2.7 Step 6. Confirm the Effectiveness of the Design

The final step of a remediation program is confirmation of the effectiveness of the system in achieving the remediation goal. The effectiveness of the design can be evaluated simply by reviewing the data from a performance-monitoring program (as described in Chapter 35) and assessing whether or not the remediation goals and the performance criteria are being satisfied. The effectiveness of the design should be evaluated both during the startup of operations (if the system contains components that require operation and maintenance) and at specified time intervals after startup. As an example of confirming the effectiveness of a groundwater pump-and-treat system design that is intended to prevent migration of contaminants beyond a line of extraction wells, the system could be monitored to ensure that the groundwater downgradient of the extraction wells does not exceed design cleanup criteria; similarly, if the system is designed to remove contaminants in the aquifer, then groundwater collected by the removal wells could be tested to confirm that groundwater quality is improving as predicted. As a subsequent step after evaluating the performance of the remediation system, the system should be modified as needed to improve the system’s ability to achieve the remediation goal and performance criteria. As described in Section 36.6, the performance-monitoring program should contain criteria for “shutting off” the system (as verified using data collected during performance monitoring); also, the performance-monitoring program should contain criteria for reevaluating the adequacy of the design if the system has failed to meet the remedial design criteria.

36.3 Hydraulic Containment of Groundwater

36.3.1 Overview

There are many alternatives for providing hydraulic containment of groundwater. The most widely used features for containing groundwater are physical barriers and hydraulic barriers. Therefore, this section

TABLE 36.10 Types of Physical Barriers

Barrier type	Width (ft)	Maximum depth (apx) (ft)	Unit cost (\$[1995]/sf)	Production rate per 10 hrs (sf)
Soil bentonite	2–3	80	2–8	2,500–15,000
Cement bentonite	2–3	80	5–18	1,000–8,000
Biopolymer drain	2–3	70	7–36	1,500–5,000
Deep mixing	2.5	90	6–15	1,000–8,000
DM structural	2.5	90	15–30	1,000–3,000
Jet grouting	1.5–3	200	30–80	300–2,500
Grout curtain	One row	200	40–100	200–1,000
Sheet piling	One sheet	150	15–40	2,000–15,000
Geochemical barrier	Varies	Hundreds	8–400	100–1,000

Source: From Rumer, R.R. and Mitchell, J.K. 1995. *Assessment of Barrier Containment Technologies—A Comprehensive Treatment for Environmental Remediation Applications*. National Technical Information Services, Springfield, VA.

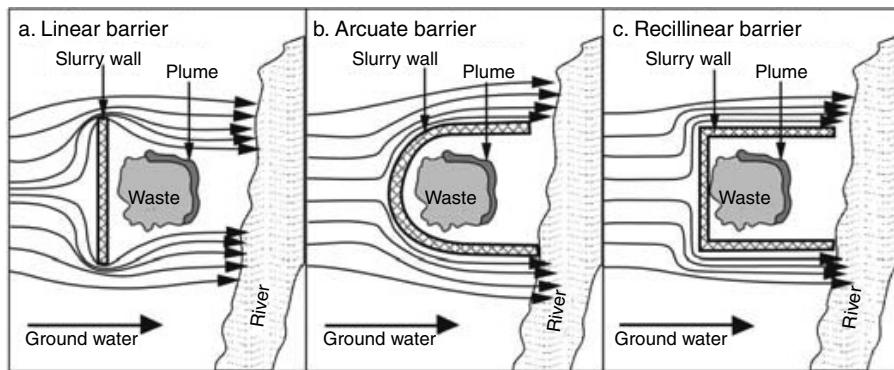
is focused on physical barriers (Section 36.3.2) and hydraulic barriers (Section 36.3.3). Also, innovative hydraulic containment methods are described in Section 36.3.4. Recommended methods for performance monitoring of hydraulic containment features are summarized in Section 36.6.

36.3.2 Physical Barriers

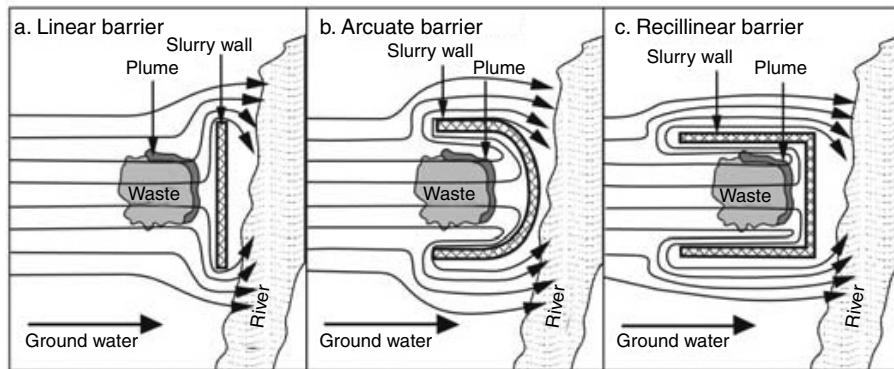
Physical barriers are vertical features in the ground that provide a barrier to groundwater flow. Physical barriers can be formed using a variety of materials and can be either nonselective (i.e., a barrier to the flow of all groundwater) or selective (i.e., a barrier to only the migration of target contaminants). In this section, only nonselective physical barriers are discussed; selective physical barriers (i.e., “permeable reaction barriers”) are described in Section 36.5.2.5. A list of several of the most commonly used physical barriers is presented in Table 36.10, along with their typical dimensions, range of unit costs, and approximate production (i.e., construction) rates.

The oldest and most commonly used types of physical barriers are cutoff walls; several types are illustrated in Figure 36.3. Cutoff walls limit the migration of groundwater by forming a physical impediment to groundwater flow. The effectiveness of the wall is a function of its continuity, its resistance to degradation by contaminants in groundwater, and its resistance to physical degradation. Note that, because it impedes groundwater flow, the elevation of groundwater may rise (or “mound”) on the upgradient side of the barrier; to prevent overtopping or flanking of the barrier by contaminated groundwater, other controls may be needed, such as (1) groundwater extraction wells or trenches adjacent to the barrier to route mounded groundwater to a discharge point; or (2) permeable segments in the barrier to allow water to pass through controlled locations in the wall.

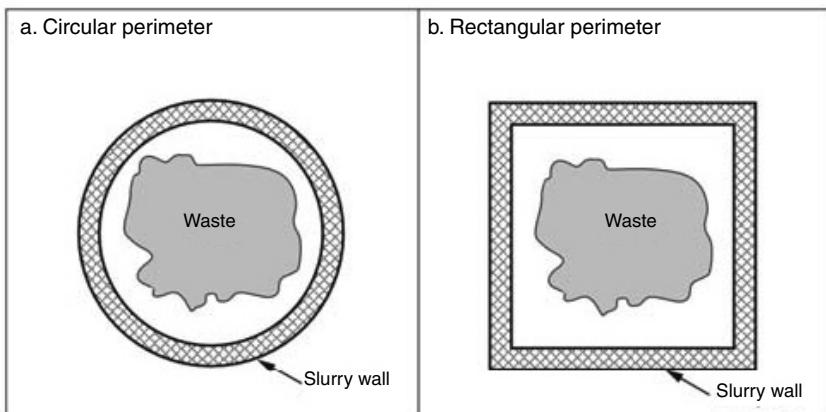
Soil-bentonite slurry walls are, by far, the most common type of physical barrier used in geotechnical and environmental remediation projects. Slurry walls have been used in the U.S. since the 1940s. Construction of a slurry wall consists of installing a mixture of soil (or other material) and bentonite clay into a vertical trench to form a very low-permeability, vertical barrier to groundwater flow. A general approach for slurry wall construction is illustrated in Figure 36.4 and includes: (1) excavation of a narrow (i.e., about 2 to 3 ft, or 0.6 to 1 m, wide) vertical trench (if the unsupported trench could cave, then the trench should be temporarily supported using a bentonite slurry); and (2) placing the soil-bentonite mix in the slurry-filled trench in a manner that displaces the bentonite slurry and forms a continuous vertical wall of low-permeability soil-bentonite material. A typical soil-bentonite mix contains about 2 to 4% bentonite clay, 20 to 35% water, 15 to 40% soil finer than the No. 200 U.S. standard sieve, and about 40 to 60% material coarser than the No. 200 U.S. standard sieve (Inyang, 1992). In many cases, the material excavated from the trench can be used as an economical source of the soil component of the soil-bentonite mix. Factors that must be considered in selecting a source for soil to be mixed with the bentonite include: (1) the grain-size



Plan views of upgradient slurry walls for partial containment of buried waste and/or plume



Plan views of downgradient slurry walls for partial containment of buried waste and/or plume



Plan views of waste totally enclosed by slurry walls

FIGURE 36.3 Examples of physical barrier configurations. (From GeoSyntec Consultants. 1994. *Subsurface Barriers*. Report to the United States Navy, Contract SBIR Topic N93-130. Atlanta, Georgia. 80 pages.)

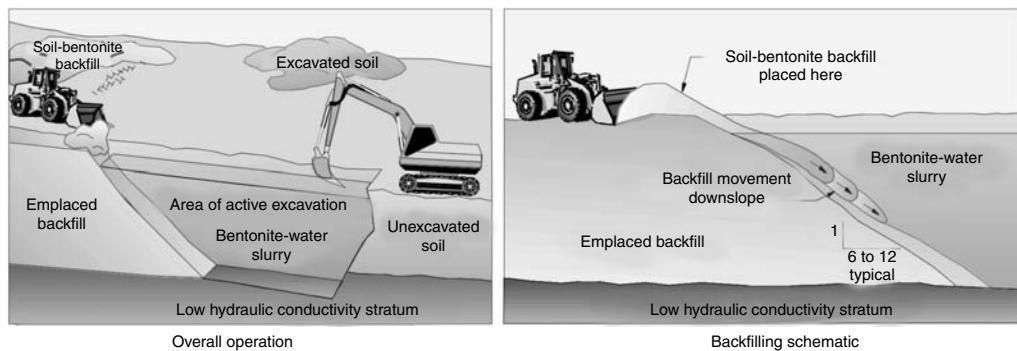


FIGURE 36.4 Soil-bentonite slurry wall construction. (From GeoSyntec Consultants. 1994. *Subsurface Barriers*. Report to the United States Navy, Contract SBIR Topic N93-130. Atlanta, Georgia. 80 pages.)

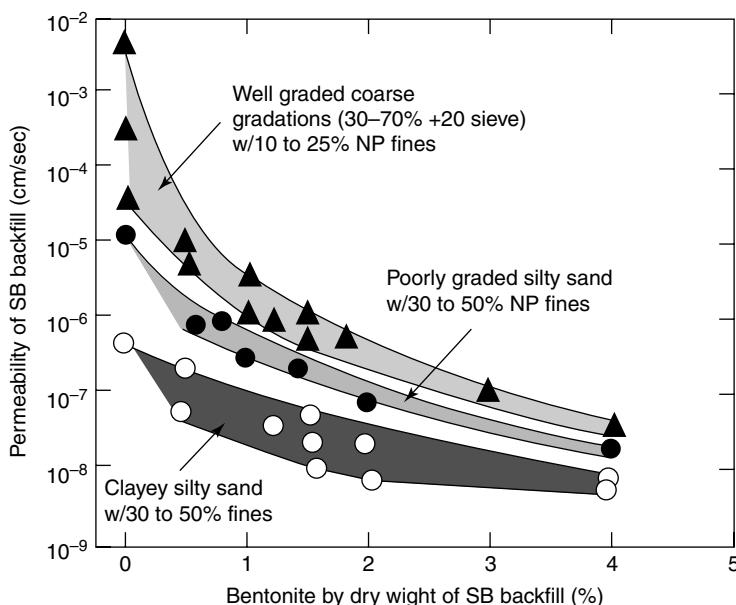


FIGURE 36.5 Permeability as a function of percent bentonite. (From GeoSyntec Consultants. 1994. *Subsurface Barriers*. Report to the United States Navy, Contract SBIR Topic N93-130. Atlanta, Georgia. 80 pages.)

distribution of the soil; (2) treatment and disposal costs for excavated trench soils that are contaminated; and (3) the chemical composition of the soil (e.g., soil that is contaminated with metals or VOCs could inhibit the ability of the bentonite to hydrate and, therefore, compromise the permeability of the wall). An excellent reference for design and construction of slurry walls is provided by Xanthakos (1979).

The effectiveness of a soil-bentonite slurry wall in preventing flow through the wall is a function of: (1) the environment of the wall (i.e., the nature of the chemicals in groundwater to which the soil-bentonite mix will be exposed); (2) the percentage of bentonite clay used in the soil-bentonite mix (as illustrated in Figure 36.5); (3) the type and gradation of soil used typically measured as a percentage of the soil that is finer than the No. 200 U.S. standard sieve (as illustrated in Figure 36.6); and (4) the quality of wall construction. When designing a soil-bentonite slurry wall, laboratory hydraulic conductivity tests are typically performed on candidate mixtures of soil and bentonite clay, and the test results are used to select the proportions of soil and bentonite that will produce the desired *in situ* hydraulic conductivity.

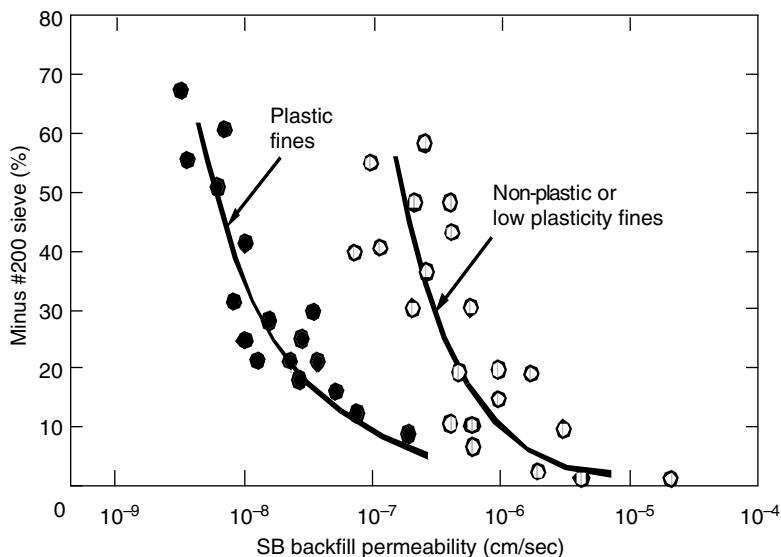


FIGURE 36.6 Permeability as a function of soil fines content. (From GeoSyntec Consultants. 1994. *Subsurface Barriers*. Report to the United States Navy, Contract SBIR Topic N93-130. Atlanta, Georgia. 80 pages.)

Based on widespread experience with soil-bentonite slurry walls over the past half century, it is generally recognized that the lowest in-place hydraulic conductivity that can reasonably be achieved for a soil-bentonite slurry wall is typically between 10^{-6} cm/sec and 10^{-7} cm/sec, with hydraulic conductivities in the lower end of this range attainable only under ideal circumstances and using excellent construction practices (GeoSyntec 1994). Finally, note that the performance of a soil-bentonite slurry wall can be adversely affected by exposure of the wall to VOCs, aqueous salts (such as those contained in seawater), or NAPLs. Sai and Anderson (1992) reported an increase of two to three orders of magnitude in hydraulic conductivity for soil-bentonite slurries that were permeated with the VOCs xylene and methanol. To improve the resistance of slurry walls to permeation by VOCs, some researchers are experimenting with the addition of activated carbon to the slurry.

36.3.3 Hydraulic Barriers

The term *hydraulic barrier* is used to describe a feature that causes a depression in the piezometric surface of groundwater and acts as a barrier beyond which groundwater within the zone of influence of the barrier should not flow. Hydraulic barriers may be formed using trenches, wells, or other features that remove groundwater, thus depressing the potentiometric surface. Groundwater collection trenches and groundwater extraction wells, which are the most common types of hydraulic barriers, are described in this section.

Groundwater Collection Trenches. Trenches can be installed in an aquifer to produce an induced hydraulic barrier by depressing the potentiometric surface of the aquifer. A typical groundwater extraction trench system is illustrated in Figure 36.7. During design, the designer must identify the depth to which the trench must be constructed to induce a hydraulic barrier. The depth of the trench should be carefully selected based on consideration of the seasonal variations in the elevation of groundwater, the location of contaminants (i.e., “floating” LNAPLs, dissolved constituents, or “sinking” DNAPLs), and the depth of contaminants in the aquifer. Trenches may be more cost-effective than wells for extracting groundwater if contaminants are located at a shallow depth below ground (i.e., less than about 50 ft [15 m]) and the aquifer materials are relatively easy to excavate (e.g., no excavation of bedrock). The depth of influence of a groundwater collection trench in an isotropic, homogeneous aquifer is illustrated in Figure 36.8 and

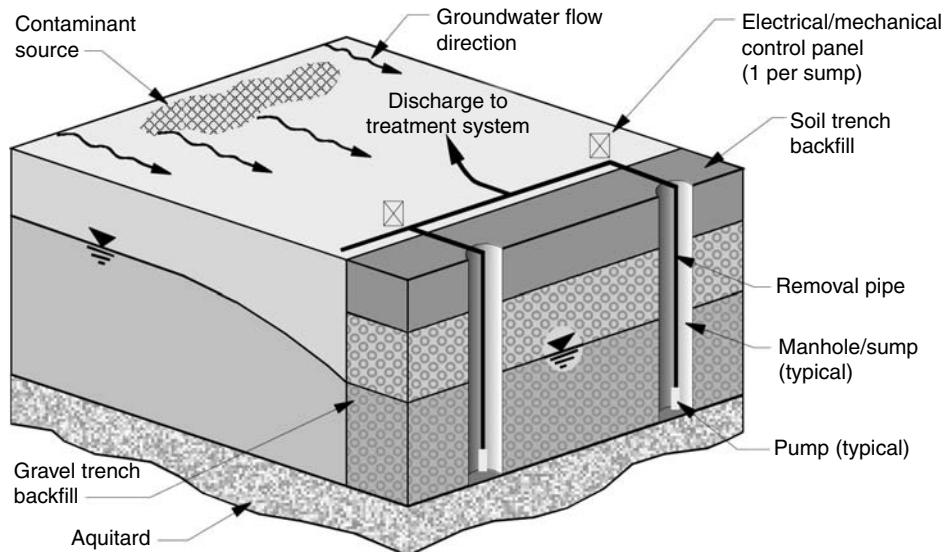


FIGURE 36.7 Typical groundwater extraction trench system.

can be estimated using the equation (Zheng et al., 1988):

$$D = \left(\frac{2S}{\pi} \right)^{1/2} \quad (36.1)$$

where S is defined by the following equation:

$$S = 2a(h_0 - h_d) \quad (36.2)$$

and D is the depth of influence of the ditch (ft), a is equal to one half the width of the ditch (ft), h_0 is the head in the aquifer beneath the ditch (ft), h_d is the head in the ditch (ft), and i is the uniform water table gradient in the aquifer. The flow rate of groundwater to the trench per unit width of trench can be estimated using the following equation:

$$Q = KiD \quad (36.3)$$

where Q is the volumetric flow to the trench (ft^3/sec), K is the hydraulic conductivity of the aquifer material, i is the uniform slope of the potentiometric surface in the aquifer, and D is the depth of influence of the trench calculated from Equation 36.3. Equation 36.1 and Equation 36.3 apply only to isotropic, homogeneous aquifers; for aquifers that contain vertical or horizontal variations in stratigraphy (e.g., clay lenses, sand seams, coarsening-with-depth deposits, etc.), or anisotropies, a more complicated model (see, for example, Chapter 18 and Chapter 19) may be needed. In such environments, the designer must be aware of the presence of such features and, if necessary, extend the depth or lateral extent of the trench to intercept contaminated groundwater that could flow through these features.

Groundwater Extraction Wells. Wells can also be installed in an aquifer to induce a hydraulic barrier to groundwater flow. In fact, because most groundwater plumes are located at depths or in locations where collection trenches are not practical, groundwater extraction wells are more commonly used. A typical groundwater extraction well is illustrated in Figure 36.9. The radius of influence and depth of influence of a single well can be calculated based on the properties of the aquifer, the geometry of the well (i.e., depth and radius), and the pumping rate of the well. Analytical solutions are provided for several combinations

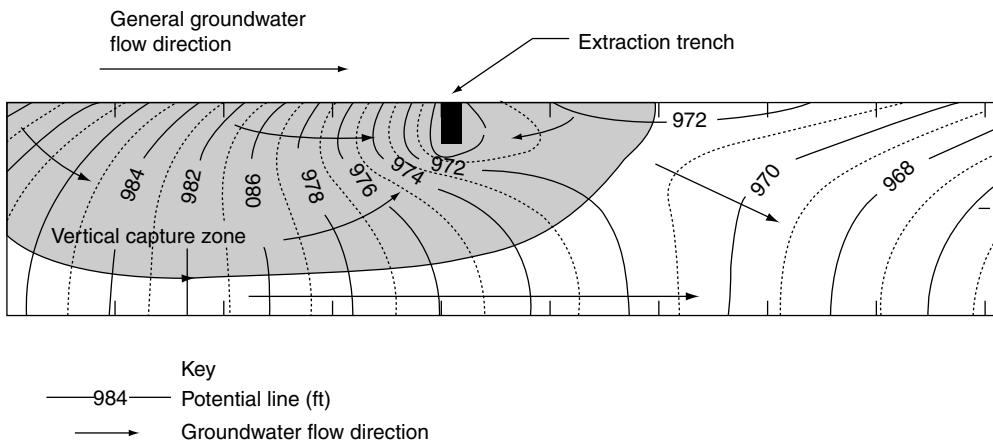


FIGURE 36.8 Vertical zone of influence of partially penetrating groundwater extraction trench.

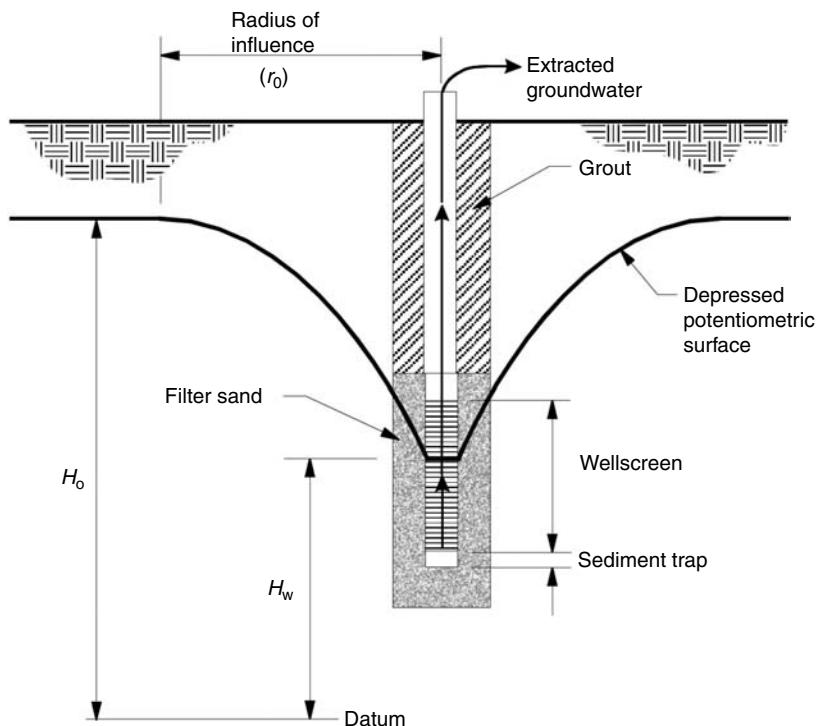
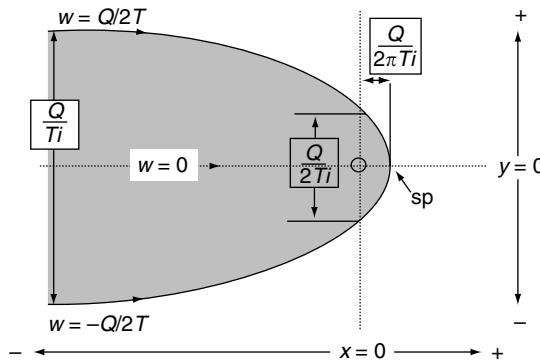


FIGURE 36.9 Groundwater extraction well.

of these variables as summarized by Cohen and Miller (1983). The zone of influence of a single well in a homogenous, isotropic, confined aquifer having a non-zero gradient is illustrated in Figure 36.10 and can be estimated as (USEPA, 1994):

$$W = \frac{Q}{Ti} \quad (36.4)$$

where W is the width of the zone of influence (ft), Q is the pumping rate (cfs) of the extraction well, T is the transmissivity of the confined aquifer (or the equivalent transmissivity of the unconfined aquifer) (sec^{-1}),



Sp = stagnation point

FIGURE 36.10 Single-well capture zone. (Data from USEPA. 1994. *Monitoring the Performance of Groundwater Pump and Treat Systems*. EPA/600/R-94/123. Washington, DC.)

and i is the hydraulic gradient (dimensionless). Zones of influence can be estimated for wells in unconfined aquifers using this equation by substituting Kb (where b is the full thickness of the unconfined aquifer at the pumping well) for T if drawdown is small compared to the thickness of the aquifer. More accurate estimates of capture zone geometry in unconfined aquifers can be made using numerical models.

The value Q in Equation 36.4 can be estimated for fully penetrating wells in an unconfined, zero gradient aquifer, where drawdown is small compared to the depth of the aquifer, as (Bear, 1979):

$$Q = \pi K \left[\frac{H_o^2 - H_w^2}{\ln(r_o/r_w)} \right] \quad (36.5)$$

where Q is the pumping rate (cfs), K is the hydraulic conductivity of the aquifer (ft/s), H_o is the hydraulic head at the limit of the zone of influence, H_w is the hydraulic head at the well (ft), r_o is the distance to the limit of the zone of influence (ft), and r_w is the radius of the pumping well screen (ft). The relationship between Q , H_o , H_w , r_o , and r_w is best estimated based on the results of a pump test, performed as described in Chapter 10 and Chapter 11.

When evaluating the combined zone of influence of a network of extraction wells, the location of the stagnation point and the overlap of adjacent zones of influence should be carefully examined. For aquifers where the hydraulic gradient is steep, the path taken by water flowing toward the network of wells could pass between and beyond the network of wells. To minimize this problem, zones of influence should be overlapped significantly (e.g., wells should be spaced at a distance of approximately $Q/(\pi T_i)$ for two extraction wells and slightly greater for increasing numbers of extraction wells [USEPA, 1994]). The capture zone of a group of extraction wells is illustrated in Figure 36.11. To perform a more accurate estimate of the dimensions of the capture zones or to more accurately evaluate the possibility of groundwater flow past the induced barrier, a particle tracking groundwater model can be used, as described in Chapter 18 and Chapter 19.

36.3.4 Other Options for Hydraulic Containment

Numerous alternative approaches to hydraulic containment have been developed over the past two decades. Some promising approaches are described below.

- *Horizontal drains* (Cohen and Miller, 1983) induce a barrier to groundwater flow in a manner similar to trenches. With recent advances in horizontal “directional” drilling technologies, the accuracy of placement for horizontal wells has improved significantly. Horizontal drains may be

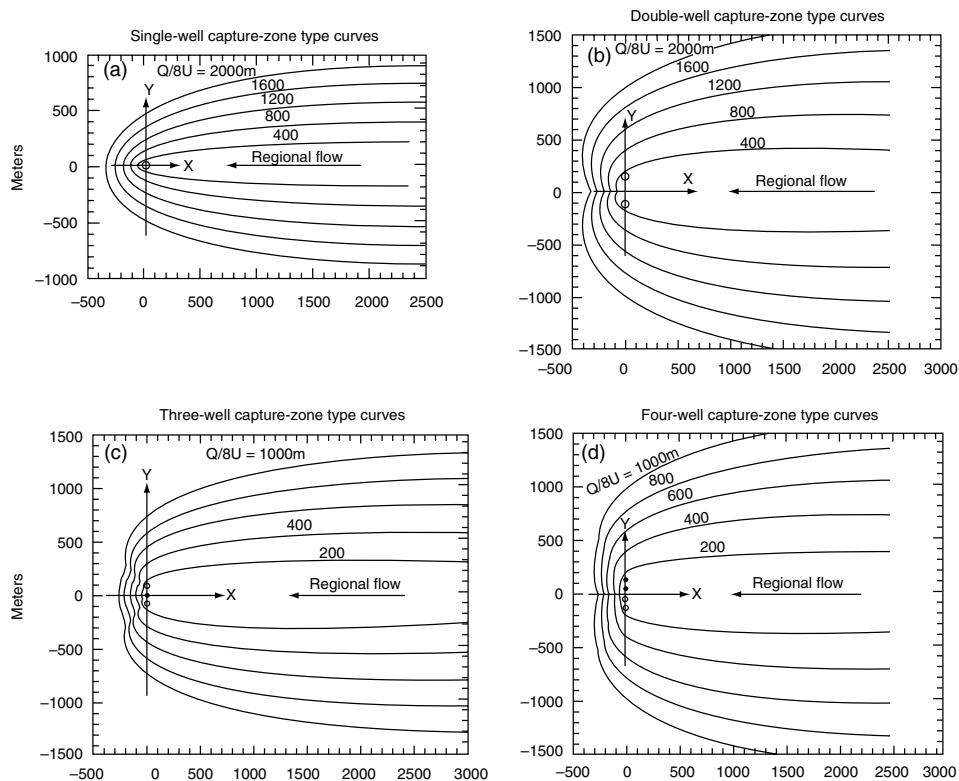


FIGURE 36.11 Capture zone of extraction well network. (From USEPA. 1994. *Monitoring the Performance of Groundwater Pump and Treat Systems*. EPA/600/R-94/123. Washington, DC.)

particularly useful in locations where access for excavation of a trench is not available (e.g., beneath a building or roadway).

- **Geomembranes** (see Rumer and Mitchell, 1995) can be incorporated into vertical barriers (as shown in Figure 36.12) to provide an extremely low-permeability barrier to groundwater flow. Geomembranes offer excellent compatibility with subsurface contaminants and have highly uniform properties. However, for the barrier to perform properly, joints between adjacent geomembrane panels must be carefully made to prevent leaks at these locations. Some commonly used installation methods for geomembrane barrier walls are summarized in Table 36.11.
- **Wellpoints** (see Driscoll, 1986) are discrete groundwater extraction devices that are typically jetted or driven into the ground. Wellpoints are useful in aquifer formations where a filter pack is not needed and where groundwater is at a relatively shallow depth. Typically, groundwater is extracted by suction or by using air-lift techniques from wellpoints; pumps are typically not used in wellpoint systems.

A good discussion of other remediation techniques is presented in USEPA (2003), and a market evaluation of these technologies is presented in USEPA (2004b).

36.4 Design of Groundwater Extraction Systems

36.4.1 Introduction

In Section 36.3, the design of groundwater extraction features was addressed, particularly groundwater extraction wells and trenches. In this section, the design of the other features of groundwater extraction

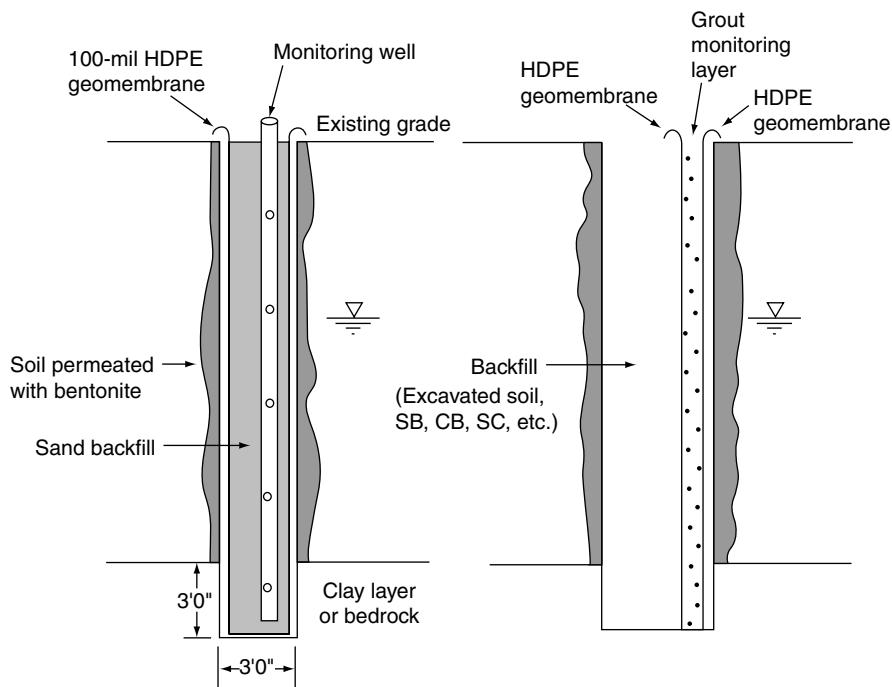


FIGURE 36.12 Geomembrane lined cutoff trench. (From Rumer, R.R. and Mitchell, J.K. 1995. *Assessment of Barrier Containment Technologies — A Comprehensive Treatment for Environmental Remediation Applications*. National Technical Information Services, Springfield, VA.)

TABLE 36.11 Installation Methods for Geomembrane Barrier Walls

Method No.	Method or technique	Geomembrane configuration	Trench support	Typical trench width mm (in.)	Typical trench depth m (ft.)	Backfill type
1	Trenching machine	Continuous	None	300–600 (12–24)	1.5–4.5 (5–15)	Sand or native soil
2	Vibrated insertion plate	Panels	None	100–15 (4–6)	1.5–6.0 (5–20)	Native soil
3	Slurry supported	Panels	Slurry	600–900 (24–36)	No limit, except for trench stability	SB, SC, CB, SCB, sand or native soil
4	Segmented trench box	Panels or continuous	None	900–1200 (36–48)	3.0–9.0 (10–30)	Sand or native soil
5	Vibrating beam	Panels	Slurry	150–220 (6–9)	No limit	SB, SC, CB, SCB slurry

Key: SB = soil bentonite; SC = soil cement; CB = cement bentonite; SCB = soil cement bentonite.

From Rumer, R.R. and Mitchell, J.K. 1995. *Assessment of Barrier Containment Technologies — A Comprehensive Treatment for Environmental Remediation Applications*. National Technical Information Services, Springfield, VA.

systems (e.g., pumps, controls, etc.) is presented. Groundwater extraction systems are a combination of both subsurface features (i.e., wells or trenches, which provide access to the contaminated groundwater) and above-ground features (which are used to regulate and monitor the extraction process). In Section 36.4.2 and Section 36.4.3, features of groundwater extraction well and trench systems are described. Then, in Section 36.4.4, methods for estimating the volume of water and time required to

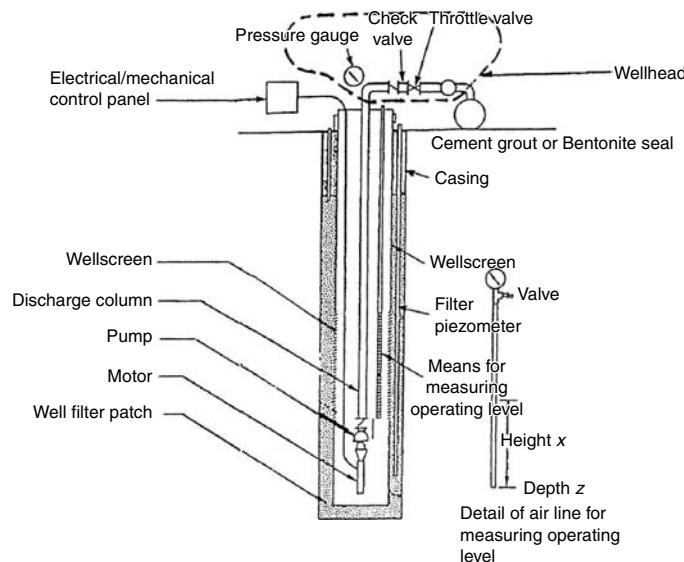


FIGURE 36.13 Typical groundwater extraction system. (From USEPA, 1995b. *Ground-Water and Leachate Treatment Systems*. EPA/636/R-94/005. Washington, DC.)

remove a plume of contaminants from the subsurface by extraction are presented and critiqued. Finally, innovative groundwater extraction approaches are described in Section 36.4.5. Note that groundwater extraction is only one step of the two-step extraction/treatment process; treatment processes are described in Section 36.5.

36.4.2 Extraction Well Systems

Groundwater extraction wells are typically more cost-effective for removing groundwater than trenches if the depth of contamination is great or if construction of the trench to the required depth in the aquifer is complex. In addition to extraction wells, extraction well systems contain the pumps, power sources, and controls that are needed to actually remove the contaminated groundwater from the aquifer and transmit it to the treatment or disposal system. A schematic diagram of a typical groundwater extraction well system is presented in Figure 36.13. The typical features of the system are briefly described below.

- **Well:** The well provides access to the groundwater in the aquifer. The well may be designed using techniques as described in Section 24.3.2.1, except the design of an extraction well should be different from the design of a groundwater monitoring well to improve the long-term performance of the extraction well and minimize the need for maintenance. Design of groundwater extraction well components (i.e., wellscreen, filter pack, etc.) is addressed by Driscoll (1986).
- **Pump and Motor:** The pump raises the groundwater from the aquifer to the above-ground wellhead assembly. The pump and motor should be selected according to the specific operating conditions that will exist during groundwater extraction; for example, some pumps are very well suited for operation under low-flow conditions, and other pumps are better suited for operation under high-flow conditions. In Table 36.12, information is presented for several types of pumps that are commonly used in groundwater extraction wells.
- **Wellhead:** The wellhead is the above-ground assembly that contains the features used to transfer groundwater from the extraction well to the transmission piping. Typical components of the wellhead assembly include valves, fittings, flowmeters, pressure gauges, and sampling ports. The wellhead can be designed to be either in an above-ground shed or in a below-ground vault, depending on the use of the area surrounding the well.

TABLE 36.12 Groundwater Extraction Pump Information

Type of pump	Practical suction lift ^a ft (m)	Usual well-pumping depth ft (m)	Usual pressure heads ft (m)	Advantages	Disadvantages
1. Positive displacement	22–36 (6.7–7.6)	22–36 (6.7–7.6)	100–2,200 (30.5–670.6)	1. Pumps water containing sand and silt 2. Especially adapted to low capacity and high lifts	1. Subject to vibration and noise 2. Maintenance cost may be high
2. Ejector				1. High capacity at low heads 2. Simple operation	1. Capacity reduces as lift increases 2. Air in suction or return line will stop
a. Shallow well	15–20 (4.6–6.1) below ejector	15–20 (4.57–6.10) below ejector	80–150 (24.4–45.7)		
b. Deep well	15–20 (4.6–6.1) below ejector	36–120 (7.62–36.58) below ejector	80–150 (24.4–45.7)	1. Same as shallow well jet 2. Well straightness not critical	1. Same as shallow well jet 2. Lower efficiency, especially at greater lifts
3. Centrifugal					
Shallow well					
a. Straight centrifugal (single stage)	20 ft (6.10) max.	(3–6.1)	100–150 (30.5–45.7)	1. Pumps water containing sand and silt 2. Usually reliable and good service life	1. Loses prime easily 2. Efficiency depends on operating under
b. Regenerative vane turbine type (single impeller)	28 (8.5) max.	28 (8.5)	100–200 (30.48–45.72)	1. Same as straight centrifugal except not suitable for pumping water containing sand or silt 2. Self-priming	1. Same as straight centrifugal except maintains priming easy 2. Efficiency depends on operating under design head and speed
Deep well					
a. Vertical line shaft turbine (multistage)	N/A (impellers submerged)	50–300 (15.2–91.4)	100–800 (30.5–243.8)	1. Same as shallow well turbine 2. All electrical components are accessible, above ground	1. Requires large, straight well 2. Repair requires pulling from well
b. Submersible turbine (multistage)	N/A (pump and motor submerged)	50–400 (15.2–121.9)	50–400 (15.2–121.9)	1. Same as shallow well turbine 2. Short pump shaft to motor 3. Quiet operation	1. Sealing of electrical equipment from water vapor critical 2. Abrasion from sand

^a Practical suction lift at sea level. Reduce lift 1 ft for each 1000 ft above sea level.

Source: From Corbitt, R.A., editor. 1990. *Standard Handbook of Environmental Engineering*. McGraw-Hill, New York. With permission.

- *Transmission Piping:* The transmission piping is used to transmit the extracted groundwater from the wellhead to the treatment or discharge location. The following should be considered when designing the piping: (1) the required quantity of flow (as described by Corbitt [1990]); (2) damage due to overburden stresses (using analysis techniques described by Merritt [1989]), including stress caused by vehicles if appropriate; (3) resistance of the piping to degradation by the contaminants in groundwater; and (4) constructability of the piping system (note that difficult-to-construct systems are less likely to be well constructed).
- *Power Source:* A source of power is typically needed for operation of the extraction system components. Power can be provided by electricity, compressed air, solar energy, or by natural gas. Local building codes should be checked to identify any specific requirements for design of electrical components of the system.
- *Electrical/Mechanical Control Panel:* The control panel contains the electrical controls (e.g., alarms, programmable logic controls, etc.) for the extraction system. Functions that are typically controlled at a panel include pump cycle time, pump temperature, monitoring for alarm conditions, and transmittal of the data collected at the wellhead (e.g., groundwater elevation, pump flow rate, etc.) to a central processing unit. Control systems often incorporate remote monitoring and/or control components that can help to optimize system O&M.
- *Discharge Point:* The discharge point is the feature at which the extracted groundwater is transferred to either the treatment system or the ultimate discharge point.

It is recommended that the following factors be considered when designing a groundwater extraction system (Driscoll, 1986):

- Because of the high quantity of groundwater that flows through an extraction well, very durable materials (such as stainless steel wellscreen) should be specified, and the open area of the well should be maximized, but carefully balanced against the strength requirements of the wellscreen.
- The chemical compatibility of each component of the extraction, transmission, and pump systems with the contaminated groundwater should be carefully examined before the final selection of the extraction system components and materials.
- The health and safety of the well installation crew should be considered when developing the plan for installation of the groundwater extraction wells (the well installer should be required to prepare a health and safety plan for well installation activities).
- The impact of naturally occurring compounds in extracted groundwater on the durability and performance of the extraction system components should be evaluated, including the effects of low and high pH (a pH significantly lower than 7 indicates acidic, possibly corrosive conditions), iron (as little as 0.36 mg/l can cause fouling of a wellscreen or groundwater treatment system), hydrogen sulfide (as little as 1 mg/l can be corrosive), dissolved solids (>1000 mg/l can cause electrical corrosion), carbon dioxide (>50 mg/l can produce carbonic acid and result in corrosive conditions), and encrusting conditions (which are typically caused by conditions of $\text{pH} > 7.5$, hardness (as CaCO_3) > 300 mg/l, iron > 0.5 mg/l, or manganese > 0.2 mg/l).

36.4.3 Extraction Trench Systems

Groundwater extraction trenches will likely be more cost-effective than extraction wells if the depth of contamination is relatively shallow and when it is relatively simple to install a trench to the required depth. Before beginning detailed design of the extraction trench system, the groundwater professional should have already identified the depth of the trench and the location where the trench must be installed (using techniques described in Section 36.3.3). In addition to an extraction trench, an extraction trench system consists of the pumps, power sources, and controls that are needed to actually remove the contaminated groundwater from the trench. A schematic diagram of a typical groundwater extraction trench system is presented in Figure 36.7. The typical features of the system are briefly described below.

- *Trench:* Trenches used for extraction of groundwater may be installed using any of a number of methods (e.g., backhoe, slurry trench backfilled with aggregate, biodegradable slurry trenches, trench machines that install the pipe in a single operation, etc. Backfill materials used in most types of groundwater extraction trenches include collection piping and aggregate; these materials should be designed to withstand degradation by the chemicals in groundwater that they will be exposed to and to resist damage caused by the stresses that they would be exposed to. Design of plastic pipes (which are typically used in extraction trenches because of their resistance to chemical degradation) is addressed by Merritt (1989).
- *Pump:* The pump should be designed using the criteria presented in Section 36.4.2 for extraction well pumps. Because groundwater extraction trenches are typically several feet wide, large-diameter sumps can be placed in the trench, which can accommodate relatively large pumps. Large pumps can be designed to be resistant to clogging by soil particles, typically have a much longer service life, and are typically more reliable and less costly to maintain than smaller pumps. A sump should be located at each low point in the trench.
- *Transmission Piping, Power Source, Control Panel, Discharge Point:* These features may be designed using the recommendations presented for extraction well systems in Section 36.4.2.

36.4.4 Time Required to Extract a Plume of Contaminated Groundwater

In this section, the concept of restoring an aquifer to its useful purpose and the reasonableness of restoration as a cleanup goal are addressed. Aquifer restoration is based on the concept that the quantity of groundwater that must be removed to restore the aquifer can be calculated based on the physical properties of the aquifer. If there were no continuing source of contamination and if contaminated groundwater flowed uniformly unimpeded to the extraction wells or trench, then the time required to "restore" the aquifer could be estimated simply by dividing: (1) the distance from the farthest part of the plume of contaminated groundwater to the extraction wells or trench (ft) by (2) the average linear groundwater flow velocity (ft/day). The volume of groundwater pumped under this scenario would be equal to the cleanup time (as calculated above) times the combined pumping rate of all of the groundwater extraction features (Figure 36.14).

Experience has shown (NRC, 1994) that, in nearly all cases, other factors affect the time and the quantity of groundwater that must be extracted to remediate a contaminated aquifer. Some of the more significant of these other factors are sorption, hydrolysis, cosolvation/ionization, volatilization/ dissolution, complexation, oxidation/reduction, and biodegradation, each of which may increase or decrease the time required

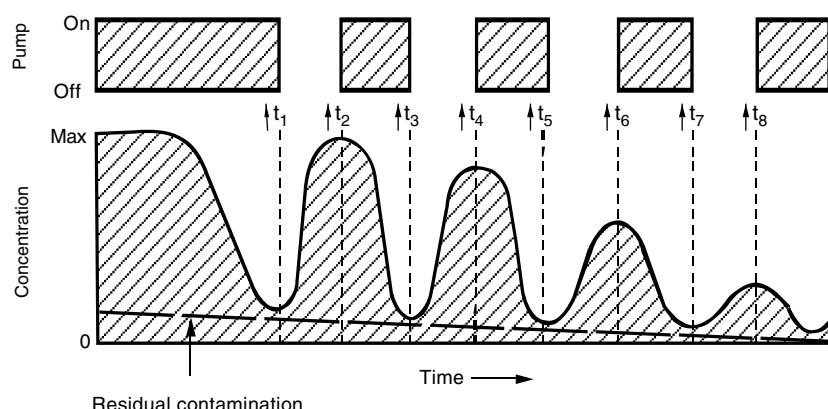


FIGURE 36.14 Use of a pulsed pumping scheme to maximize reduction of contaminant concentration. (From Keely, J.F. 1989. *Performance of Pump-and-Treat Remediations*, EPA Superfund Groundwater Issue, EPA/540/4-89/005. Washington, DC.)

TABLE 36.13 Partitioning Characteristics for Selected Chemicals

Chemical	Water solubility (mg/l)	Vapor pressure (torr)	Henry's Law constant, K_H (atm-m ³ /mol)	Organic carbon partition coefficient, K_{ow} (unitless)
Aldrin	1.7×10^{-2}	6.0×10^{-6}	1.7×10^{-4}	5.5
Benzene	1.8×10^3	9.5×10^{-1}	5.6×10^{-3}	2.1
DDT	2.5×10^{-2}	1.6×10^{-1}	1.6×10^{-6}	6.4
Tetrachloroethene (PCE)	2.0×10^2	1.9×10^1	1.8×10^{-2}	3.4
Trichloroethene (TCE)	1.5×10^3	7.3×10^1	1.0×10^{-2}	2.4
Vinyl chloride	8.8×10^3	3.0×10^3	2.7×10^{-2}	1.4

Source: From USEPA 2004c. *Superfund Chemical Data Matrix*. Washington, DC.

for remediation. A concise discussion of these processes and their impact on remediating a contaminated aquifer is presented in USEPA (1989). Sorption typically has a significant impact on remediation, although other factors could have a greater impact depending on the specific site features and the particular contaminants that are present in groundwater. Sorption is described in Chapter 18. In general, sorption is the dual process of: (1) adsorption of dissolved contaminants in groundwater to the surface of soil particles, and (2) subsequent desorption of the contaminants from the soil particles when equilibrium conditions in the aquifer change. As equilibrium conditions change in the aquifer, “partitioning” of contaminants between the dissolved phase and the solid (i.e., adsorbed) phase changes, causing contaminants to either sorb to soil particles from groundwater or to desorb from soil particles to groundwater. Contaminants become sorbed to soil particles as a result of electrochemical forces (i.e., either ion exchange reactions or van der Waals forces) between the soil particles and the dissolved contaminants. Partitioning coefficients for some common organic contaminants are presented in Table 36.13. Ions that have a high valence charge (e.g., Cr⁺⁶, Al⁺³) and molecules that are either large (e.g., chlorine-based compounds, such as TCE) or long-chained (e.g., polynuclear hydrocarbons) tend to sorb strongly to soil particles. Because desorbed contaminants can adversely impact groundwater quality, aquifer materials having sorbed contaminants act as a continuing “source” of groundwater contamination, allowing uncontaminated groundwater that flows past the sorbed contaminants to become contaminated until all of the sorbed material has been desorbed.

There are numerous models available for evaluating the time required to remediate an aquifer for which the effects of sorption are significant. The models range from simple (e.g., the batch-flush model using simple hydraulic “pore-volume” approaches, or advective-dispersive approaches) to complex (e.g., finite-element models incorporating data from bench-scale tests, pilot tests, etc.). A simple and reasonably accurate analytical method for estimating the time and quantity of extracted groundwater required to remediate an aquifer is the batch-flush model. The batch-flush model is based on the ideal assumptions that: (1) upon becoming contaminated, groundwater is perfectly mixed within the aquifer; and (2) sorption is linear, reversible, and rapid. The analysis does not account for nonhomogeneities or anisotropies within the aquifer, dispersion, or any other mechanism of contaminant transport described in Chapter 18 and Chapter 19. The batch-flush approach is based on the explicit, finite-difference approximation of the one-dimensional differential equation for groundwater transport, which is:

$$V_o \frac{dC}{dt} + \rho \frac{V_o}{n} \frac{dS}{dt} = -QC \quad (36.6)$$

where C is the concentration of the contaminant in water (mg/l), S is the sorbed-phase concentration of the contaminant (mg/l), V_o is the volume of contaminated water (i.e., the pore-volume water) (m³), ρ is the bulk density of the porous media (g/m³), n is the effective porosity (dimensionless), Q is the volumetric groundwater extraction rate (m³/sec), and t is time (sec). Because sorption is linear, reversible, and rapid,

S can be substituted with K_{dC1} , where K_d is the soil-water partitioning coefficient. Finally, substituting and rearranging gives the equations:

$$T_t = R \left[-\ln \frac{C_t}{C_o} \right] \quad (36.7)$$

$$t_t = \frac{V_o}{Q} T_t \quad (36.8)$$

where C_o is the initial concentration of contaminant in the aquifer, T_t is the number of pore volumes that must be extracted to “restore” the aquifer to the desired cleanup concentration (C_t), and t_t is the time required to achieve C_t . In Equation 36.9, the term R is known as the retardation factor and is defined as:

$$R = 1 + \frac{\rho}{n} K_d \quad (36.9)$$

where ρ , n , and K_d are as previously defined. Values of K_d may be measured in the laboratory; however, because the value of K_d is primarily a function of sorption to organic material, a useful approximation of K_d can be found using the equation:

$$K_d = K_{oc} \times f_{oc} \quad (36.10)$$

where K_{oc} is a proportionality constant (which can be estimated based on the octanol-water partitioning coefficient, K_{ow} , of the contaminant) and f_{oc} is the fraction of organic carbon in the aquifer soils. Estimates of K_{oc} for several common organic contaminants are provided in Table 36.13. Because desorption is a function of the ratio of the concentration of sorbed contaminant to the concentration of dissolved contaminant, sorbed contaminants could be expected to desorb indefinitely, which (depending on the desired cleanup concentration, C_t) could make it impossible to restore groundwater quality in a reasonable period of time, if at all.

Batch flush models are useful for estimating cleanup times for simple aquifer systems in which sorption is nearly linear. However, the model typically underestimates cleanup times because it does not account for any of the following: (1) heterogeneities; (2) continuous releases from a source of contamination; (3) the presence of separate-phase contaminants; or (4) the effects of nonlinear sorption. These limitations also apply to models that are more complex than the batch flush model. Because of the effects of sorption, the cost to clean up an aquifer to a very low level of contamination can be significantly greater than the cost to achieve a slightly higher level of contamination; an example of this problem is shown in Figure 36.15 and in Table 36.14, which illustrates groundwater quality as a function of the number of pore volumes removed for a simple pump-and-treat groundwater remediation system in a homogeneous sand aquifer. Remediation cost can be expected to increase substantially as the retardation factor increases (as shown in Figure 36.16), and a substantial additional cost can be expected for achieving very high cleanup efficiencies (as shown in Figure 36.17 and Table 36.14). If it is found that the groundwater remediation goal cannot be achieved through extraction, then either a different groundwater remediation goal should be selected or a complementing remediation technique should be considered to enhance remediation of groundwater.

36.4.5 Other Groundwater Extraction Approaches

There are many groundwater extraction approaches available to the groundwater professional in addition to vertical wells and trenches, including both innovative approaches for removing groundwater and approaches that enhance the extraction of contaminants from the subsurface. Other groundwater extraction techniques include the use of horizontal wells and wellpoints. Some of the more promising innovative approaches to enhance extraction efficiency include pneumatic hydraulic fracturing

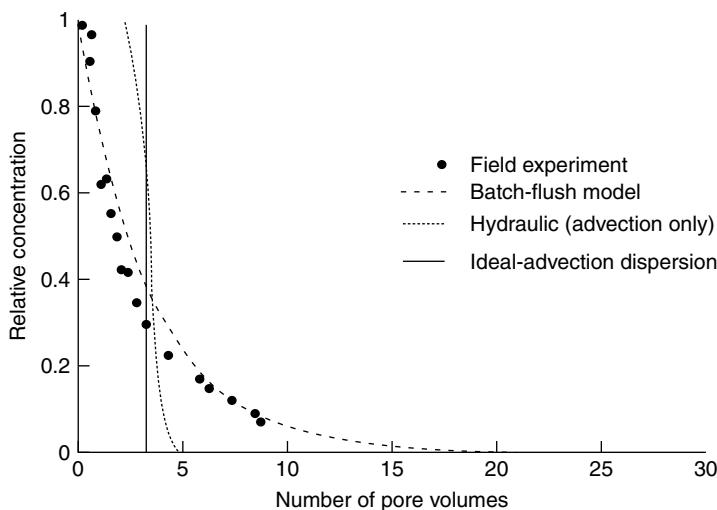


FIGURE 36.15 Elution curve for diethylether. (From Brusseau, M. 1996. Evaluation of Simple Methods of Estimating Contaminant Removal by Flushing. *Groundwater*, p. 19–22. With permission.)

TABLE 36.14 Impact of Cleanup Goal on Cost of a Conventional Pump-and-Treat System

Percent removal required	Calculated years to achieve goal	Present worth
80	15	\$ 2,800,000
90	21	\$ 3,360,000
99	42	\$ 4,750,000
99.9	63	\$ 5,600,000
99.99	84	\$ 6,000,000

Source: Adapted from National Research Council (NRC), Water Science and Technology Board. 1994. *Alternatives for Groundwater Cleanup*. Washington, DC. With permission.

(i.e., fracturing subsurface media using water pressure or explosives to increase the permeability of the media and the groundwater extraction efficiency), flushing (i.e., reinjecting treated groundwater upgradient of the source area to increase the hydraulic gradient toward the extraction feature), surfactant flushing or chemical extraction (i.e., flushing using surfactants or other chemicals that are designed to increase the mobility or solubility of the contaminants in the aquifer), electroacoustical decontamination (i.e., use of electrokinetic and acoustical waves to reduce surface tension and viscosity of separate-phase contaminants, increasing their mobility and the opportunity to extract them), and steam extraction (i.e., injection of steam into the aquifer to either decrease the viscosity of a contaminant or to volatilize contaminants). When evaluating reinjection techniques that involve the use of chemicals, the groundwater professional should be careful to evaluate the potential adverse impact of the chemicals, which may be toxic, on human health and the environment.

When evaluating groundwater extraction approaches, separate attention should be given to extraction of NAPLs. Because they represent a significant source of potential contamination, removal of NAPLs can significantly decrease the amount of time required to achieve a groundwater remediation goal using pump-and-treat techniques. Dual-phase pumping approaches can be effective for removing both LNAPL and DNAPL products; examples of dual-extraction systems for LNAPL and DNAPL products are illustrated in Figure 36.18.

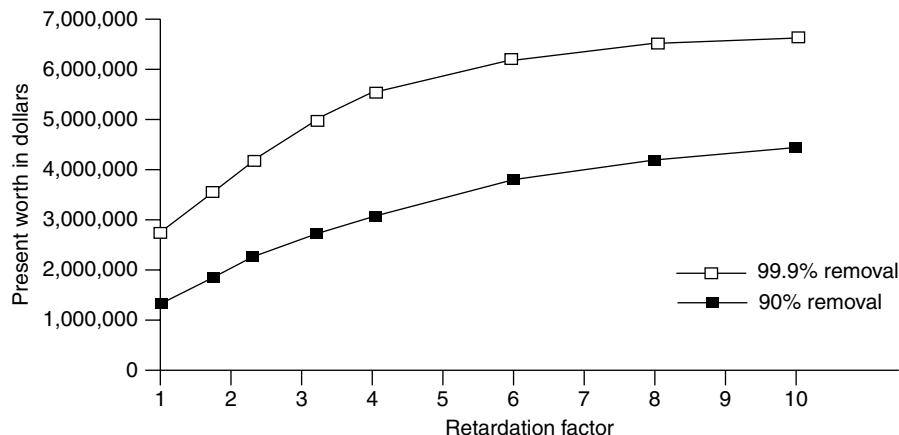


FIGURE 36.16 Cost of pump-and-treat system as a function of the retardation factor of the contaminant. (Adapted from National Research Council (NRC), Water Science and Technology Board. 1994. *Alternatives for Groundwater Cleanup*. Washington, DC. With permission.)

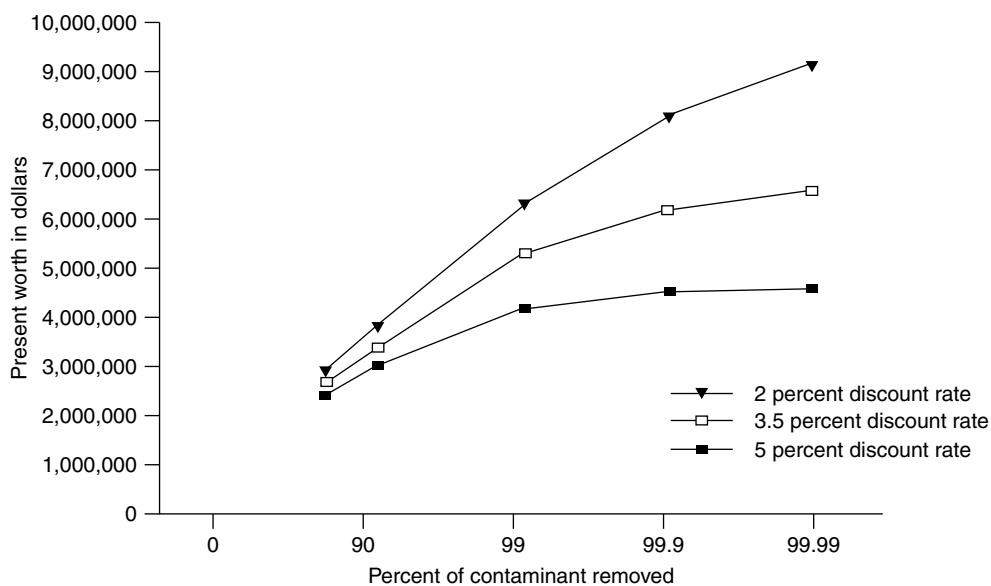


FIGURE 36.17 Cost of pump-and-treat system as a function of the cleanup goal. (Adapted from National Research Council (NRC), Water Science and Technology Board. 1994. *Alternatives for Groundwater Cleanup*. Washington, DC. With permission.)

36.5 Treatment of Contaminated Groundwater

36.5.1 Introduction

There are many techniques available for treatment of contaminated groundwater. Over the past 20 years, a large number of groundwater treatment techniques has been developed; these techniques are described by the Federal Remediation Technology Roundtable (2002), USEPA (1995c), and GAO (2005). In this section, brief descriptions of the most widely used standard treatment techniques and the most promising innovative treatment techniques are presented. Candidate treatment technologies that are addressed in this

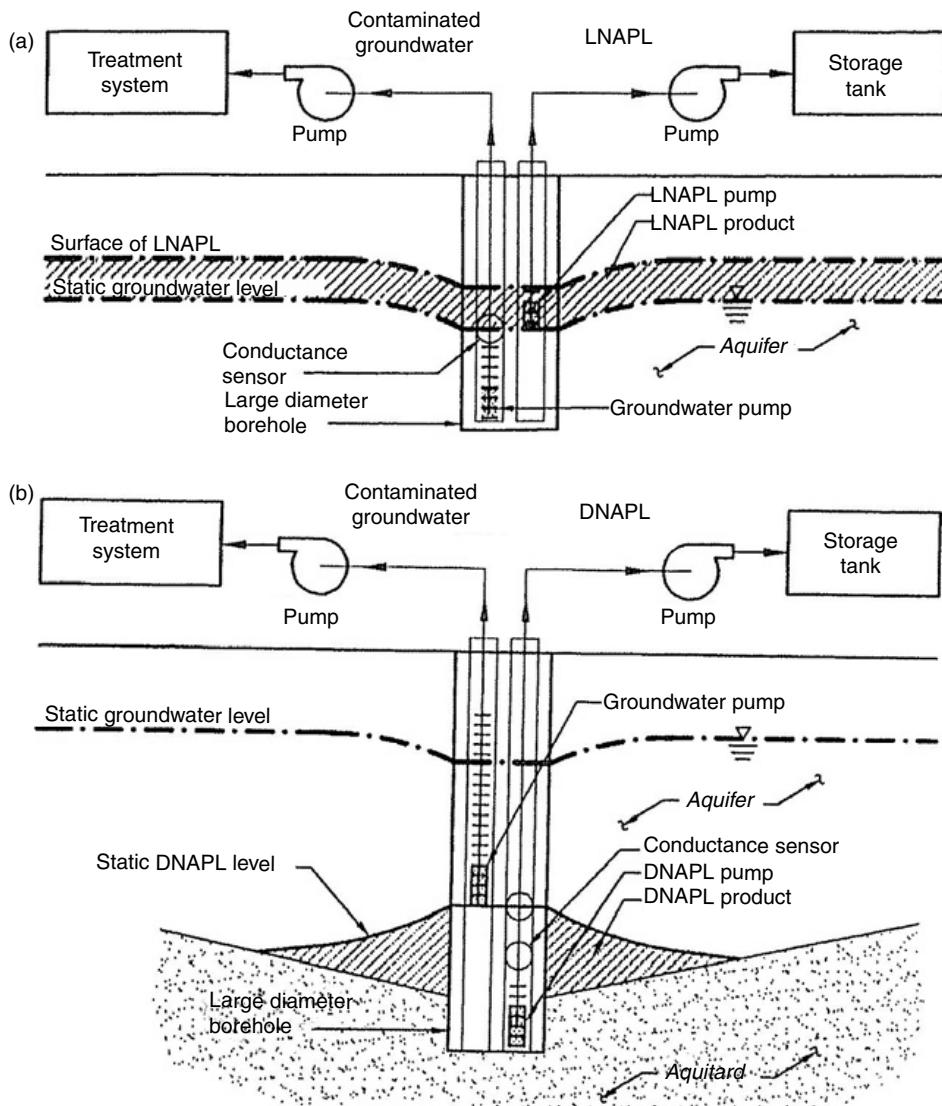


FIGURE 36.18 (a) LNAPL Dual-extraction system; (b) DNAPL Dual-extraction system.

section are presented in Table 36.15. For the purposes of this section, these technologies are categorized as either *in situ* or *ex situ* techniques. *In situ* techniques are generally intended to either render a contaminant nontoxic through treatment (e.g., bioremediation) or to enhance extraction of the contaminant from the aquifer (e.g., air sparging). *Ex situ* techniques can be used only to treat groundwater that has been extracted from the aquifer. *In situ* treatment techniques are described in Section 36.5.2, and *ex situ* treatment techniques are described in Section 36.5.3.

36.5.2 *In Situ* Treatment

36.5.2.1 Introduction

In situ treatment techniques are described in this section in three categories. First, techniques that incorporate biological processes (e.g., bioventing, bioremediation) are described in Section 36.5.2.2. Next, techniques that incorporate volatilization processes (e.g., air sparging, soil vapor extraction) are described

TABLE 36.15 Summary of Candidate Treatment Technologies

	Treatment type	Process description
<i>In Situ</i>		
Bioremediation	Destruction	Biological degradation of contaminants by stimulation of naturally occurring microbes in soil or non-native (exogenous) cultures introduced to target-specific contaminants
Soil vapor extraction	Destruction	Volatilization of contaminants that are present in the vadose zone
Air sparging	Extraction	Volatilization of contaminants in the saturated zone
Permeable reaction barriers	Destruction	Physical or chemical treatment in a trench
Chemical Oxidation/Reduction	Destruction	Chemical degradation of contaminants using injected oxidants (e.g., permanganate) or reducing agents (e.g., nanoscale zero-valent iron)
<i>Ex Situ</i>		
Bioreactor	Destruction	Biological degradation of contaminants (activated sludge, fixed-film biological reactor, biophysical treatment or slurry phase treatment)
Air stripping	Extraction	Volatilization of contaminants
Carbon adsorption	Extraction	Adsorption of contaminants to activated carbon
Ion exchange	Extraction	Exchange-type attachment of contaminants to ion-exchange resin
Chemical precipitation	Extraction	Alteration of water quality (usually pH adjustment) to conditions in which concentration exceeds the solubility limit of the compound, causing precipitation
Membrane Separation	Extraction	Separation of solids from water using membranes (reverse osmosis, ultrafiltration)
Wetlands treatment	Destruction/Extraction	Uptake of contaminants by wetland features
Electrokinetic decontamination	Destruction	Desorption of contaminants by “acidic front” of groundwater caused by hydrolysis of the groundwater
Advanced oxidation process	Destruction	Destruction of contaminants by contacting them with oxidizing environments typically containing free hydroxyl radicals

in Section 36.5.2.3. Then, techniques that incorporate a chemical or physical process (e.g., permeable reaction barriers) are described in Section 36.5.2.4. Finally, some promising innovative *in situ* treatment approaches are described in Section 36.5.2.5. In some cases, *in situ* treatment technologies may offer a viable alternative to more conventional remediation approaches (e.g., pump-and-treat) and may improve the chances of achieving the ultimate remediation goal (in some cases, even aquifer restoration).

36.5.2.2 Biological Remediation

In situ biological treatment of groundwater — commonly referred to as enhanced *in situ* bioremediation (EISB) — involves creating conditions in the subsurface that promote the growth of microorganisms that can degrade contaminants. The popularity and use of EISB for groundwater remediation increased significantly during the 1990s, and today EISB is transitioning from a developing technology to mature technology. Technical guidance for designing and applying EISB is available in a variety of government reports, including USEPA, 2000c; Hughes et al., 2002; ITRC, 2002; and USAFCEE, 2004. Many of these reports are available on-line (see links to world-wide-web provided in the reference list). In this section, a brief overview of biological remediation is presented; a more detailed discussion of bioremediation is presented in Chapter 31 (i.e., Bioremediation).

EISB exploits the capabilities of naturally occurring subsurface microorganisms, which can derive energy and often reproduce while metabolizing contaminants. Biodegradation reactions may involve destruction of organic contaminants, such as fuel hydrocarbons and chlorinated solvents, but can also involve transformation of inorganic contaminants to less toxic or immobile forms (e.g., nitrate to nitrogen gas or hexavalent chromium to trivalent chromium). EISB involves the perfusion of treatment agents that stimulate biological growth and catalyze biodegradation reactions. Bioremediation agents can include edible organic substrates (e.g., soybean oil, ethanol, molasses), air, oxygen gas, hydrogen gas, nutrients

(nitrogen and phosphorus), wood mulch, specialized bacterial cultures, and vitamins. These agents can be delivered via batch-type injections into the subsurface, or via groundwater recirculation type designs.

Numerous and diverse types of bacteria are naturally present in the subsurface, and many of these bacteria possess the capability to degrade common groundwater contaminants (e.g., fuel hydrocarbons or nitrate). Many organic contaminants (e.g., polycyclic aromatic hydrocarbons and fuel hydrocarbons) are degraded by the process of heterotrophic metabolism, in which bacteria use the contaminant as an electron donor and gain energy for maintenance and growth via microbial respiration. The by-products of complete biodegradation include carbon dioxide, water, and biomass. Biodegradation of synthetic organic contaminants in the subsurface may occur in a series of several steps, in which a contaminant is degraded into intermediate products before it is completely degraded. For example, under certain aquifer conditions, tetrachloroethylene is biodegraded to ethene and ethane by the following sequential steps: tetrachloroethylene => trichloroethene => *cis*-1, 2-dichloroethene => vinylchloride => ethene => ethane.

Bioremediation is often most successful in aquifers that are contaminated with chemicals that are easily metabolized by the indigenous microorganisms; in general, the more closely a contaminant resembles a naturally occurring compound, the more likely it is that there exists a microorganism in the aquifer that is capable of biodegrading the contaminant. Fuel hydrocarbons such as benzene and naphthalene, for example, are naturally occurring compounds that can be easily biodegraded under aerobic conditions by many subsurface bacteria. Certain synthetic organic compounds, such as chlorinated ethenes or chlorofluorocarbons, are more difficult to biodegrade, and there are fewer species of bacteria capable of completely biodegrading these compounds. An illustration of the relative biodegradability of organic contaminants under aerobic conditions is shown in Figure 36.19. Note that while some inorganic compounds are easily biodegraded (e.g., nitrate=>nitrogen gas), many inorganic compounds (e.g., heavy metals) cannot be biodegraded. Nevertheless, EISB systems can be used to engineer favorable biological treatment reactions, such as the reductive precipitation of hexavalent chromium and hexavalent uranium to precipitated end products. Also, note that the ambient conditions within the aquifer have a significant effect on the potential for successful bioremediation; a summary of the environmental factors that affect bioremediation and the range of optimum values for each factor are presented in Table 36.16.

The rate of biodegradation (and, generally, the consumption of any material at relatively low concentration by biological activity) is often expressed by the first-order rate equation defined below:

$$\ln \left(\frac{C_t}{C_o} \right) = -kt \quad (36.11)$$

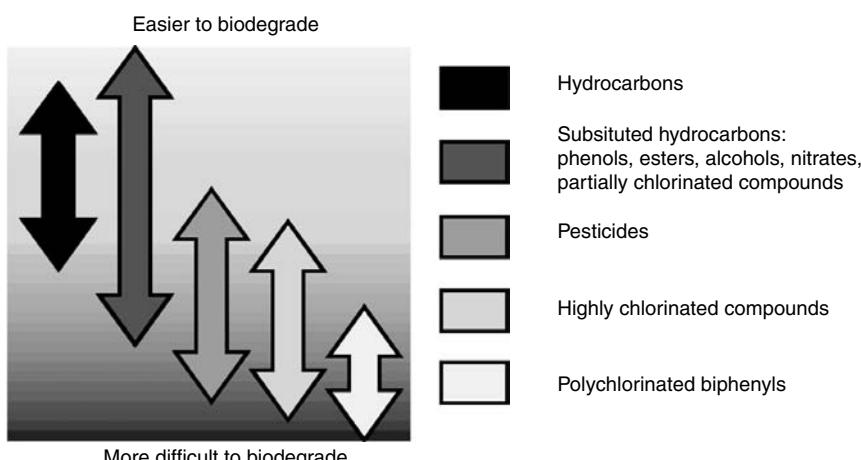


FIGURE 36.19 Relative biodegradability of organic compounds.

TABLE 36.16 Environmental Factors That Affect Biodegradability

Environmental factor	Optimum levels
Available soil water	36–85% of water holding capacity; -0.01 Mpa
Oxygen	Aerobic metabolism: >0.2 mg/l dissolved oxygen, minimum air-filled pore space of 1% by volume Anaerobic metabolism: oxygen concentrations <1% by volume
Redox potential	Facultative anaerobes: >50 mV
pH	Obligate anaerobes: <50 mV pH values of 5.5–8.5
Nutrients	Sufficient nitrogen, phosphorus, and other nutrients so as to not limit microbial growth (Suggested C:N:P ratio of 120:10:1)
Temperature	15–45°C (Mesophiles)

Source: Modified from USEPA. 1991. *Handbook — Ground Water, Volume II: Methodology*. EPA/636/690/016b. Washington, DC.

where k is the first-order reaction rate coefficient (sec^{-1}), C_t is the concentration of constituent remaining at time t (mg/l), C_0 is the initial concentration of the constituent (mg/l), and t is time (sec). Biodegradation rate coefficients are often estimated based on the results of bench-scale testing performed prior to field implementation. In the case of pre-design analyses, the time required for degradation of a particular plume of contamination can be estimated by solving Equation 36.11 for t . In evaluating the field performance of EISB systems, the rate coefficient can be estimated by comparing contaminant concentrations over time and space relative to the initial concentration (C_0). It should be noted that, at high contaminant concentrations, biodegradation reaction rates are often more accurately described by variable- or zero-order (Monod) kinetics. In the case of contaminant source areas or high-concentration plumes, use of first-order rate models to quantify biodegradation rates can significantly underestimate the time needed to achieve cleanup criteria.

The design effort for a bioremediation system is typically focused on (1) identifying the additives (e.g., organic substrates, nutrients, oxygen, and microorganisms, etc.) that will promote biodegradation; and (2) determining effective approaches for delivering those additives to the subsurface. The most effective EISB systems are typically those that have been designed via the traditional engineering scale-up process of bench testing, followed by pilot testing, followed by full-scale application (EPA, 2000a). Laboratory bench testing is a simple and low-cost step for identifying effective treatment additives and assessing treatment rates for design of full-scale systems. Bench testing, as well as pilot-scale testing, is a prudent approach for determining effective additives, identifying potential inhibitory factors, and avoiding wasteful expenditures on unsuccessful full-scale designs in the field.

36.5.2.3 Volatilization Processes

Volatilization is the transfer of a chemical from the liquid state to the gaseous state. Remediation techniques that employ volatilization processes promote the change of contaminants from liquid phase (including dissolved) to the vapor phase. Vapor-phase contaminants are typically easier to remediate than liquid-phase contaminants because aquifer materials are significantly more permeable to vapors than to liquids, making vapors easier to remove from the aquifer (see discussion of intrinsic permeability in Chapter 2). Remediation technologies that employ volatilization processes include both soil vapor extraction (which is applicable to contaminants adsorbed to soil in the vadose zone) and air sparging (which is applicable to both dissolved and adsorbed contaminants in the saturated zone). Both processes are based on the principle of molecular diffusion of contaminants from the dissolved or nonaqueous phase to the vapor phase, which is described by Fick's Law for diffusion. The principles of soil vapor extraction (SVE) and air sparging (AS) are described below.

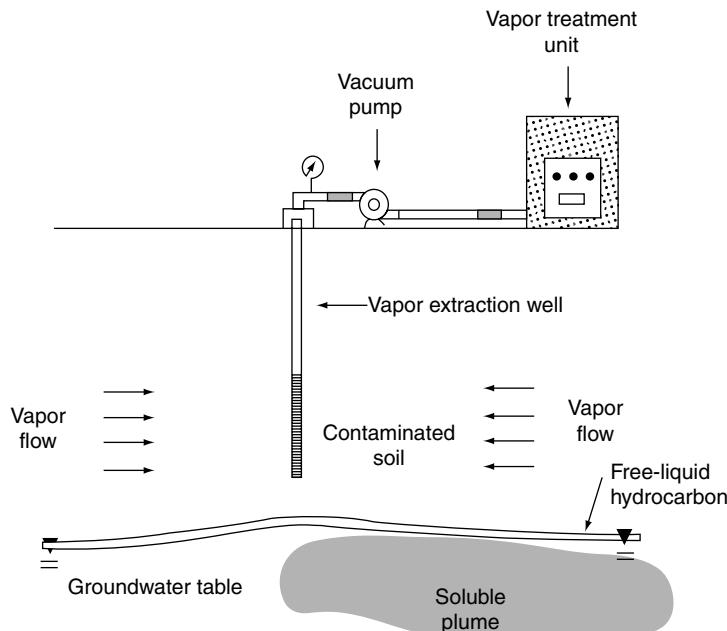


FIGURE 36.20 “Basic” soil vapor extraction system. (From Johnson, P.C., Stanley, C.C., Kembowski, M.W., Byers, D.I., and Colhart, J.D. 1990. A practical approach to the design, operation, and monitoring of *in situ* soil venting systems. *Groundwater Monitoring Review*, Spring 1990, 159–178. With permission.)

Soil Vapor Extraction. An example of an SVE system is illustrated in Figure 36.20. The concentration of a particular contaminant in the extracted vapor can be estimated using Fick’s Law as follows:

$$C_{\text{est}} = \sum_i \frac{x_i P_i^{\gamma} M_w i}{RT} \quad (36.12)$$

where C_{est} is the estimated concentration of contaminant in the vapor (mg/l), x_i is the mole fraction of component i in the liquid-phase contaminants, P_i^{γ} is the vapor pressure of component i at temperature T (atm), M_w is the molecular weight of component i (mg/mole), R is the gas constant (i.e., 0.0821 l-atm/mole °K), and T is the absolute temperature of the contaminant residual (K). Note that contaminant concentrations will decline with time as the concentration of the contaminant adsorbed to the subsurface soils decreases; therefore, the contaminant vapor concentration will change over time. SVE is potentially suitable for contaminants that have a vapor pressure conducive to volatilization (usually greater than about 1×10^{-4} atm at the average subsurface temperature); a summary of vapor pressures for some typical contaminants is presented in Table 36.13. Using the value of C_{est} , the removal rate can be estimated by multiplying C_{est} by the expected air flow rate Q . The achievable air flow rate can be estimated using the following equation (Johnson et al., 1990):

$$\frac{Q}{H} = \pi \frac{k}{\mu} P_w \frac{[1 - (P_{\text{atm}}/P_w)^2]}{\ln(R_w/R_i)} \quad (36.13)$$

where Q/H is the flow rate per unit thickness of wellscreen (cm^3/sec), k is the soil permeability to air flow (cm^2), μ is the viscosity of air (1.8×10^{-4} g/cm-s, or 0.018 cp), P_w is the absolute pressure at the extraction well (atm), P_{atm} the absolute atmospheric ambient pressure (i.e., 1.01×10^6 g/cm-s 2 , or 1 atm), R_w is the radius of the extraction well (cm), and R_i is the radius of influence of the extraction well (cm). The reasonableness of the expected gas flow rates for the permeability of the site soils

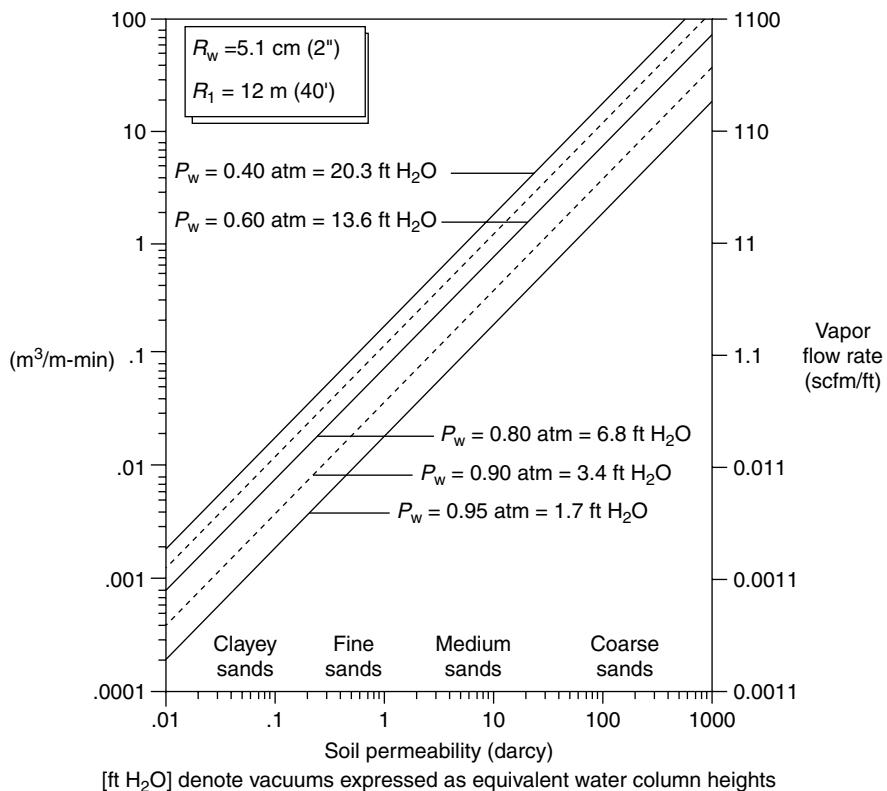


FIGURE 36.21 Predicted steady-state flow rates (per unit wellscreen thickness) for a range of soil permeabilities and applied vacuums (P_w). (From Johnson, P.C., Stanley, C.C., Kembrowski, M.W., Byers, D.I., and Colthart, J.D. 1990. A practical approach to the design, operation, and monitoring of *in situ* soil venting systems. *Groundwater Monitoring Review*, Spring 1990, 159–178. With permission.)

may be confirmed using Figure 36.21. In some cases, there may be a significant quantity of residual contamination adsorbed to soil particles located a short depth below the water table as a result of changes in groundwater elevation since contaminants were released. In this case, it may be efficient to implement a “dual extraction” (i.e., air and groundwater) system similar to the dual extraction system illustrated in Figure 36.18A for groundwater and LNAPLs. By depressing the groundwater table below the residual contamination, the residual contamination can be extracted using SVE techniques, which are much more efficient than groundwater extraction techniques for removing adsorbed or residual contamination.

Air Sparging/Biosparging/Bioventing. An illustration of an air-sparging (AS) system is presented in Figure 36.22. Air sparging involves introducing air into the aquifer to volatilize and biodegrade groundwater contaminants. To prevent uncontrolled migration of contaminated sparge air, AS systems are typically used in conjunction with SVE systems, which capture the sparged air and volatilized organic compounds after they rise from groundwater into the vadose zone, as shown in Figure 36.22. Biosparging is a form of AS that involves injection of air into contaminated groundwater at low flow rates (e.g., 2 to 5 cfm), to maximize *in situ* biodegradation and minimize volatilization. Similarly, bioventing involves injection of air and, in some cases, gaseous nutrients) above the water table to stimulate biodegradation of vapor- and sorbed-phase contaminants in the vadose zone. The US Department of Defense has applied AS, biosparging, and bioventing technologies at hundreds of sites, and achieved success particularly at fuel spill sites. Useful guidance manuals for these technologies include (Leeson and Hinchee, 1996a and 1996b; and Leeson et al., 2002).

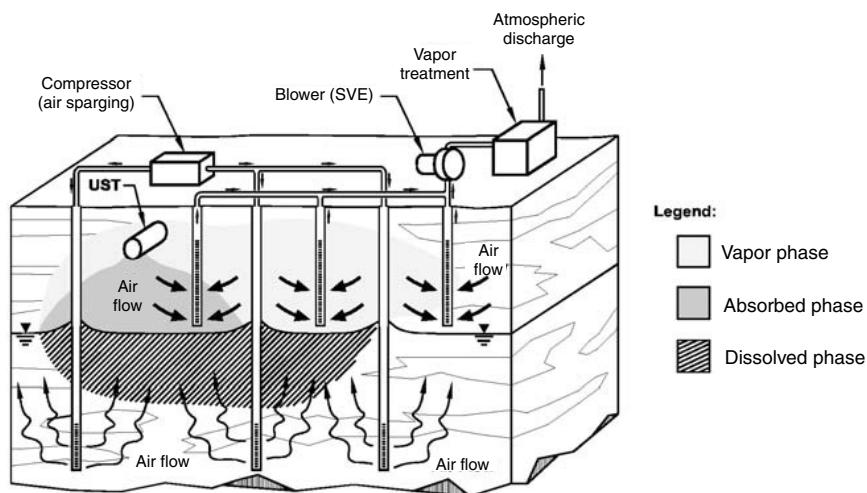


FIGURE 36.22 Typical air-sparging system configuration. (From Hinchee, R.E. 1994. *Air Sparging for Site Remediation*. Lewis Publishers, Chelsea, Michigan.)

In general, the physics of air movement in the saturated zone is currently not well understood, which limits the predictability of AS system performance; in fact, the decisions regarding sparge well location and spacing are usually based on experience rather than calculation. Still, because it has a relatively low cost, AS can be a costeffective approach to groundwater remediation. As presently understood, air discharged to groundwater by a sparge well typically flows upward to the vadose zone through air channels (as illustrated in Figure 36.22) that are established at first operation of the system. These air channels directly impact only a very small quantity of groundwater and soil. Contaminants partition between the sparge air, groundwater, and soils; the portion that partitions to vapor in the sparge air rises with the sparge air to the ground surface. For AS to be effective, adsorbed contaminant mass must be transported by diffusion from the surface of soils through the groundwater to the air–water interface of the air channels, which is a slow process.

The minimum air pressure required to operate a sparge system is provided by the equation (Johnson et al., 1993):

$$P_H = r_w g (L_s - L_{gw}) \quad (36.14)$$

where P_H is the hydrostatic pressure corresponding to the water column height that is displaced (atm), r_w is the density of water (1000 kg/m^3), g is the acceleration due to gravity (9.8 m/sec^2), L_s is the depth below ground to the top of the sparge wellscreen (m), and L_{gw} is the depth to groundwater (m).

Two concerns exist related to the performance of air sparge systems, both of which can be addressed through monitoring. First, the introduction of air below the water table could raise the elevation of the water table, possibly changing groundwater flow direction. This first concern can be addressed by monitoring groundwater elevations near the sparge area to confirm that groundwater elevations are not changing. The second concern is that sparge gas may migrate horizontally in more permeable zones resulting in an increase in the lateral extent of contamination. This concern can be addressed by monitoring the chemical characteristics of soil vapor at locations beyond the sparge wells to verify that contaminated vapors are not migrating offsite in an uncontrolled manner and, if necessary, by using SVE to prevent offsite migration of vapors.

In Situ Thermal Treatment. Beginning in the 1980s, *in situ* thermal remediation technologies were developed and have more recently been employed to enhance the mobility and extraction of volatile

and semi-volatile contaminants from groundwater. Variations of *in situ* thermal treatment technologies include steam injection, resistive heating, *in situ* thermal desorption, hot air injection, and radio frequency (RF) heating. A useful summary of *in situ* thermal treatment technologies is provided by EPA (2004d). Each of the variations of *in situ* thermal treatment functions by increasing subsurface temperatures, thereby increasing the mobility of subsurface contaminants and facilitating their removal from the subsurface. Therefore, they typically are used in conjunction with extraction technologies, primarily SVE. The mechanisms by which *in situ* thermal treatment technologies enhance mobilization, even within highly contaminated source areas, for less volatile contaminants, or in low permeability aquifers, have been described by USEPA (Davis, 1997; USEPA, 2004). In addition to increasing mobility, some applications of *in situ* technologies have the added effect of increasing subsurface temperatures to an extreme where contaminants may be thermally destroyed. One potential concern with *in situ* thermal treatment technologies is the potential for spreading of contaminant plumes if appropriate vapor or groundwater capture systems are not employed.

36.5.2.4 Chemical and Physical Processes

Chemical and physical processes are often used to promote *in situ* treatment of contaminated groundwater. The range of chemical processes that can be applied *in situ* includes reduction with chemical reducing agents (e.g., sodium dithionite), reduction with zero-valent iron, and chemical oxidation with permanganate. For a chemical treatment process to be carried out to a high rate of completion, groundwater must be well mixed with reagents, which can be difficult to achieve in heterogeneous formations, low-permeability formations, and large plumes. Still, under some conditions, *in situ* chemical and physical treatment of contaminated groundwater is possible, particularly in situations where contaminated groundwater can be routed to a collection point before it is treated and where contaminated groundwater can be well mixed with treatment reagents.

Chemical processes that can be considered for *in situ* application include adsorption, precipitation, oxidation/reduction, fixation, and physical transformation. Such processes can be promoted by: (1) installing a permeable reaction barrier (also known as a “reaction trench,” or “*in situ* reactive barrier,” or “selective” barrier) in the aquifer to intercept contaminated groundwater and provide a controlled environment for the reaction; or (2) by applying reagents directly to groundwater through wells or by infiltration from the ground surface. The barrier can be constructed using the techniques described in Section 36.4.3 and can be backfilled with reagents, such as activated carbon (for treatment of volatile organic contaminants), limestone (to promote precipitation of dissolved metal ions), or iron filings (to capture organic carbon compounds [Focht et al., 1996]), to provide the desired treatment of groundwater. A variation of the permeable reaction barrier is the “funnel and gate” reaction barrier, which is illustrated in Figure 36.23. This approach uses relatively impermeable barrier walls (i.e., the “funnels”) to route groundwater to the permeable, reactive portion of the barrier (i.e., the “gate”) and minimizes the length of barrier that must be maintained or reconditioned as the reaction sites in the barrier wall become exhausted.

The range of treatment technologies that can be used for permeable reaction barriers is illustrated in Table 36.17. The primary advantages of permeable reaction barriers are a result of the passive nature of the technology; for example, there is no need for a power source, which typically represents the majority of long-term operation and maintenance costs for active (e.g., pump-and-treat) remediation systems. There is a large variety of technical guidance in the public domain regarding design and application of *in situ* chemical treatment technologies and permeable reactive barriers. Example guidance regarding zero-valent iron reactive barriers includes Blowes et al. (1999a, 1999b) and Blowes and Mayer (1999); example guidance regarding chemical oxidation is provided in USEPA (1998a); ITRC (2005); Yin and Allen (1999).

36.5.2.5 Monitored Natural Attenuation

Monitored natural attenuation (also known as “MNA”) involves the evaluation and monitoring of naturally occurring processes that prevent the migration of contaminants to receptors. Currently, MNA is most frequently used to describe biodegradation of contaminants by naturally occurring microorganisms;

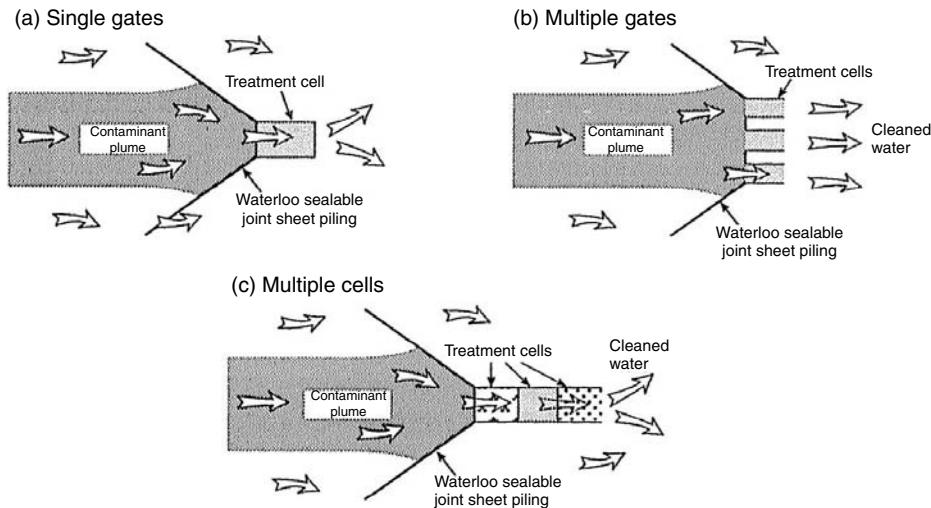


FIGURE 36.23 Funnel and gate system configuration. (From Rumer, R.R. and Mitchell, J.K. 1995. *Assessment of Barrier Containment Technologies—A Comprehensive Treatment for Environmental Remediation Applications*. National Technical Information Services, Springfield, VA.)

TABLE 36.17 Treatment Technologies Applied in Permeable Reactive Barriers

Treatment medium	Target contaminants	Technology status
Zero-valent iron	Halocarbons	Commercially applied
Zero-valent iron	Reducible metals (Cr^{+6} , U^{6+})	Commercially applied
Limestone	Metals, acid waters	In practice (mining)
Precipitation agents (gypsum, hydroxyapatite)	Metals	Laboratory studies
Sorptive agents (Fe hydroxide, GAC ^a , zeolites, coal)	Metals and organics	Field demonstration and/or laboratory studies
Chemical reducing agents (dithionite, hydrogen sulfide)	Reducible metals	Field demonstration
Chemical oxidants (e.g., potassium permanganate)	Halocarbons	Field demonstration
Metal couples ^b	Halocarbons	Laboratory studies
Oxygen to stimulate aerobic bioremediation (e.g., via hydrogen peroxide, air, or oxygen gas)	Fuel hydrocarbons	Commercially applied
Methane, propane, and oxygen to stimulate co-cometabolic bioremediation (e.g., biosparging in horizontal wells)	Halocarbons, MTBE	Commercially applied
Organic electron donors (organic compost and mulch, alcohols, sugars, soybean oil, and fatty acids)	Halocarbons	Commercially applied

^a GAC = granulated activated carbon.

^b Coupled oxidation of metal and reduction of halocarbon to produce chloride and Fe^{+2} in solution.

Source: Modified from Rumer, R.R. and Mitchell, J.K. 1995. *Assessment of Barrier Containment Technologies—A Comprehensive Treatment for Environmental Remediation Applications*. National Technical Information Services, Springfield, VA.

however, other natural physical or chemical processes (e.g., adsorption) may effectively prevent the migration of contaminants. Under the MNA approach, natural forces are allowed to act on the contaminated groundwater, without any human intervention, to remediate groundwater. Because microorganisms are present in nearly all aquifers and because other naturally occurring processes (e.g., sorption, complexation, precipitation, etc.) can mitigate groundwater contamination by attenuating contaminants, these naturally occurring processes can have a beneficial impact on groundwater quality. A key component of MNA is a program of groundwater monitoring to verify that natural processes are, indeed, resulting in the degree of attenuation required. MNA can be used alone or to supplement conventional remediation techniques. For example, a contaminant source could be removed prior to beginning a program of MNA to reduce contaminant loading in the aquifer and thereby improve the performance of MNA (National Academy of Sciences, 2000, USEPA, 1998b).

36.5.2.6 Innovative *In Situ* Treatment Approaches

A large number of innovative techniques for *in situ* remediation of groundwater have been developed in the past several years that promise to provide a higher degree of permanence, a higher contaminant removal efficiency, or lower remediation cost for some groundwater contamination problems. A good summary of available innovative technologies is presented in USEPA (2003). Descriptions of some promising innovative *in situ* remediation techniques are presented below.

- *Nanoscale Zero-Valent Iron.* Since the late 1990s, nanoscale zero-valent iron (nZVI) has emerged as a promising new technology for *in situ* remediation of halogenated organics and heavy metals in groundwater. Laboratory Research over the past several years has demonstrated that nZVI and bimetallic Pd-Fe nanoparticles can destroy chlorinated biphenyls, ethenes, methanes, benzenes, and certain pesticides at rates that are orders of magnitude faster than can be achieved by granular ZVI and ZVI powder (Wang and Zhang, 1997; Lien and Zhang, 1999; Xu and Zhang, 2000; Masciangioli and Zhang, 2003). The faster transformation by nZVI has been shown to be due to higher specific surface area (the surface area of NZVI and commercial iron powder are $33.5 \text{ m}^2/\text{g}$, and $< 1 \text{ m}^2/\text{g}$, respectively). In addition to rapid transformation rates, a key benefit of nZVI is that it can be delivered to the subsurface via aqueous suspension. In contrast to granular ZVI, nZVI does not require trenching. A field pilot test in 2001 demonstrated that nZVI injected into the subsurface was well distributed and achieved effective treatment of chloroethenes and chloromethanes (Elliott and Zhang, 2001). However, nZVI is still an emerging technology, and additional development is required before the technology can be widely deployed. In particular, research and development is required regarding methods for improving delivery of nZVI to the subsurface.
- *Groundwater Circulation Well (GCW) Technologies.* GCW is a general term that describes groundwater remediation technologies that treat groundwater by creating a circulation cell around each individual treatment well. The treatment cells are typically created using wells with two screened intervals at the top and bottom of the treatment zone; groundwater is drawn in the bottom screen and reintroduced to the aquifer through the upper screen. Different variations of this technology employ different mechanisms to effect groundwater treatment. One variation used to strip VOCs from groundwater injects air into the bottom of the well; as the air lifts groundwater through the well, VOCs are stripped from the groundwater. Off gas from such a system can be treated above ground. Allmon et al. (1999) provides a detailed description and assessment of this variation of the technology. Other configurations of GCW technology may employ active pumping and other treatment technologies, such as bioremediation or carbon adsorption. GCW technology has been tested at more than 50 sites in the United States with mixed results. The application of GCW technology may be limited by the presence of highly anisotropic conditions or layers of variable permeability, which can preclude the formation of a recirculation cell, or the potential for biofouling in the circulation well screens (Allmon et al., 1999; Parson, 2001).
- *Surfactant Enhanced Aquifer Remediation.* A primary challenge of remediating chlorinated solvent DNAPL source areas is the slow rate of dissolution of DNAPL. Surfactant Enhanced Aquifer

Remediation (SEAR) is an emerging technology that offers promise for effective remediation of DNAPL source areas via enhanced solubilization of DNAPL. Surfactants, or Surface Active Agents, are chemicals that possess a molecular structure that is hydrophobic at one end of the molecule, and hydrophilic at the other end. In the presence of hydrophobic compounds, such as fats or chlorinated solvents, surfactants can form micelles, where the hydrophobic end of the surfactant joins with fat or solvent, and the hydrophilic tail of the surfactant facilitates solubilization of the compound. Chemical engineers have exploited surfactants for decades, using them in household detergents and as food-grade additives. Over the past several years, engineers and academic researchers have performed laboratory column and sand box tests to demonstrate that food-grade surfactants can be highly effective for enhanced recovery of DNAPLs. More recently, a handful of carefully controlled field pilot tests have shown that while SEAR can achieve effective results in the field (e.g., Ramsburg et al., 2005; Abriola et al., 2005), more testing is required before the technology can be widely applied. The ability of SEAR to solubilize and mobilize DNAPL requires that groundwater extraction systems be carefully designed to prevent any uncontrolled migration of DNAPL during remediation with SEAR. Technical guidance on SEAR is provided in NAVFAC, 2002.

36.5.3 *Ex Situ* Treatment

36.5.3.1 Biological Processes

Ex situ biological processes involve treatment of extracted groundwater in vessels until the concentration of the remaining contaminants is below a predefined level. Biological processes have been widely used for the past 40 years in the United States to treat municipal wastewater. A good general reference for biological treatment is presented by Corbitt (1990). The processes used to treat municipal wastewater are generally applicable to treatment of contaminated groundwater except in circumstances where the contaminants in groundwater are toxic to the microbes in the biological treatment system. *Ex situ* biological treatment processes include bioreactor and slurry-phase treatment, among others. A brief description of these two processes is presented below.

- *Bioreactors.* In this approach, contaminants are degraded by placing them in contact with microbes in an environment where biological growth is promoted. This can be accomplished using either suspended (i.e., activated sludge) or attached (i.e., rotating biological contactors or trickling filters) growth systems. The microbial population can be either derived from the contaminated groundwater or added to the system.
- *Slurry-Phase Treatment.* In this approach, extracted groundwater is mixed with soil and nutrients and then routinely agitated in a controlled environment to promote biodegradation of the contaminants. The slurry is typically placed in a lined area and left until the desired degree of biodegradation occurs. Finally, the slurry is dewatered, and then the water and soil (both should be uncontaminated after biodegradation is completed) are properly disposed of as nonhazardous wastes. This process offers much better control of process variables than the bioreactor approach but is slow and requires a larger area for implementation.

36.5.3.2 Volatilization Processes

As described in Section 36.5.2.5 and as shown in Table 36.13, volatilization processes are effective in treating groundwater that is contaminated with organic chemicals that have a high vapor pressure. Volatilization processes may also be effective at removing nonorganic contaminants that have a relatively high vapor pressure (e.g., mercury). Volatilization processes, or “air stripping,” involve mixing of contaminated groundwater with air to allow the volatile contaminants that are dissolved in groundwater to partition from the dissolved phase to the vapor phase, allowing them to be removed from groundwater. The efficiency of the volatilization processes depends on the number of air/water contacts. Air/water contacts can be maximized by using spray aeration, diffused-air aeration, packed column air

stripping, tray aerators, or rotating disk aerator techniques. The primary advantage of *ex situ* over *in situ* volatilization approaches is the flexibility of *ex situ* systems and the high efficiency rates of these systems.

36.5.3.3 Chemical and Physical Processes

Numerous well-developed technologies are available for treating extracted groundwater by chemical or physical means. Useful summaries of these processes are presented by the Federal Remediation Technology Roundtable (FRTR) (2002) and by Corbitt (1990). In general, chemical and physical processes are available for *ex situ* treatment of nearly all types of contaminants that occur in groundwater. *Ex situ* chemical and physical treatment techniques typically have the following advantages over *in situ* treatment techniques: (1) the ability to maintain uniformity of the influent by using equalization features (i.e., tanks etc.) at the point of inflow to the treatment system; (2) the ability to easily evaluate the effectiveness of the treatment process by monitoring the quality of treated groundwater; and (3) ease in adjusting the treatment system components in response to changes in influent quality to optimize groundwater treatment efficiency. In Table 36.18, several chemical and physical treatment technologies are identified, and examples of their use are illustrated. Below, descriptions are provided for the *ex situ* treatment techniques that are most

TABLE 36.18 Summary of Selected *Ex Situ* Groundwater Treatment Technologies

Treatment technology	Representative examples	Residual streams	Status of technology
<i>Organic Contaminants</i>			
Air stripping	Packed towers, surface or diffused aeration removal of volatile compounds; soil venting	Air stream with VOCs	Commercial
Liquid-phase	GAC removal of broad spectrum of VOCs	GAC for regeneration or disposal	Commercial
Steam stripping	Packed tower with steam stripping, removal of low volatile organics	Recovered solvent	Some commercial
Membranes	Ultrafiltration for removal of selected organics	Concentrated brine side stream	Commercial
Oxidation	Ozone/UV, or ozone/H ₂ O ₂ , destruction of chlorinated organics	None	Some commercial
Activated sludge	Oxygen or air biological oxidations for removal/destruction of degradable organics	Sludge	Commercial
Fixed-film biological reactors	Fixed-film fluidized bed, for oxidation of less degradable organics	Sludge	Commercial
Biophysical	Powdered carbon, with activated sludge, treatment of high strength waste waters	Powdered carbon and bacterial	Commercial, PACT process
<i>Inorganic Contaminants</i>			
Alkaline precipitation	Heavy metals removal	Hazardous sludge	Commercial
Coagulation	Ferric sulfate or alum for heavy metals removal	Hazardous sludge	Commercial
Ion exchange	Heavy metals; nitrate	Regeneration stream	Commercial
Adsorption	Selenium removal on activated alumina	Regeneration stream	Commercial
Filtration	Removal of clays, other particulates	Backwash wastes	Commercial
Reduction	SO ₂ reduction of Cr (VI)	Sludge	Commercial
Membranes	Reverse osmosis, ultrafiltration for removal of metals, other ions	Concentrated liquid waste	Commercial, new membranes are under development
Oxidation	Fe (II) and Mn (II)	Sludge	Commercial

Source: From USEPA. 2003. *Superfund Innovative Technology Evaluation Program, Technology Profiles*, Eleventh Edition. EPA/540/C-03/501. Washington, DC.

commonly used to treat contaminated groundwater.

- *Carbon Adsorption:* For this approach, groundwater is typically pumped through a series of canisters containing activated carbon, to which contaminants from the groundwater are adsorbed. Because activated carbon offers “attractive” locations for adsorption of organic carbons (i.e., because of the similarity between the carbon atoms on the contaminant molecule and the carbons in the activated carbon), carbon adsorption is a very effective technique for removing VOCs from contaminated groundwater. Periodically, the activated carbon must be either replaced or regenerated as the adsorption sites become saturated with VOCs. For some application, biologically activated carbon (BAC) treatment may be employed to enhance removal efficiency of minimize the need for carbon replacement.
- *Precipitation:* Precipitation occurs when a constituent exists at a concentration in groundwater that is greater than its solubility limit. Solubility is a function of many factors, including pH, temperature, and the presence of other dissolved constituents. Precipitation of a target constituent is caused by manipulating these factors to change the equilibrium in the direction that promotes precipitation of that constituent. Typically, the least-soluble constituent at a given pH will precipitate from the solution. Note that precipitation can also occur *in situ* in response to changes in groundwater chemistry. Because solubility is affected by many factors, changes in groundwater quality over time may cause changes in the solubility of constituents of concern. Precipitation and dissolution may be slow reactions, depending on the solubility of the constituent; therefore, after a change occurs in groundwater chemistry that increases the constituent’s solubility (e.g., drawing groundwater having lower pH through a contaminated aquifer), the previously precipitated minerals may require a long time to redissolve and be removed from the aquifer.
- *Reverse Osmosis:* In reverse osmosis (RO) (USEPA, 1996), solids are separated from water by creating a concentration gradient across a semipermeable membrane (i.e., a membrane that is “selectively” permeable to only the solids of choice, based on size of molecules, but not to other compounds or solids). To prevent water from also passing through the membrane, a pressure gradient is applied to the opposite side of the membrane that exceeds the osmotic pressure (hence the use of the term “reverse” osmosis). The technology is extremely effective at removing dissolved solids from groundwater, but RO capital equipment is very expensive.

36.6 Performance Monitoring of Groundwater Remediation Systems

Monitoring the performance of a groundwater remediation system allows evaluation of the success of the system in meeting the remediation goal and the design criteria for the system. Different approaches are typically used to monitor the performance of groundwater extraction, barrier, and treatment remediation systems. Performance monitoring of groundwater extraction systems typically includes at least the following activities (USEPA, 1994):

- Confirming that the system is extracting groundwater at flow rates similar to those predicted in the predesign pump test (by measurement using a flowmeter at the wellhead)
- Verifying that groundwater elevations near the extraction system are being affected as needed to control groundwater flow direction and gradient (by measuring groundwater elevations at wells or piezometers located near the extraction well(s))
- Verifying that there are no significant breaches in the groundwater extraction system that would allow contaminated groundwater to flow past the extraction system (using the groundwater elevation data collected at the wells and piezometers)
- Confirming that the system components perform consistently throughout the operational life of the facility (by routinely inspecting the components of the system and checking for signs of operational problems).

Performance monitoring of a groundwater treatment system (both *in situ* and *ex situ*) typically includes the following:

- Periodic analysis of influent groundwater quality to evaluate changes in the quality of water requiring treatment
- Periodic analysis of effluent groundwater quality, to evaluate the success of the system at meeting treatment goals
- Routine inspection of the components of the system to identify potential future problems.

Performance monitoring of subsurface barriers (see Rumer and Mitchell, 1995) typically includes the following activities:

- Verifying that the barrier is having the desired impact on groundwater flow direction and gradient (by measuring groundwater elevations at wells or piezometers near the barrier)
- Verifying that contaminated groundwater is not passing through or around the barrier (by monitoring groundwater quality beyond the barrier)
- Verifying that the barrier is not degrading over time.

These activities can be monitored in a variety of ways using a system of piezometers or wells, as described in Section 36.3.5. When the purpose of the performance-monitoring program is to verify that nearby receptors are not at risk of contaminated groundwater, external monitoring features should be located downgradient of the hydraulic containment system near receptors or near a property boundary.

Glossary

Adsorption The adherence of ions or molecules in solution to the surface of solids.

Anisotropy The condition under which one or more of the hydraulic properties of an aquifer vary with direction.

Biodegradation The biologically mediated conversion of a compound to simple products.

Contaminant An undesirable substance that is not normally present in groundwater or a substance that may naturally occur in groundwater, but that is present at an undesirably high concentration.

Desorption The release of sorbed molecules from the solid into solution (the reverse of adsorption).

DNAPL An acronym for denser-than-water nonaqueous-phase liquid; a liquid, composed of one or more contaminants, that does not mix with water and is denser than water.

Fraction of Organic Carbon The organic carbon content of soil, expressed as a mass fraction of the dry soil.

Hydraulic Control Prevention of the spread of contaminated groundwater using physical features to control groundwater flow.

Ion A molecule or atom that has a positive or negative electric charge.

LNAPL An acronym for less-dense-than-water (i.e., “light”) nonaqueous-phase liquid.

Maximum Contaminant Level (MCL) The maximum amount of a compound allowed in drinking water under the Safe Drinking Water Act.

Octanol-Water Partition Coefficient A measure that indicates the extent to which a compound is attracted to an organic phase (for which octanol is a proxy) and, hence, the tendency of the compound to sorb to subsurface materials.

Partitioning A chemical equilibrium condition in which the concentration of a chemical is apportioned between two different phases according to the partition coefficient.

Plume A zone of contaminated groundwater.

Receptor An organism that is exposed to a contaminant.

Retardation The movement of a solute through a geologic medium at a velocity less than that of the groundwater as a result of phenomena that separate a fraction of the solute mass from the groundwater.

Sorption Refers to a reversible process involving physical reaction of aquifer material and dissolved constituents.

Upgradient In the direction of increasing hydrostatic head.

Volatilization The transfer of a chemical from the liquid to the gas phase.

References

- Abriola, L.M., Drummond, C.D., Hahn, E.J., Hayes, K.F., Kibbey, T.C.G., Lemke, L.D., Pennell, K.D., Petrovskis, E.A., Ramsburg, C.A., and Rathfelder, K.M. 2005. Pilot-scale demonstration of surfactant-enhanced PCE solubilization. 1. Site characterization and Test Design. *Environ. Sci. Technol.* 39: 1778–1790.
- Allmon, W.E., Everett, L.G., Lighter, A.T., Alleman, B., Boyd, T.J., and Spargo, B.J. 1999. *Groundwater Circulating Well Technology Assessment*. Report No. NRL/PU/6115-99-384. Naval Research Laboratory, Washington, DC. <http://www.estcp.org/projects/cleanup/199602v.cfm>.
- Bear, J. 1979. *Hydraulics of Groundwater*, McGraw-Hill, New York.
- Blowes, D.W., Gillham, R.W., Ptacek, C.J., Puls, R.W., Bennett, T.A., O'Hannesin, S.F., Hanton-Fong C.J., and Bain, J.G. 1999a. *An In Situ Permeable Reactive Barrier for the Treatment of Hexavalent Chromium and Trichloroethylene in Ground Water: Volume 1 Design and Installation*. Office of Research and Development. EPA/600/R-99/095a. http://www.epa.gov/ada/download/reports/prbdesign_v1.pdf.
- Blowes, D.W., Puls, R.W., Gillham, R.W., Ptacek, C.J., Bennett, T.A., Bain, J.G., Hanton-Fong, C.J., and Paul, C.J. 1999b. *An In Situ Permeable Reactive Barrier for the Treatment of Hexavalent Chromium and Trichloroethylene in Ground Water: Volume 2 Performance Monitoring*. Office of Research and Development. EPA/600/R-99/095b. http://www.epa.gov/ada/download/reports/prbperformance_v2.pdf
- Blowes, D.W. and Mayer, K.U. 1999. *An In Situ Permeable Reactive Barrier for the Treatment of Hexavalent Chromium and Trichloroethylene in Ground Water: Volume 3 Multicomponent Transport Modeling*. Office of Research and Development. EPA/600/R-99/095c. http://www.epa.gov/ada/download/reports/prbmodel_v3.pdf
- Brusseau, M. 1996. Evaluation of simple methods of estimating contaminant removal by flushing. *Groundwater* 34:19–22.
- Cohen, R.M. and Miller, W.J. III, 1983. Use of analytical models for evaluating corrective actions at hazardous waste sites. Proceedings of the Third National Symposium on Aquifer Restoration and Ground-Water Monitoring. Nielson, D.M., ed. *National Water Well Association*. Worthington, OH, 86–97.
- Corbitt, R.A. (ed.) 1990. *Standard Handbook of Environmental Engineering*. McGraw-Hill, New York.
- Davis, E.L. 1997. *Ground Water Issue: How Heat Can Enhance In-situ Soil and Aquifer Remediation: Important Chemical Properties and Guidance on Choosing the Appropriate Technique*. EPA/540/S-97/502. Washington, DC.
- Driscoll, F. 1986. *Groundwater and Wells*, Johnson Division, St. Paul, MN, Second Edition.
- Elliott, D. and Zhang, W. 2001. Field assessment of nanoparticles for groundwater treatment. *Environmental Science & Technology* 35: 4922–4926.
- Federal Remediation Technology Roundtable, 2002. *Remediation Technologies Screening Matrix and Reference Guide*. EPA 542-B-93-005. Washington, DC.
- Focht, R., Vaughn, S., and O'Hannesin. 1996. Field application of reactive iron walls for in-situ degradation of volatile organic compounds. *Remediation*. Summer 1996.
- GAO. 2005. *Groundwater Contamination: DOD Uses and Develops a Range of Remediation Technologies to Clean Up Military Sites*. Washington, DC.
- GeoSyntec Consultants. 1994. *Subsurface Barriers*. Report to the United States Navy, Contract SBIR Topic N93-130. Atlanta, GA. 80 pages.
- Hinchee, R.E. 1994. *Air Sparging for Site Remediation*. Lewis Publishers, Chelsea, MI.

- Hughes, J.B., Duston, K.L., and Ward, C.H. 2002. *Engineered Bioremediation. Technology Evaluation Report* TE-02-03. Ground-Water Remediation Technologies Analysis Center, Pittsburgh, PA. http://www.groundwatercentral.info/org/pdf/E_bio.pdf
- Hutson, S.S., Barber, N.L., Kenny, J.F., Linsey, K.S., Lumma, D.S., and Maupin, M.A. 2004. Estimated Use of Water in the United States in 2000. U.S. Geological Survey Circular. 1268, Reston, VA.
- Interstate Technology and Regulatory Council. 2002. *A Systematic Approach to In Situ Bioremediation in Groundwater*. <http://www.itrcweb.org/Documents/ISB-8.pdf>.
- Interstate Technology and Regulatory Council. 2005. *Technical and Regulatory Guidance for In Situ Chemical Oxidation of Contaminated Soil and Groundwater*. <http://www.itrcweb.org/Documents/ISCO-2.pdf>.
- Inyang, 1992. *Selection and design of slurry walls as barriers to control pollutant migration*. A seminar Presentation, Office of Solid Waste and Emergency Response, USEPA, Washington, DC.
- Johnson, P.C., Stanley, C.C., Kemblowski, M.W., Byers, D.L., and Colthart, J.D. 1990. A practical approach to the design, operation, and monitoring of in-situ soil-venting systems. *Groundwater Monitoring Review*, Spring 1990, 159–178.
- Johnson, R.I., Johnson, P.C., McWhorter, D.B., Hinchee, R.E., and Goodman, I. 1993. An overview of in situ air sparging. *Groundwater Monitoring and Remediation*. Fall 1993, 127–135.
- Keely, J.F. 1989. *Performance Evaluations of Pump-and-Treat Remediations*, EPA Superfund Groundwater Issue, EPA/540/4-89/005. Washington, DC. McCoy and Associates. September/October 1992. *The Hazardous Waste Consultant*. 4.1–4.38.
- Leeson A., Johnson, P.C., Johnson, R.L., C.M. Vogel, R.E. Hinchee, M. Marley, T. Peargin, C.L. Bruce, I.L. Amerson, C.T. Coonfare, R.D. Gillespie, and D.B. McWhorter. 2002. Air Sparging Design Paradigm. Battelle Memorial Institute, Columbus, Ohio. http://www.estcp.org/documents/techdocs/Air_Sparging.pdf.
- Leeson, A. and Hinchee, R.E. 1996a. *Principles and Practices of Bioventing Volume I: Principles and Practices*. Battelle Memorial Institute, Columbus, OH.
- Leeson, A. and Hinchee, R.E. 1996b. *Principles and Practices of Bioventing Volume II: Bioventing Design*. Battelle Memorial Institute, Columbus, OH.
- Lien, H. and Zhang, W. 1999. Reactions of Chlorinated Methanes with Nanoscale Metal Particles in Aqueous Solutions. *J. Environ. Eng.*, 136: 1042–47.
- Masciangioli, T. and Zhang, W. 2003. Environmental Nanotechnology: Potential and Pitfalls. *Environmental Science & Technology*, 37:102A–108A.
- Merritt, F.S. 1989. *Standard Handbook for Civil Engineers*. McGraw-Hill, New York.
- National Academy of Sciences. 2000. *Natural Attenuation for Groundwater Remediation*. Water Science and Technology Board. Washington, DC.
- National Research Council (NRC), Water Science and Technology Board. 1994. *Alternatives for Groundwater Cleanup*. Washington, DC.
- Naval Facilities Engineering Command. 2002. *Surfactant-Enhanced Aquifer Remediation (SEAR) Design Manual*. Technical Report TR-2206-ENV. Washington, DC.
- Parsons Engineering Science, Inc. 2001. *Groundwater Circulation Well Technology Evaluation at Facility 1381 Cape Canaveral Air Station, Florida*. Technical Summary Report.
- Ramsburg, C.A., Pennell, K.D., Abriola, L.M., Daniels, G., Drummond, C.D., Gamache, M., Hsu, H.-L., Petrovskis, E.A., Rathfelder, K.M., Ryder, J.L., and Yavaraski, T.P. 2005. Pilot-scale demonstration of surfactant-enhanced PCE solubilization. 2. System operation and evaluation. *Environ. Sci. Technol.* 39:1791–1801.
- Rumer, R.R. and Mitchell, J.K. 1995. *Assessment of Barrier Containment Technologies — A Comprehensive Treatment for Environmental Remediation Applications*. National Technical Information Services, Springfield, VA.
- Russel, M., Colglazier, E.W., and English, M.R. 1991. *Hazardous Waste Site Remediation: The Task Ahead*. Waste Management Research and Education Institute. University of Tennessee, Knoxville, TN.

- Sai, J.O. and Anderson, D.C. 1992. Barrier wall materials for containment of dense nonaqueous phase liquid (DNAPL). *Hazardous Waste and Hazardous Materials*. 9, 4, 317–330.
- U.S. Air Force Center for Environmental Excellence. 2004. *Principles and Practices of Enhanced Bioremediation of Chlorinated Solvents*. San Antonio, TX. <http://www.afcee.brooks.af.mil/products/techtrans/bioremediation/downloads/PrinciplesandPractices.pdf>.
- USEPA. 1989. Seminar Publication — *Transport and Fate of Contaminants in the Subsurface*. EPA/636/4-89/019. Washington, DC.
- USEPA. 1991. *Handbook — Groundwater, Volume II: Methodology*. EPA/636/6-90/016b. Washington, DC.
- USEPA. 1993. *Guidance for Evaluating the Technical Impracticability of Ground-Water Restoration*. EPA-540-R-93-080. Washington, DC.
- USEPA. 1994. *Monitoring the Performance of Groundwater Pump and Treat Systems*. EPA/600/R-94/123. Washington, DC.
- USEPA. 1995a. *Remedial Design/Remedial Action Handbook*. EPA 540/R-95/059. Washington, DC.
- USEPA. 1995b. *Ground-Water and Leachate Treatment Systems*. EPA/636/R-94/005. Washington, DC.
- USEPA. 1995c. *Evaluation of Technologies for In-Situ Cleanup of DNAPL Contaminated Sites*. EPA/600/R-94/0120. Washington, DC.
- USEPA. 1996. *Capsule Report: Reverse Osmosis Process*. EPA/636/R-96/009. Washington, DC.
- USEPA. 1998a. *Field Applications of In Situ Remediation Technologies: Chemical Oxidation*. Office of Solid Waste and Emergency Response. EPA 542-R-98-008. <http://www.clu-in.org/download/remed/chemox.pdf>.
- USEPA. 1998b. *Technical Protocol for Evaluating Natural Attenuation of Chlorinated Solvents in Ground Water*. Office of Research and Development. Washington, DC.
- USEPA. 2000a. *A Guide to Developing and Documenting Cost Estimated During the Feasibility Study*. EPA 540-R-00-002. Washington, DC.
- USEPA. 2000b. Risk Assessment Guidance for Superfund: Interim Final Guidance (Parts A through E). Office of Emergency and Remedial Response. Washington, DC.
- USEPA. 2000c. *Engineered Approaches to In Situ Bioremediation of Chlorinated Solvents: Fundamentals and Field Applications*. Office of Solid Waste and Emergency Response. EPA 542-R-00-008. <http://www.epa.gov/tio/download/remed/engappinsitbio.pdf>.
- USEPA. 2002. *A Framework for the Economic Assessment of Ecological Benefits*. Ecological Benefit Assessment Workgroup, Social Sciences Discussion Group, Science Policy Council. Washington, DC.
- USEPA. 2003. *Superfund Innovative Technology Evaluation Program, Technology Profiles*, Seventh Edition. EPA/540/R-94/526. Washington, DC.
- USEPA. 2004a. *Handbook of Groundwater Protection and Cleanup Policies for RCRA Corrective Action for Facilities Subject to Corrective Action Under Subtitle C of the Resource Conservation and Recovery Act*. EPA 530-R-04-030. Washington, DC.
- USEPA. 2004b. *Cleaning up the Nation's Waste Sites: Markets and Technology Trends*. 2004 Edition. EPA 542-R-04-015. Washington, DC.
- USEPA. 2004c. *Superfund Chemical Data Matrix*. EPA/540-R-94-009. Washington, DC.
- USEPA. 2004d. *In Situ Thermal Treatment of Chlorinated Solvents: Fundamentals and Field Applications*. Office of Solid Waste and Emergency Response, Office of Superfund Remediation and Technology Innovation. EPA-542-R-04-010. <http://www.clu-in.org/download/remed/epa542r04010.pdf>
- USGS, 2005. *Estimated Use of Water in the United States in 2000*. NRL/PU/6115-99-384.
- Wang, C. and Zhang, W. 1997. Nanoscale Metal Particles for Dechlorination of PCE and PCBs, *Environmental Science & Technology*, 31: 2154–2156.
- Xanthakos, P. 1979. *Slurry Walls*. McGraw-Hill, New York.
- Xu, Y. and Zhang, W. 2000. Subcolloidal Fe/Ag Particles for Reductive Dehalogenation of Chlorinated Benzenes. *Industrial & Engineering Chemistry Research*, 39: 2238–2244.
- Yin, Y. and Allen, H.E. 1999. *In Situ Chemical Treatment*. Ground-Water Remediation Technologies Analysis Center. Technology Evaluation Report No. TE-99-01. http://www.groundwatercentral.info/org/pdf/E_inchem.pdf.

Zheng, C., Wong, H.F., Anderson, M.P., and Bradbury, K.R. 1988. Analysis of interceptor ditches for control of groundwater pollution. *J. Hydrol.* 98: 67–81.

Further Information

USEPA (2000b) presents a good discussion of the basic approaches involved in performing a risk assessment.

A discussion of selection of cleanup goals and current issues in groundwater remediation technology is presented by NRC (1994).

USEPA (2003) evaluates emerging remediation technologies and routinely publishes updates of the results of the program.

37

Geosynthetics

37.1	Introduction.....	37-1
37.2	Geosynthetic Functions	37-3
	Design by Function of Geosynthetics • Separation Function	
	• Reinforcement Function • Filtration Function •	
	Drainage Function • Barrier Function • Protection	
	Function	
37.3	Geosynthetic Types	37-10
	Geotextiles • Geomembranes • Geogrids • Geosynthetic	
	Clay Liners • Geocomposite Sheet Drains • Geocomposite	
	Strip (Wick) Drains • Geocells • Erosion Control Products	
	• HDPE Vertical Barrier Systems	
37.4	Geosynthetic Applications in Landfill Design	37-27
37.5	Case Histories	37-28
	Case History of Vertical Barrier System • Case History of	
	Multiple Use of Geosynthetics in Landfill Cover Design	
	Glossary.....	37-31
	References	37-34
	Further Information	37-35

Jorge G. Zornberg
University of Texas at Austin

Barry R. Christopher
Independent Consultant

37.1 Introduction

Geosynthetics can be defined as planar products manufactured from polymeric material, which are used with soil, rock, or other geotechnical engineering-related material as an integral part of a man-made project, structure, or system (ASTM, 1995). Geosynthetics are widely used in many geotechnical, environmental, and hydraulic applications related to groundwater quality and control. One of the most common examples is the use of geotextile filters in trench (i.e., French) drains. Base and cover liner systems for modern landfills also make extensive use of geosynthetics with the main purpose of minimizing the potential for groundwater contamination. Furthermore, the use of geosynthetics is rapidly increasing in applications related directly to groundwater control. This is the case of high density polyethylene (HDPE) vertical barrier systems, which can be used instead of traditional soil-bentonite cutoff walls in projects involving groundwater remediation and control.

The geosynthetics market is strong and rapidly increasing due to the continued use of geosynthetics in well-established applications and, particularly, due to the increasing number of new applications that make use of these products. The strength of the geosynthetics market can be appreciated by evaluating the growth in the estimated amount of geosynthetics in North America over the years. Table 37.1 shows the estimated North American shipments of geosynthetics for 2001. While the total amount of geosynthetics

TABLE 37.1 North American Shipments of Geosynthetics
Material 1995–2001 (in million m²)

	1995	1996	1998	2001
Geotextiles	346.2	356.2	419.7	477.4
Geomembranes	62.4	64.4	74.6	86.8
Geogrids	22.4	24.3	29.1	36.8
Geosynthetic clay liners	5.0	5.4	6.1	8.2
Erosion-control products	72.7	77.8	82.8	93.6
Specialty geosynthetics	16.7	20.1	25.9	31.8

Source: Industrial Fabrics Association International (1996).

produced in North America was slightly over 83 million m² in 1980, the production of geosynthetics was approximately 1 billion m² in 2005.

Geosynthetics applications are very diverse. To fulfill different functions in the design of geotechnical-, environmental-, and hydraulic-related systems, the geosynthetic industry has developed a number of products to meet engineers' needs. In addition to the use of geotextiles as filters in trench drains, geomembranes in landfill liner systems, and HDPE vertical panels in groundwater control projects, other examples of geosynthetics applications include the use of geotextiles as filtration elements in dams and waste containment systems, the use of geocomposites as erosion control elements in channels and slopes, and the use of geogrids as reinforcement elements in soil embankments, to mention just a few.

Numerous tests have been developed to characterize the hydraulic and mechanical properties of geosynthetics. Many of these properties are important in the manufacture and quality control of geosynthetics; however, many others are also important in design. The material properties that are primarily related to the manufacture and quality control of geosynthetics are generally referred to as index properties and those related to the design are referred to as performance properties. When properly correlated to performance properties, some index properties may also be used for design. As the various geosynthetic products can perform different functions, they should be designed to satisfy minimum criteria to adequately perform these functions in a given design. The different functions performed by geosynthetics are discussed in Section 37.2. The functions that geosynthetics may provide are as follows: separation, reinforcement, filtration, drainage, barrier, and/or protection (or stress relief).

Geosynthetics are manufactured in sheet form in a factory-controlled environment. They are most often packaged in rolls for transporting to the site. They may also be folded or cut and stacked and placed in cartons. At the project site the geosynthetic sheets are unrolled on the prepared subgrade surface, overlapped to form a continuous geosynthetic blanket, and often physically joined to each other, for example, by welding (geomembranes) or sewing (geotextiles). The different types of geosynthetics are discussed in Section 37.3. They include geotextiles, geomembranes, geogrids, geosynthetic clay liners (GCLs), geocomposite sheet drains, geocomposite strip (wick) drains, geocells, erosion control products and HDPE vertical barrier systems.

Different types of geocomposite drains and HDPE vertical barriers, a special form of geomembranes, are considered separately in the list above. These geosynthetics are described separately in this chapter because of their particular relevance in groundwater-related applications.

Geotechnical, environmental, and hydraulic systems frequently incorporate several types of geosynthetics, which are designed to perform more than one function in the system. The bottom and cover liner systems of waste containment facilities are good examples of applications that make use of geosynthetics for multiple purposes. In these facilities, the different geosynthetic products are combined to create the liner system, the components fulfill the functions of infiltration barrier, filtration, separation, drainage, protection, and reinforcement. The multiple uses of geosynthetics in the design of modern landfills are described in Section 37.4. Finally, a case history illustrating the use of HDPE panels as vertical barrier in a groundwater control project is presented in Section 37.5. A glossary of relevant terms and a list of sources are also included for further information.

37.2 Geosynthetic Functions

37.2.1 Design by Function of Geosynthetics

As with other engineering materials, there are several design approaches that could be used during the selection process of geosynthetic products. The most common methods are design by experience, by specification, or by function (Koerner, 2005).

Design-by-experience is generally based on the use of manufacturer's literature and of the designer's experience and familiarity with specific geosynthetic products. An extension of design by experience is design-by-specification, which consists of selecting geosynthetic products for common application areas based on basic minimum or maximum specified property values. These methods may be acceptable for routine, repeat, non-critical applications.

Design-by-function is the preferred approach. Design-by-function should be performed as a check for applications covered by specifications and required for applications not covered by specifications or of such a nature that large property or personal damage would result in the event of a failure. A generic design process that applies to the different geosynthetic functions is summarized as follows (Koerner, 2005):

1. Evaluate the critical and severe nature of the application
2. Determine the function(s) of the geosynthetic
3. Calculate, estimate, or otherwise determine the required property value for the function(s)
4. Test or otherwise obtain the allowable property of the candidate geosynthetic material(s)
5. Calculate the factor of safety (FS) as follows:

$$FS = \frac{\text{allowable (test) value}}{\text{required (design) value}} \quad (37.1)$$

6. Determine if the resulting factor of safety is adequate for the site-specific situation under consideration
7. Prepare specifications and construction documents
8. Observe construction and post-construction performance

If the factor of safety is sufficiently high for the specific application, the candidate geosynthetic(s) is (are) deemed acceptable. The same process can be repeated for a number of available geosynthetics, and the final selection among acceptable products is based on availability and cost.

The design-by-function approach is the general approach to be followed in the majority of the projects. As mentioned, the primary function(s) of geosynthetics can be separation, reinforcement, filtration, drainage, infiltration barrier, or protection. However, a certain geosynthetic product can perform different functions and, similarly, the same function can often be performed by different types of geosynthetics. Geosynthetic applications are usually defined by the primary or principal function. In addition, geosynthetics can perform one or more secondary functions, which must also be considered when selecting the geosynthetic characteristics for optimum performance. For example, a geotextile can provide separation of two dissimilar soils (e.g., gravel from clay in a road), but the geosynthetic may also be required to provide the secondary function of filtration to minimize the build up of excess pore water pressure in the soil beneath the separator. The specific function(s) of the different geosynthetic(s) are presented in Table 37.2. Each of these functions is described in Section 37.2.2 to Section 37.2.7.

37.2.2 Separation Function

Separation is the introduction of a flexible, porous geosynthetic product between dissimilar materials so that the integrity and functioning of both materials can remain intact or be improved. For example, a major cause of failure of roadways constructed over soft foundations is contamination of the aggregate base course with the underlying soft subgrade soils (Figure 37.1a). Contamination occurs both due to: (1) penetration of the aggregate into the weak subgrade due to localized bearing capacity failure under

TABLE 37.2 Function of Different Geosynthetic Products

	Geotextile	Geo-membrane	Geogrid	GCL	Geocomposite sheet drain	Geocomposite strip (wick) drain	Geocell	Erosion control product	HDPE vertical barrier
Separation	X	X			X				
Reinforcement	X		X					X	
Filtration	X				X				
Drainage	X				X		X		
Barrier	X ^a	X		X					X
Protection	X			X	X			X	X

^a Asphalt-saturated geotextiles.

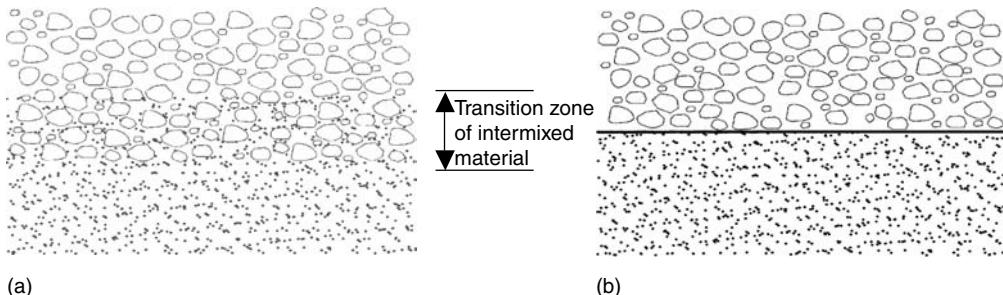


FIGURE 37.1 Separation function of a geotextile placed between road aggregate and soft saturated subgrade. (a) Without geotextile and (b) With geotextile.

stresses induced by wheel loads, and (2) intrusion of fine-grained soils into the aggregate because of pumping or subgrade weakening due to excess pore water pressure. Subgrade contamination results in inadequate structural support that often leads to premature failure of the system. A geotextile can be placed between the aggregate and the subgrade to act as a separator and prevent the subgrade and aggregate base course from mixing (Figure 37.1b).

Among the different geosynthetics, geotextiles have been the products generally used in the function of separation. Examples of separation applications are the use of geotextiles between subgrade and stone base in roads and airfields, and between geomembranes and drainage layers in landfills. In addition to these applications, in which separation is the primary function of the geotextile, it could be said that most other geosynthetic applications generally include separation as a secondary function. The design of geosynthetics for separation applications is provided by Holtz et al. (1997, 1998) and Koerner (2005).

37.2.3 Reinforcement Function

Geosynthetic inclusions within a soil mass can provide a reinforcement function by developing tensile forces that contribute to the stability of the geosynthetic-soil composite (a reinforced soil structure). Design and construction of stable slopes and retaining structures within space constraints are major economical considerations in geotechnical engineering projects. For example, when geometry requirements dictate changes of elevation for a retaining wall, or dam project, the engineer faces a variety of distinct alternatives for designing the required earth structures. Traditional solutions have been either a near vertical concrete structure or a conventional, relatively flat, unreinforced slope (Figure 37.2). Although simple to design, concrete wall alternatives have generally led to elevated construction and material costs. On the other hand, the construction of unreinforced embankments with flat slope angles dictated by stability considerations is an alternative often precluded in projects where design is controlled by space constraints. As shown in

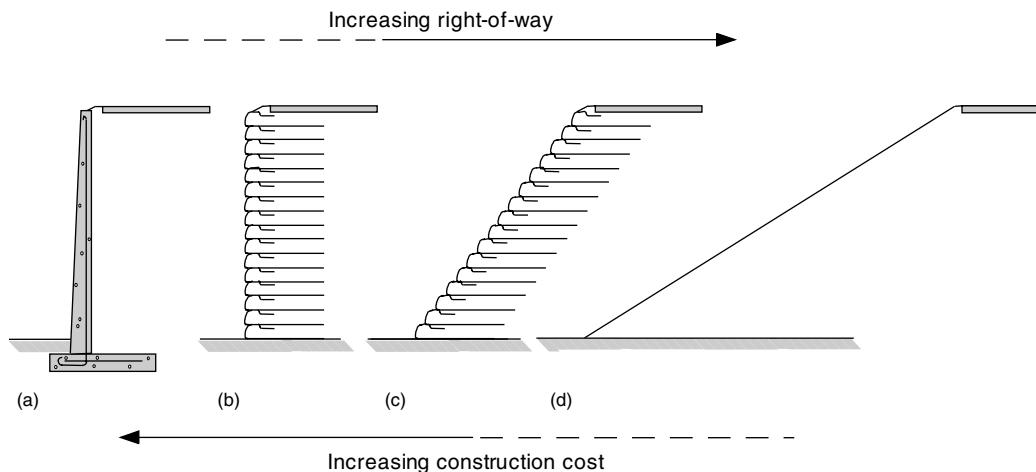


FIGURE 37.2 Reinforcement function of geosynthetics used to optimize the design of earth retaining structures: (a) concrete retaining wall; (b) reinforced wall; (c) reinforced slope; (d) unreinforced slope.

Figure 37.2, an alternative would be to place horizontal, geosynthetic reinforcing elements in the soil to allow construction of very steep embankment side slopes or a vertical reinforced soil wall. For example, vertical reinforced soil walls have been used to both construct dams for reservoirs and increase the height of existing dams.

Geosynthetic products typically used as reinforcement elements are geotextiles and geogrids. Additional products include geocells and fiber reinforcement. Reinforced soil walls generally provide vertical grade separations at a lower cost than traditional concrete walls. Reinforced wall systems involve the use of facing elements such as precast panels, cast-in-place concrete panels, or modular block systems. Alternatively, steepened reinforced slopes (with facing inclination below approximately 70°) may eliminate the use of facing elements, thus saving material costs and construction time in relation to vertical reinforced walls. As indicated in Figure 37.2, a reinforced soil system generally provides an optimized alternative for the design of earth retaining structures by combining lower cost and decreased right-of-way requirements.

The effect of geosynthetic reinforcements on the stability of slopes is illustrated in Figure 37.3, which shows a reduced scale geotextile-reinforced slope model built using dry sand as backfill material. The maximum slope inclination of unreinforced sand under its own weight is the angle of repose of the sand, which is well below the inclination of the slope face of the model in the figure. Horizontal geotextile reinforcements placed within the backfill provided stability to the steep sand slope. In fact, not only did the reinforced slope model not fail under its own weight, but its failure only occurred when the unit weight of the backfill was increased 67 times by placing the model in a geotechnical centrifuge (Zornberg et al., 1998). Figure 37.4 shows the reinforced slope model after centrifuge testing.

The use of inclusions to improve the mechanical properties of soils dates to ancient times. However, it is only within the last 35 years (Vidal, 1969) that analytical and experimental studies have led to the contemporary soil reinforcement techniques. Soil reinforcement is now a highly attractive alternative for embankment and retaining wall projects because of the economic benefits it offers in relation to conventional retaining structures. Its acceptance has also been triggered by a number of technical factors including aesthetics, reliability, simple construction techniques, good seismic performance, and the ability to tolerate large deformations without structural distress. The design of reinforced soil walls is based on earth pressure theory while the design of reinforced slopes is based on the use of limit equilibrium methods. The design process involves evaluation of the external (global), internal, and compound stability of the structure. The required tensile strength of the reinforcements is selected in design so that the margins of safety to prevent internal failure, such as that shown in Figure 37.4, are adequate. Guidance in soil reinforcement design procedures is provided by Elias et al. (2001), Holtz et al. (1997, 1998), and NCMA (1997).



FIGURE 37.3 Model of a sand slope reinforced with geosynthetics.

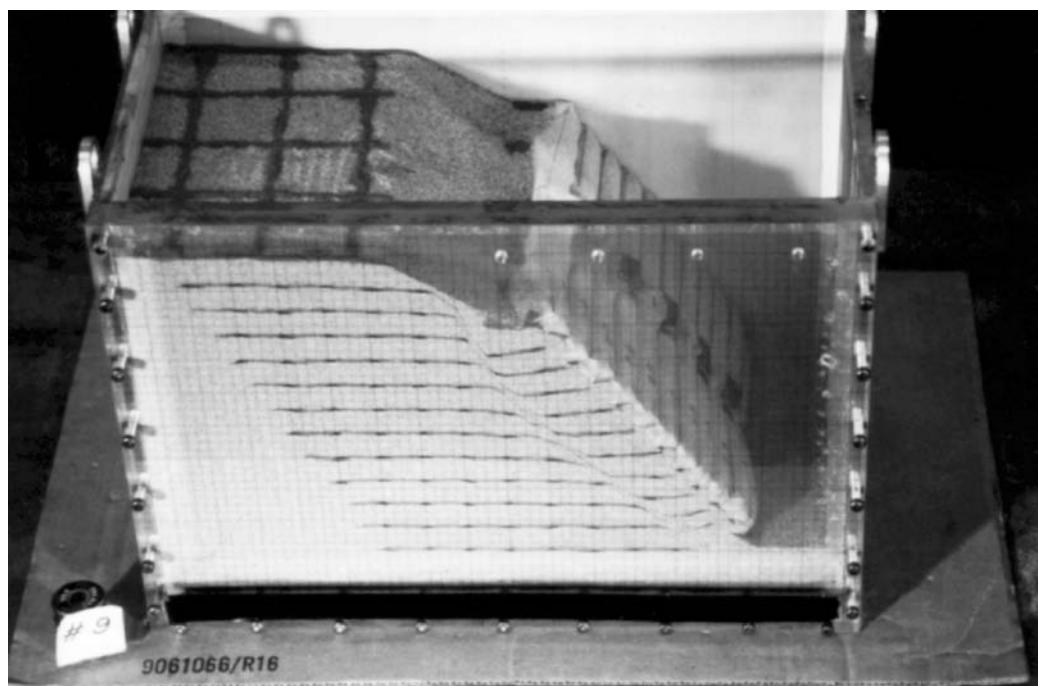


FIGURE 37.4 Reinforced slope model brought to failure by increasing the unit weight of the backfill.

A special form of geosynthetic reinforcement is the mixing of fibers with the soil. Fiber reinforcement is used in applications such as stabilization of thin soil veneers, localized repair of failed slopes and increasing the seismic performance. Randomly distributed fibers can provide isotropic strength increases to the soil and avoid the existence of the potential planes of weakness that can develop on the soil-reinforcement interface. The design of fiber-reinforced soil slopes has typically been performed using composite approaches, where the fiber-reinforced soil is considered as a single homogenized material. Accordingly, fiber-reinforced soil design has required non-conventional laboratory testing of composite fiber-reinforced soil specimens, which has discouraged implementation of fiber-reinforcement in engineering practice. A new discrete approach was recently proposed (Zornberg, 2002), which predicts the “equivalent” shear strength of the fiber-reinforced soil based on the independent properties of fibers (e.g., fiber content, fiber aspect ratio) and soil (e.g., friction angle and cohesion).

37.2.4 Filtration Function

The filtration function involves movement of liquid through the geosynthetic and, at the same time, retention of soil on its upstream side. As indicated in Table 37.2, geotextiles are the product generally used for the function of filtration. Applications include geotextile filters for trench drains, blanket drains, interceptor drains, structural drains, toe drains in dams, filters for hard armor (e.g., rip-rap, gabions, fabric-form) erosion control systems, silt fences, and silt curtains. Both adequate hydraulic conductivity (provided by a geotextile with a relatively porous structure) and adequate soil retention (provided by a geotextile with a relatively tight structure) should be offered by the selected product. In addition, considerations should be made regarding the long-term soil-to-geotextile flow compatibility such that the flow through the geotextile will not be excessively reduced by clogging during the lifetime of the system. The geosynthetic-to-soil system should then achieve an equilibrium that allows for adequate liquid flow with limited soil loss across the geotextile throughout a service lifetime compatible with the application under consideration. Filtration concepts are well established in the design of soil filters, and similar concepts are used in the design of geotextile filters.

As the flow of liquid is perpendicular to the plane of the geosynthetic, filtration refers to the cross-plane hydraulic conductivity. Some of the geosynthetics used for this purpose are relatively thick and compressible. For this reason, geosynthetics are generally characterized by their permittivity, which is defined as:

$$\psi = k_n / t \quad (37.2)$$

where ψ is the permittivity, k_n is the cross-plane hydraulic conductivity, and t is the geosynthetic thickness at a specified normal pressure.

Testing procedures for geotextile permittivity follow similar guidelines used for testing the hydraulic conductivity of the soil. Some designers prefer to work directly with hydraulic conductivity and require the geotextile's hydraulic conductivity to be some multiple of the adjacent soil's hydraulic conductivity (Christopher and Fischer, 1992).

As the flow of liquid through the geotextile increases, the geotextile voids should be larger. However, large geotextile voids can lead to an unacceptable situation called soil piping, in which the soil particles are continuously carried through the geotextile, leaving large soil voids behind. The liquid velocity then increases, which accelerates the process and may lead to the collapse of the soil structure. This process can be prevented by selecting a geotextile with voids small enough to retain the soil on the upstream side of the fabric. It is the coarser soil fraction that must be initially retained. The coarser-sized particles eventually create a filter “bridge,” which in turn retains the finer-sized particles, building up a stable upstream soil structure (Figure 37.5).

Several methods have been developed for soil retention design using geotextiles; most of them compare the soil particle size characteristics to the 95% opening size of the geotextile (defined as O_{95} of the geotextile). The test method used in the United States to determine the geotextile opening size is called the apparent opening size (AOS) test (ASTM D 4751).

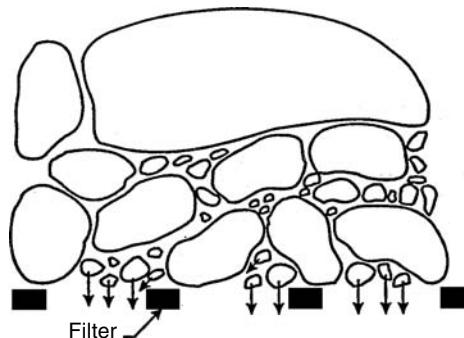


FIGURE 37.5 Geotextile providing adequate filtration through selection of adequate opening size and formation of soil filter bridge.

Some of the soil particles will rest on or embed them within the geotextile structure, and will cause a reduction in the hydraulic conductivity or permittivity of the geotextile. Although some partial clogging should be expected, the designer should ensure that the geotextile will not become excessively clogged, that is, that the flow of liquid will not be decreased to a point in which the system will not adequately perform its function. Thus, the geotextile voids should be large enough to allow the finer soil particles to pass. Clogging potential creates a special problem when geotextiles are used to wrap pipes due to the restricted area available for flow (i.e., a portion of the geotextile is covered by the pipe walls). Any clogging will significantly reduce the flow into the pipe. Either the flow capacity of the geotextile should be increased proportionally to the covered area (i.e., the total pipe area divided by geotextile area available for flow, which is usually the area of the holes in the pipe) or the geotextile should only be used to wrap gravel placed around the pipe. Design guidelines are available for clogging evaluation of non-critical, non-severe cases (Holtz et al., 1997, 1998; Koerner, 2005), but laboratory testing is strongly recommended in important applications. The gradient ratio test (ASTM D 5101), the long-term flow test (Halse et al., 1987), or the hydraulic conductivity ratio test (ASTM D 5567) should be performed. An evaluation of the filtration function of geotextiles is provided by Giroud (1996), Bhatia and Smith (1996a,b), Bhatia et al. (1996), Holtz et al. (1997, 1998), and Palmeira and Fannin (2002).

Proper construction is a critical factor in the performance of the filtration function. To develop and maintain the filter bridge, the geotextile must be placed in continuous (*intimate*) contact with the soil to be filtered such that no void spaces occur behind the geotextile. Placement granular material (e.g., gravel, drain rock, or rip rap) must also be consistent with the strength and durability of the geosynthetic to prevent damage during construction.

37.2.5 Drainage Function

Geosynthetics provide a drainage function by transmitting liquid within the plane of their structure. As shown in Table 37.2, the geosynthetics generally used for drainage purposes are geotextiles and geo-composites. The drainage function of geosynthetics allows for adequate liquid flow with limited soil loss within the plane of the geotextile over a service lifetime compatible with the application under consideration.

Thick, needle-punched nonwoven geotextiles have considerable void space in their structure and can convey liquid in their plane (on the order of 0.01 to 0.1 l/sec/m width of geotextile). Geocomposite drains can transmit one to two orders of magnitude more liquid than geotextiles. Proper design should dictate what type of geosynthetic drainage material is necessary.

The soil retention and the long-term compatibility considerations regarding the drainage function of geosynthetics are the same as those discussed in Section 37.2.4 regarding the filtration function of geosynthetics. As the geosynthetic thickness decreases with increasing normal stress, the in-plane drainage

of a geosynthetic is generally quantified by its transmissivity, which is defined as:

$$\theta = k_p \cdot t \quad (37.3)$$

where θ is the transmissivity, k_p is the in-plane hydraulic conductivity, and t is the geosynthetic thickness at a specified normal pressure. The in-plane flow capacity of geosynthetic will reduce under compression and with time (due to creep). Therefore, the transmissivity should be evaluated under normal pressures that are comparable to field conditions with a factor of safety included in the design to account for the design life of the project. Design guidance is provided by Holtz et al. (1997, 1998) and Koerner (2005).

Calculating the thickness of the liquid in or above a liquid collection layer is an important design step because one of the design criteria for a liquid collection layer is that the maximum thickness of the liquid collection layer must be less than an allowable thickness. The thickness of liquid in a liquid collection layer depends on the rate of liquid supply. A typical case of liquid supply is that of liquid impinging onto the liquid collection layer. Two examples of liquid collection layers with such a type of liquid supply can be found in landfills: (1) the drainage layer of the cover system, where the liquid that impinges onto the liquid collection layer is the precipitation water that has percolated through the soil layer overlying the drainage layer; and (2) the leachate collection layer, where the liquid that impinges onto the leachate collection layer is the leachate that has percolated through the waste and through the protective soil layer overlying the leachate collection layer. Equations are available (Giroud et al., 2000) to calculate the maximum thickness of liquid in a liquid collection layer located on a single slope with a perfect drain at the toe.

37.2.6 Barrier Function

The barrier function can be performed by geosynthetic products that have adequately low hydraulic conductivity as to provide containment to liquid or vapor. As shown in Table 37.2, the barrier function may be provided by several types of geosynthetics, namely, geomembranes and geosynthetic clay liners (GCLs). Other geosynthetic products also used as infiltration barriers include membrane-encapsulated soil layers (MESLs) used with paved or unpaved road construction, asphalt-saturated geotextiles used in the prevention of bituminous pavement crack reflection problems, and geofoam used for insulation against moisture and extreme temperatures.

Geosynthetic barriers are commonly used as liner for surface impoundments storing hazardous and nonhazardous liquids, as covers above the liquid surface of storage reservoirs, and as liner for canals used to convey water or chemicals. Geosynthetic barriers are also used as secondary containment for underground storage tanks, and in applications related to dams and tunnels. Of particular relevance for groundwater applications is the use of geosynthetic barriers for seepage control (HDPE vertical barrier systems). A common application of geosynthetics as infiltration barriers is base and cover liners for landfills. In landfill applications, infiltration barriers are typically used instead of (or in addition to) low-hydraulic conductivity soils. Base liners are placed below the waste to prevent liquids from the landfill (leachate) contaminating the underlying ground and the groundwater. Geosynthetic cover liner systems are placed above the final waste configuration to keep precipitation water from entering the waste and generating leachate. If a building or other structure is constructed on a landfill, a geosynthetic barrier may be placed under the building foundation to provide a barrier for vapors such as landfill gas. The use of geosynthetics in infiltration barriers is further described in Koerner (2005).

Dams are among the most critical hydraulic structures and stand to benefit the most from the use of geosynthetics. Deterioration and structural damage due to seepage are major concerns that have been addressed by placement of geomembrane liners. Geosynthetic systems as hydraulic barriers in dams are effective solutions to the problem of degradation. Accordingly, geosynthetics have been used in a wide range of dam types and dam sizes. The first installation of a geomembrane in a dam was in Italy in 1959. Since then, geosynthetics have been installed in dams worldwide as part of new projects and rehabilitation projects. Most recently, the Olivenhain Dam in San Diego was built with a geosynthetic system as an infiltration barrier. Significant advances have taken place in geosynthetics engineering since

geomembranes were first used in hydraulic structures. A good example is the current confidence on the extended service life of geomembranes. Additional information on the use of geosynthetics in dams is provided by and can be found in Christopher and Dahlstrand (1989), Zornberg and Weber (2003), Zornberg (2005), and Scuero et al. (2005).

37.2.7 Protection Function

Geosynthetics (mainly geotextiles) can be used to provide stress relief and protect other materials such as geosynthetics (mainly geomembranes) against damage. A common example is the use of geotextiles to provide protection against puncture of geomembranes in waste and liquid containment systems. Adequate mechanical protection must be provided to resist both short-term equipment loads and long-term loads imparted by the waste. Experience has shown that geotextiles can play an important role in the successful installation and long-term performance of geomembranes by acting as a cushion to prevent puncture damage of the geomembrane. In the case of landfill base liners, geotextiles can be placed (1) below the geomembrane to resist puncture and wear due to abrasion caused by sharp-edged rocks in the subgrade, and (2) above the geomembrane to resist puncture caused either by the drainage aggregate or direct contact with waste materials. In the case of landfill cover liners, geotextiles can be placed below the geomembrane to reduce risk of damage by sharp objects in the landfill and above the geomembrane to prevent damage during placement of drainage aggregate or cover soil. Key characteristics for the geotextile cushions are polymer type, mass density, method of manufacture, and construction survivability. The selection process of a geotextile that fulfills a protective function of a geomembrane involves the following three steps: (1) selection of polymer type and method of manufacture; (2) evaluation of the geotextile's capacity to provide puncture protection for the geomembrane; and (3) evaluation of construction survivability. Detailed procedures and methods for conducting these evaluations are described by Holtz et al. (1997, 1998), Koerner (2005), Narejo et al. (1996), and Wilson-Fahmy et al. (1996).

Another protection application associated with groundwater is erosion control blankets and mats. In this case, the geosynthetic protects the ground surface from the prevailing atmospheric conditions (i.e., wind, rain, snow, etc.). Specialty geocomposites have been developed for the specific purpose of erosion control. The general goal of these products is to protect soil slopes from both sheet and gully erosion, either permanently or until vegetation is established. Information on the design of geosynthetics for separation applications is provided by Holtz et al. (1997, 1998) and Koerner (2005).

37.3 Geosynthetic Types

As indicated in the introduction Section 37.1, there are a number of different geosynthetics. The characteristics of these materials vary considerably, primarily due to the method of manufacturing and the types and amount of polymers used for their production. This section provides a brief description of these materials and their primary characteristics. Additional information on the manufacturing process and characteristics of geosynthetics is provided by Koerner (2005) and Holtz et al. (1997, 1998) as well as by manufacturer's organizations, such as the Geosynthetic Manufacturer's Association (www.gmanow.com) and the PVC Geomembrane Institute (pgi-tp.ce.uiuc.edu).

37.3.1 Geotextiles

Among the different geosynthetic products, geotextiles are the ones that present the widest range of properties. They can be used to fulfill all the different functions listed in Table 2.2 for many different geotechnical, environmental, and hydraulic applications. For example, Figure 37.6 shows the construction of a reinforced slope in which geotextiles were selected as multipurpose inclusions within the fill, as they can provide not only the required tensile strength (reinforcement function), but also the required transmissivity (drainage function), needed for that particular project (Zornberg et al., 1995).



FIGURE 37.6 Placement of a high-strength nonwoven geotextile to perform a dual function of reinforcement and in-plane drainage in a reinforced slope.

Geotextiles are manufactured from polymer fibers or filaments that are later formed to develop the final product. Approximately 85% of the geotextiles used today are based on polypropylene resin. An additional 10% are polyester and the remaining 5% is a range of polymers including polyethylene, nylon, and other resins used for specialty purposes. As with all geosynthetics, however, the base resin has various additives, such as for ultraviolet light protection and long-term oxidative stability.

The most common types of fibers or filaments used in the manufacture of geotextiles are monofilament, multifilament, staple filament, and slit-film. If fibers are twisted or spun together, they are known as a yarn. Monofilaments are created by extruding the molten polymer through an apparatus containing small-diameter holes. The extruded polymer strings are then cooled and stretched to give the filament increased strength. Staple filaments are also manufactured by extruding the molten polymer; however, the extruded filaments are cut into 25- to 100-mm portions. The staple filaments or fibers may then be spun into longer yarns. Slit-film filaments are manufactured by either extruding or blowing a film of a continuous sheet of polymer and cutting it into filaments by knives or lanced air jets. Slit-film filaments have a flat, rectangular cross-section instead of the circular cross-section shown by the monofilament and staple filaments.

The filaments, fibers, or yarns are formed into geotextiles using either woven or nonwoven methods. Figure 37.7 shows a number of typical woven and nonwoven geotextiles. Woven geotextiles are manufactured using traditional weaving methods and a variety of weave types. Nonwoven geotextiles are manufactured by placing and orienting the filaments or fibers onto a conveyor belt, which are subsequently bonded by needle punching or by melt bonding. The needle-punching process consists of pushing numerous barbed needles through the fiber web. The fibers are thus mechanically interlocked into a stable configuration. As the name implies, the heat (or melt) bonding process consists of melting and pressurizing the fibers together.

Common terminology associated with geotextiles includes machine direction, cross machine direction, and selvage. Machine direction refers to the direction in the plane of fabric in line with the direction of

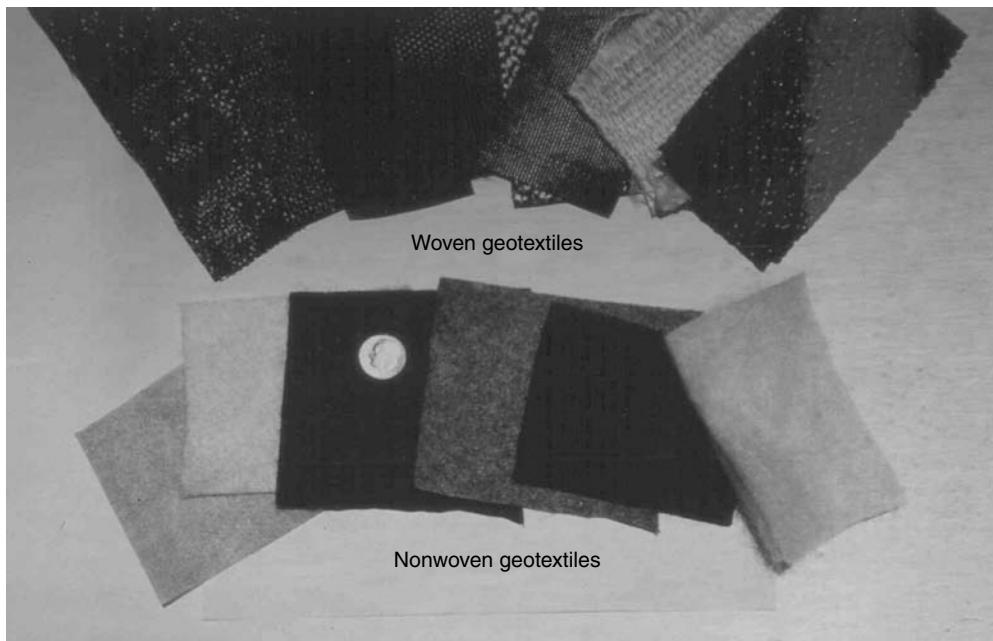


FIGURE 37.7 Typical woven and nonwoven geotextiles.

manufacture. Conversely, cross machine direction refers to the direction in the plane of fabric perpendicular to the direction of manufacture. The selvage is the finished area on the sides of the geotextile width that prevents the yarns from unraveling. Adjacent rolls of geotextiles are seamed in the field by either overlapping or sewing. Sewing is generally the case for geotextiles used as filters in landfill applications, but may be waived for geotextiles used in separation. Heat bonding may also be used for joining geotextiles used in filtration and separation applications.

Numerous tests have been developed to evaluate the properties of geotextiles. In developing geotextile specifications, it is important that the designer understands the material tests and that material properties important for the geotextile's intended use are specified. Table 37.3 describes the tests commonly performed in geotextile products (ASTM, 1995). Geotextiles can generally be specified with index and performance properties. As previously discussed, critical filtration applications are exceptions and the use of pre-qualified specific products from performance flow tests are strongly recommended.

37.3.2 Geomembranes

Geomembranes are flexible, polymeric sheets that have very low hydraulic conductivity (typically less than 10^{-11} cm/sec) and, consequently, are used as liquid or vapor barriers. The most common types of geomembranes are high density polyethylene (HDPE), very flexible polyethylene (VFPE), polyvinyl chloride (PVC), flexible polypropylene (fPP) and reinforced chlorosulfonated polyethylene (CSPE). Figure 37.8 shows a number of geomembranes currently available in the geosynthetics market.

Polyethylene is the type of geomembrane most commonly used in landfill applications for base and cover liner systems. This is primarily because of its high chemical resistance and durability. Specifically, high-density polyethylene (HDPE) is typically used in base liner systems. This material is somewhat rigid but generally has good physical properties and can withstand large stresses often imposed on the geomembrane during construction.

VFPE and PVC are the most commonly used geomembrane materials besides HDPE. The term VFPE encompasses various polyethylene grades such as very low density polyethylene (VLDPE) and certain types of linear low density polyethylene (LLDPE). The linear structure and lack of long-chain branching

TABLE 37.3 Standard Tests for Geotextiles

Property	Test standard	Test name
Thickness	ASTM D 5199	Standard Test Method for Measuring Nominal Thickness of Geotextiles and Geomembranes
Mass per Unit Area	ASTM D 5261	Standard Test Method for Measuring Mass per Unit area of Geotextiles
Grab rupture	ASTM D 4632	Standard Test Method for Breaking Load and Elongation of Geotextiles (Grab Method)
Uniaxial tensile strength geotextiles	ASTM D 4595	Standard Test Method for Tensile Properties of Geotextiles by the Wide-Width Strip Method
Geogrids	ASTM D 6637	Standard Test Method for Determining Tensile Properties of Geogrids by the Single or Multi-Rib Tensile Method
Multiaxial tensile, puncture or burst tests	ASTM D 6241	Standard Test Method for the Static Puncture Strength of Geotextiles and Geotextile Related Products Using a 50-Mm Probe
Trapezoid tear strength	ASTM D 4533	Standard Test Method for Trapezoid Tearing Strength of Geotextiles
Apparent opening size	ASTM D 4751	Standard Test Method for Determining Apparent Opening Size of a Geotextile
Permittivity	ASTM D 4491	Standard Test Methods for Water Permeability of Geotextiles by Permittivity
Gradient ratio	ASTM D 5101	Standard Test Method for Measuring the Soil-Geotextile System Clogging Potential by the Gradient Ratio
Transmissivity	ASTM D 4716	Standard Test Method for Constant Head Hydraulic Transmissivity (In-Plane Flow) of Geotextiles and Geotextile Related Products
Ultraviolet resistance	ASTM D 4355	Standard Test Method for Deterioration of Geotextiles from Exposure to Ultraviolet Light and Water (Xenon-Arc Type Apparatus)
Seam strength	ASTM D 1683	Failure in Sewn Seams of Woven Fabrics
Seam strength	ASTM D 4884	Standard Test Method for Seam Strength of Sewn Geotextiles

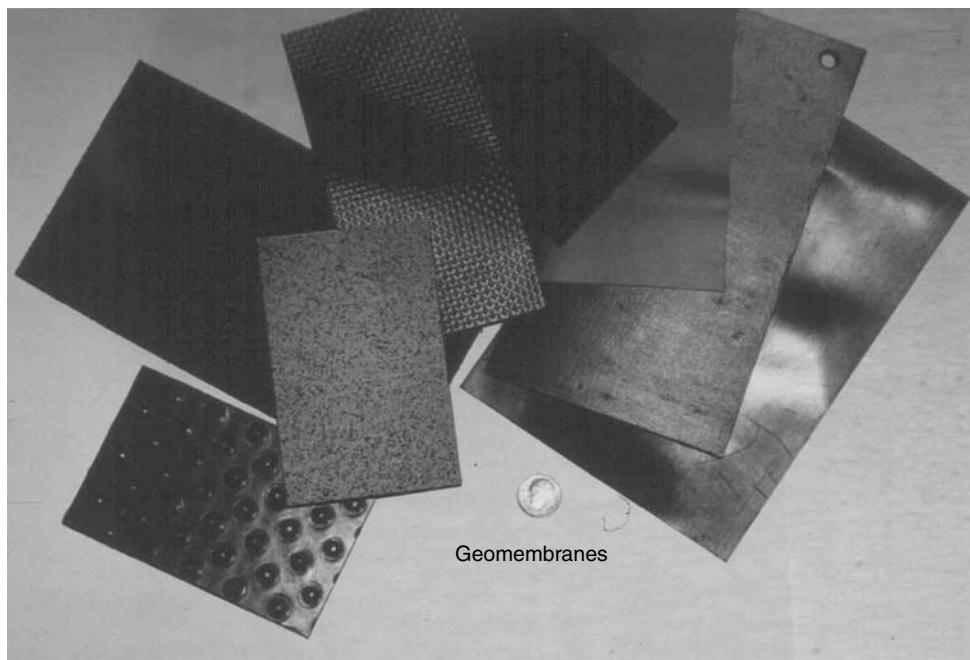
**FIGURE 37.8** Typical geomembranes.

TABLE 37.4 Criteria for Selection of HDPE, PVC, or CSPE Geomembranes

Criteria	Considerations for selection
Liquid barrier	All three polymers have acceptable characteristics as liquid barriers, although HDPE geomembranes have the best characteristics in terms of chemical resistance and long-term durability. All three have extremely low hydraulic conductivity and are impermeable for practical purposes.
Mechanical properties	Although the mechanical properties vary somewhat with geomembrane thickness, HDPE is relatively stiff and has relatively small yield strain. PVC, in contrast, is relatively extensible and does not exhibit yield. The tensile properties of CSPE often fall between those of HDPE and PVC but are difficult to generalize because CSPE is usually made with embedded reinforcing fabrics which affect tensile response.
Construction survivability	All three polymers have an acceptable ability to maintain integrity when subjected to concentrated stresses. However, the best performance is obtained with more extensible geomembranes. Therefore, based on the relative extensibility, PVC generally offers the most favorable performance.
Installation	Key considerations include ease of placement and seaming. PVC and CSPE are easier to place than HDPE because their greater flexibility makes them conform more easily to the foundation and makes them less prone to thermal expansion wrinkles. Acceptable placement and wrinkle control, however, can be achieved with all three polymers if appropriate installation procedures are used. All three polymers are easily seamed, with HDPE usually achieving the highest seam strength and quality.
Chemical resistance	HDPE has the highest degree of compatibility with a wide variety of chemicals encountered in wastes. CSPE has good resistance to many chemicals, but is attacked by some that are relatively common, namely chlorinated solvents and hydrocarbons. PVC typically is the least chemically resistant of these polymers.
Long-term durability	HDPE offers the best performance. HDPE is a highly inert and durable material that is not susceptible to chemical degradation under conditions generally encountered in landfills. In addition, HDPE is not susceptible to physical degradation (extraction). The durability of PVC geomembranes is less favorable than that of HDPE. This is because PVC geomembranes are composed of approximately two-thirds PVC resin and one-third plasticizers. Over time, physical degradation (extraction) may cause plasticizer loss which results in reduced geomembrane flexibility. The durability of CSPE geomembranes is typically between that of HDPE and PVC.

in both LLDPE and VLDPE arise from their similar polymerization mechanisms. Due to the large settlements that may occur, cover liner systems commonly require a flexible geomembrane. VFPE is often used in this application as it provides similar chemical resistance as HDPE but is more flexible and can more readily conform to underlying refuse settlements without puncturing. Flexible polypropylene (fPP) geomembranes have recently (1990) been developed by combining the polypropylene (PP) and the ethylene propylene elastomer (EPE). This product was developed to obtain physical characteristics similar to the low density polyethylene while offering the flexibility similar to the polyvinyl chloride geomembranes.

PVC geomembranes are used in liners for many waste containment applications, such as contaminated soils containment and liquid storage ponds. While PVC may not be as durable as polyethylene geomembranes, the merits of PVC geomembranes are that they are generally less expensive than polyethylene geomembrane and can be factory manufactured in relatively large panels. The large panel sizes allow easier installation as there are fewer field fabricated seams.

In landfill applications, geomembranes are typically used as a base or a cover liner in place of or in addition to low-hydraulic conductivity soils. The key performance factors related to the selection of geomembrane polymer types for landfill applications are summarized in Table 37.4. Geomembrane thickness ranges from 0.75 to 2.5 mm (30 to 100 mils). Table 37.5 summarizes the key performance factors related to

TABLE 37.5 Criteria for Minimum Thickness of HDPE Geomembrane

Criteria	Considerations for selection of thickness
Abrasion resistance	The abrasion resistance of HDPE geomembranes increases with geomembrane thickness. Experience indicates that geomembranes with thickness < 1 mm may not have acceptable abrasion resistance.
Response to differential settlements	The thicker the HDPE geomembrane, the higher its stiffness. This issue is more significant for geomembrane cover systems than for geomembrane base liner systems because the cover system must be flexible enough to accommodate differential settlements. From this viewpoint, a thickness of not more than 2 mm is desirable.
Effective welding	The thinner the HDPE geomembrane, the more difficult is the welding of adjacent panels. For most effective welding, a desirable thickness is of at least 1 mm (40 mils), and preferably 1.5 to 2 mm.

the selection of the thickness of HDPE geomembranes for landfill applications. Geomembranes are placed after subgrade preparation, and placement is followed by seaming, inspection, and backfilling. A properly designed geomembrane has the potential of hundreds of years of service lifetime, but installation must follow high quality management principles. In the early uses of geomembranes for waste containment applications, the main concerns were related to the chemical compatibility between geomembranes and waste, and to the service life of geomembranes. Now, construction quality issues are viewed as the principal limitations to the performance of geomembranes.

For continuity of the impermeable barrier, geomembranes should be seamed in the field. The fundamental mechanism of seaming polymeric geomembrane sheets together is to temporarily reorganize (melt) the polymer structure of the two surfaces to be joined in a controlled manner. This reorganization can be done either through thermal or chemical processes. These processes may involve the addition of extra polymer in the bonded area. There are four general categories of seaming methods: extrusion welding, thermal fusion or melt bonding, chemical fusion, and adhesive seaming. Thermal fusion and extrusion welding are the methods most commonly used, and are described next.

In thermal fusion or melt bonding (the most common seaming method), portions of the opposing surfaces are truly melted. Temperature, pressure, and seaming rate play important roles as excessive melting weakens the geomembrane and inadequate melting results in low seam strength. The hot wedge, or hot shoe, method consists of an electrically heated resistance element in the shape of a wedge that travels between the two sheets to be seamed. A standard hot wedge creates a single uniform width seam, while a dual hot wedge (or "split" wedge) forms two parallel seams with a uniform unbonded space between them. This space can then be conveniently used to evaluate seam quality and continuity by pressurizing the unbonded space with air and monitoring any drop in pressure that may signify a leak in the seam (Figure 37.9).

Extrusion welding is at present used exclusively on geomembranes made from polyethylene. A ribbon of molten polymer is extruded over the edge of, or in between, the two surfaces to be joined. The molten extrudate causes the surface of the sheets to become hot and melt, after which the entire mass cools and bonds together. The technique is called extrusion fillet seaming when the extrudate is placed over the leading edge of the seam, and is called extrusion flat seaming when the extrudate is placed between the two sheets to be joined. Fillet extrusion seaming is essentially the only practical method for seaming polyethylene geomembrane patches, for seaming in poorly accessible areas such as sump bottoms and around pipes, and for seaming of extremely short seam lengths.

The material properties of geomembranes are divided into the properties of the raw polymer or resin used in the manufacture of the geomembrane sheet and the manufactured geomembrane properties. Table 37.6 lists the tests commonly performed for evaluation of the raw polymer properties. Table 37.7 summarizes the tests commonly performed to evaluate the manufactured geomembrane sheet properties (ASTM, 1995). As with the geotextiles, many of these tests provide index or quality control properties.



FIGURE 37.9 Monitoring seaming of a geomembrane liner.

TABLE 37.6 Tests for Raw Geomembrane Polymers

Property	Test standard	Test name
Density	ASTM D 792	Standard Test Method for Specific Gravity and Density of Plastics by the Density-Gradient Technique
Density	ASTM D 1505	Standard Test Method for Density of Plastics by the Density-Gradient Technique
Melt index	ASTM D 1238	Standard Test Method for Flow Rates of Thermoplastics by Extrusion Plastometer
Chemical identification methods (fingerprinting)		Thermogravimetric Analysis (TGA) Differential Scanning Calorimetry (DSC) Thermomechanical Analysis (TMA) Infrared Spectroscopy (IR) Chromatography (GC) Gel Permeation Chromatography (GPC)

37.3.3 Geogrids

Geogrids constitute a category of geosynthetics designed preliminary to fulfill a reinforcement function. Geogrids have a uniformly distributed array of apertures between their longitudinal and transverse elements. The apertures allow direct contact between soil particles on either side of the installed sheet, thereby increasing the interaction between the geogrid and the backfill soil.

Geogrids are composed of polypropylene, polyethylene, polyester, or coated polyester. They are formed by several different methods. The coated polyester geogrids are typically woven or knitted. Coating is generally performed using PVC or acrylics to protect the filaments from construction damage and to maintain the grid structure. The polypropylene geogrids are either extruded or punched sheet drawn, and polyethylene geogrids are exclusively punched sheet drawn. Figure 37.10 shows a number of typical geogrid products. In the past several years, geogrids have also been manufactured by interlacing polypropylene or polyester strips together and welding them at their cross-over points (i.e., junctions).

TABLE 37.7 Standard Tests for Geomembranes

Property	Test standard	Test name
Thickness	ASTM D 5199	Standard Test Method for Measuring Nominal Thickness of Geotextiles and Geomembranes
Tensile behavior	ASTM D 6693	Test Method for Determining Tensile Properties of Nonreinforced Polyethylene and Nonreinforced Flexible Polypropylene Geomembranes
Tensile behavior	ASTM D 7003	Test Method for Strip Tensile Properties of Reinforced Geomembranes
Tensile behavior	ASTM D 882	Test Methods for Tensile Properties of Thin Plastic Sheeting (Method A Used for PVC)
Tensile behavior	ASTM D 4885	Standard Test Method for Determining Performance Strength of Geomembranes by the Wide Strip Tensile Method
Tear resistance	ASTM D 1004	Test Method for Initial Tear Resistance of Plastic Film and Sheeting
Puncture resistance	ASTM D 2065	Puncture Resistance and Elongation Test
Puncture resistance	ASTM D 5494	Test Method for the Determination of Pyramid Puncture Resistance of Unprotected and Protected Geomembranes
Puncture resistance	ASTM D 5514	Test Method for Large-Scale Hydrostatic Puncture Testing of Geosynthetics
Environmental stress crack	ASTM D 1693	Standard Test Method for Environmental Stress-Cracking of Ethylene Plastics
Environmental stress crack	ASTM D 5397	Test Method for Evaluation of Stress Crack Resistance of Polyolefin Geomembranes Using Notched Constant Tensile Load Test
Carbon black	ASTM D 1603	Standard Test Method for Carbon Black in Olefin Plastics
Carbon black	ASTM D 5596	Standard Practice for Microscopic Evaluation of the Dispersion of Carbon Black in Polyolefin Geosynthetics
Durability	ASTM D 5721	Practice for Air-Oven Aging of Polyolefin Geomembranes
chemical resistance	ASTM D 5747	Practice for Tests to Evaluate the Chemical Resistance of Geomembranes to Liquids
Seam strength	ASTM D 4437	Standard Practice for Determining the Integrity of Field Seams Used in Joining Flexible Polymeric Sheet Geomembranes
Leak detection	ASTM D 7007	Practices for Electrical Methods for Locating Leaks in Geomembranes Covered with Water or Earth Materials

Although geogrids are used primarily for reinforcement, some products are used for asphalt overlay and some are combined with other geosynthetics to be used in water proofing or in separation and stabilization applications. In waste containment systems, geogrids may be used to support a lining system over a weak subgrade or to support final landfill cover soils on steep refuse slopes. Geogrids are also used for support of liners in the design of “piggyback” landfills, which are landfills built vertically over older, usually unlined landfills. Regulatory agencies often require that a liner system be installed between the old and new landfill. As the old refuse is highly compressible, it provides a poor base for the new lining system. A geogrid may be used to support the lining system and bridge over voids that may occur beneath the liner as the underlying refuse components decompose.

As with other geosynthetics, geogrids have several physical, mechanical, and durability properties. Many of the test methods used for geotextiles and geomembranes also apply to geogrids. In particular, a key design parameter for reinforcement is the tensile strength, which is performed with the specimen width incorporating typically a few ribs of the geogrid. The allowable tensile strength of geogrids (and of other geosynthetics used for soil reinforcement applications) is typically significantly less than its ultimate tensile strength. The allowable tensile strength is determined by dividing the ultimate tensile strength by partial factors to account for installation damage, creep deformation, chemical degradation, and biological degradation. The partial factors for installation damage, chemical degradation, and biological degradation range from 1.0 to 1.6, with the partial factor of safety for creep ranging from 1.5 to 3.5 (Koerner, 2005). The design engineer should review all recommendations, available codes, national evaluations, and the like and determine appropriate partial factors for the specific projects (i.e., manufacturers' recommendations should be carefully evaluated based on project-specific requirements and not blindly accepted).

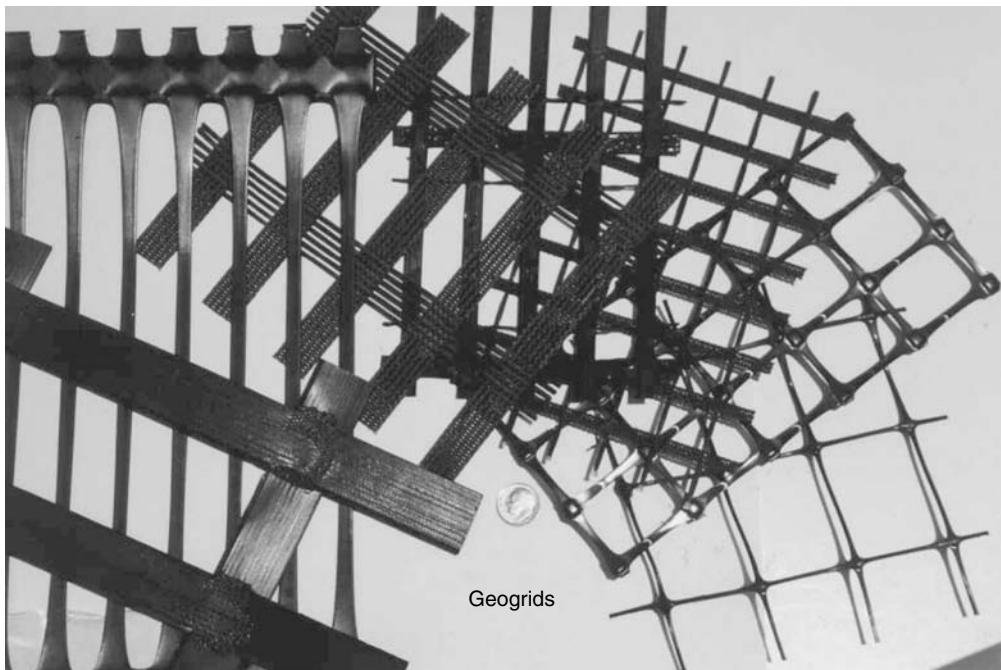


FIGURE 37.10 Typical geogrids.

37.3.4 Geosynthetic Clay Liners

Geosynthetic clay liners (GCLs) are rapidly expanding products in the geosynthetics market. GCLs are infiltration barriers consisting of a layer of unhydrated, loose granular or powdered bentonite placed between two or on top of one geosynthetic layer (geotextile or geomembrane). GCLs are produced in panels that are joined in the field by overlapping. They are generally used as an alternative to compacted clay liners (Bouazza, 2002).

Due to the inherent low shear strength of hydrated bentonite, GCL usage had initially been limited to applications where stability of the overlying materials was not a concern. In the late 1980s, however, methods were developed to reinforce the GCLs, producing a composite material with higher shear strength properties. This allowed the use of GCLs in landfill applications (Figure 37.11).

Some advantages of GCLs over compacted clay liner are that they occupy significantly less space to achieve equivalent performance, plus they are flexible, self-healing, and easy to install. In locations where low hydraulic conductivity clays are not readily available, they may offer significant construction cost savings. In addition, as they are factory manufactured with good quality control, field construction quality assurance costs are typically less than with compacted clay liners.

Bentonite is a clay formed primarily from the mineral montmorillonite. While several types of montmorillonite exist, including calcium and sodium montmorillonite, the term bentonite typically refers to a sodium montmorillonite. Water is strongly attracted to the surface of the negatively charged montmorillonite crystal and is readily absorbed by it. In its unhydrated state, the montmorillonite crystals are densely packed. Once hydrated, the structure becomes very open and swells. The high water absorption and swell characteristics of bentonite lead to its low hydraulic conductivity and low hydrated shear strength.

Geosynthetic clay liners are manufactured by laying down a layer of dry bentonite, approximately 5 mm thick, on a geosynthetic material and attaching the bentonite to the geosynthetic. Two general configurations are currently employed in commercial processes (Figure 37.12): bentonite sandwiched between two geotextiles or bentonite glued to a geomembrane. The primary purpose of the geosynthetic component



FIGURE 37.11 Installation of a GCL during construction of a landfill base liner.

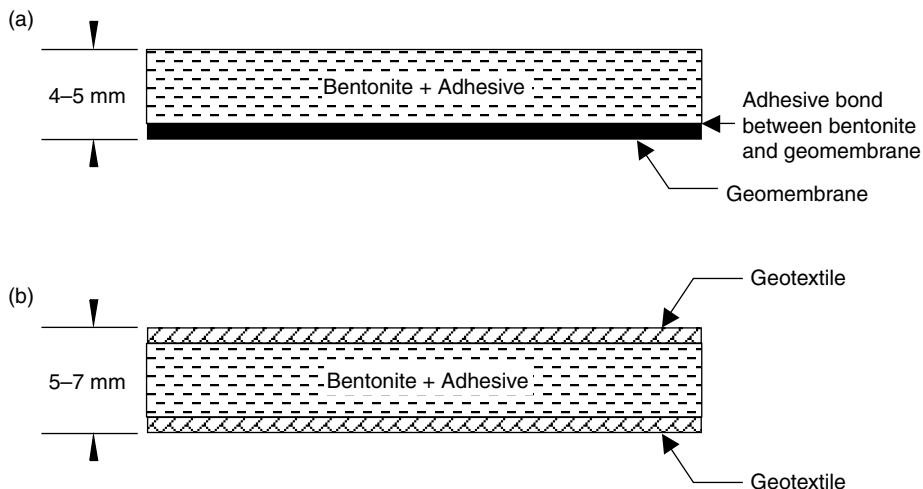


FIGURE 37.12 GCL configurations: (a) bentonite glued to a geomembrane; (b) bentonite sandwiched between two geotextiles.

is to hold the bentonite together in a uniform layer and to permit transportation and installation of the GCL without loss of bentonite.

The outer geosynthetic layer of GCLs can be mechanically bonded using stitching or needle punching (resulting in reinforced GCLs). A different process consists in using an adhesive bond to glue the bentonite to the geosynthetic (resulting in unreinforced GCLs). The mechanical bonding of reinforced GCLs increases their internal shear strength. Geosynthetic clay liners contain approximately 5 kg/m^2 of bentonite that has a hydraulic conductivity of approximately $1 \times 10^{-9} \text{ cm/sec}$. Infiltration under unit

hydraulic gradient through a material with hydraulic conductivity of 1×10^{-9} cm/sec would result in an infiltration rate of 0.3 mm per year.

As a GCL is a composite material, its relevant properties are those of the geotextile alone, of the bentonite alone, and of the composite. Geotextile properties were discussed in Section 37.3.1. The geotextile properties relevant to GCLs include mass per unit area, grab tensile, wide-width tensile, and puncture resistance. Relevant properties of the bentonite are obtained from free swell tests, which measure the absorption of water into a bentonite based on its volume change, and plate water absorption tests, which measure the ability of powdered bentonite to absorb water. The relevant properties of the composite GCL material include bentonite content, which is simply a measure of the mass of bentonite per unit area of GCL, permeameter testing (ASTM D 5084) used to estimate the GCL hydraulic conductivity, tensile strength characterized either by grab tensile or wide-width tensile tests, and puncture resistance tests performed to assess the relative puncture resistance between GCLs and geomembranes or other geosynthetics. The internal and interface shear strength of GCLs should also be determined as stability is a major concern for side slopes in bottom liner or cover systems that include GCLs. This is because of the very low shear strength of hydrated sodium bentonite. Proper shear strength characterization is needed for the different materials and interfaces in hydraulic barriers (Koerner, 2005). The analysis of a large database of test results on the internal shear strength of GCLs is presented by Zornberg et al. (2005).

37.3.5 Geocomposite Sheet Drains

A geocomposite consists of a combination of different types of geosynthetics. In particular, the geosynthetics industry has developed a number of geocomposite drains, which are polymeric drainage cores with continuously open flow channels sandwiched between geotextile filters. Geocomposite sheet drains are discussed in this section while geocomposite strip (wick) drains are discussed in Section 37.3.6.

Geocomposite sheet drainage systems have been engineered to replace costly aggregate and perforated pipe subsurface drainage systems. They have reached rapid acceptance because they provide adequate drainage and reduce the material cost, installation time, and design complexity of conventional aggregate systems.

The core of geocomposite sheet drains are extruded sheets of plastic formed into a configuration that promotes drainage. The core of the geocomposite sheet drains are most commonly composed of polyethylene but may also be composed of polypropylene, polystyrene, high-impact polystyrene, or other materials. The structure of the core drainage products ranges from a dimpled core to a geonet. Geonets, a commonly used drainage product, generally consist of two or three sets of parallel solid or foamed extruded ribs that intersect at a constant angle to form an open net configuration. Channels are formed between the ribs to convey either liquid or gases. Figure 37.13 shows a number of geonets currently available in the market. Dimpled cores type products tend to have a greater flow capacity than geonets; however, they generally have a much lower crush resistance and tend to be more compressible than geonets. As indicated in Section 37.2.5, the transmissivity of the geocomposite drain must be evaluated under anticipated site-specific normal pressures and a factor of safety included to account for creep of the product over the life of the system (Koerner, 2005).

The geotextile serves as both a separator and a filter, and the geonet or built-up core serves as a drain. There may be geotextiles on both the top and bottom of the drainage core and they may be different from one another. For example, the lower geotextile may be a thick needlepunched nonwoven geotextile used as a protection material for the underlying geomembrane, while the top geotextile may be a thinner nonwoven or woven product. Composite drainage nets are typically formed by thermally bonding the geotextile and geonet. Gluing and solvent welding can also be used to bond the geosynthetic core to the geotextile. In producing geocomposite drainage nets, the melt temperatures of the geotextile and geonet must be compatible so that the properties of each material are retained. Figure 37.14 shows a number of available geocomposite sheet drainage materials.

As the purpose of the core is drainage, the most important properties to include in specifications are thickness, crush strength, and long-term transmissivity under load. Table 37.8 summarizes the tests

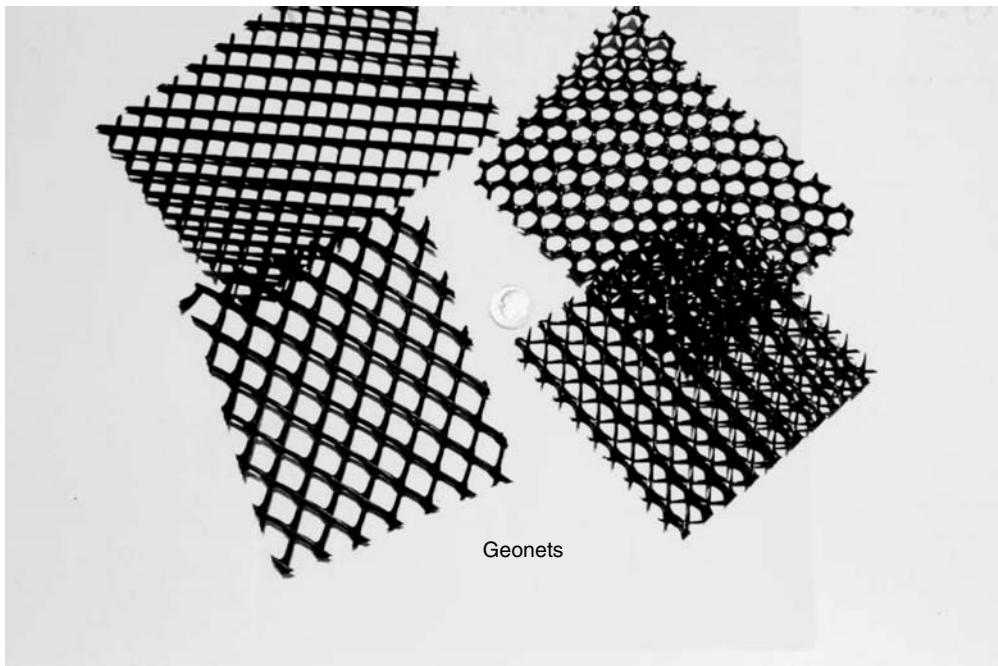


FIGURE 37.13 Typical geonets used as the core of geocomposite sheet drains.

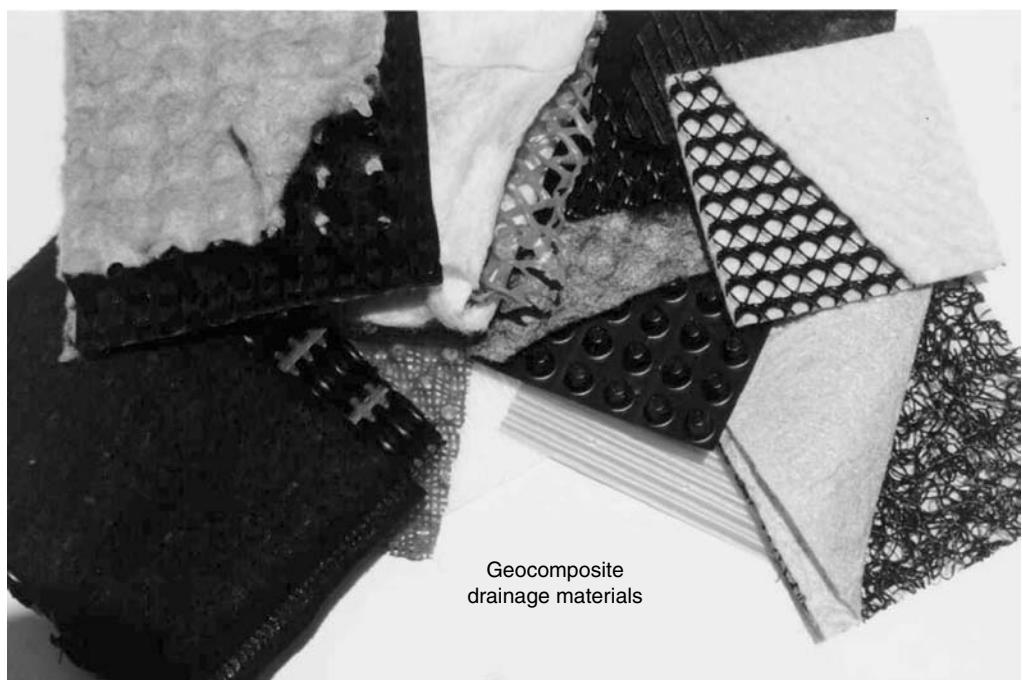


FIGURE 37.14 Typical geocomposite sheet drains.

TABLE 37.8 Standard Tests for Geocomposite Drainage Nets

Property	Test standard	Test name
Composite and core thickness	ASTM D 5199	Standard Test Method for Measuring Nominal Thickness of Geotextiles and Geomembranes
Core crush strength	ASTM D 1621	Standard Test Method for Compressive Properties of Rigid Cellular Plastics
Composite compression	ASTM D 6244	Test Method for Vertical Compression of Geocomposite Pavement Panel Drains
Composite compression	ASTM D 6364	Test Method for Determining Short-Term Compression Behavior of Geosynthetics
Composite transmissivity	ASTM D 4716	Standard Test Method for Determining the (In-plane) Flow Rate per Unit Width and Hydraulic Transmissivity of a Geosynthetic Using a Constant Head
Core carbon black	ASTM D 4218	Standard Test Method for Determination of Carbon Black Content in Polyethylene Compounds by the Muffle Furnace Technique
Geotextile filter	See Table 37.3	Grab Tensile Tear Strength CBR Puncture Strength AOS Permittivity

commonly performed to evaluate the properties of geocomposite sheet drains. It is also important to evaluate filtration requirements for the geotextile. Design and specification of composite drains is covered by Giroud et al. (2000), Koerner (2005), and Holtz et al. (1997, 1998).

37.3.6 Geocomposite Strip (Wick) Drains

Geocomposite strip drains, also called “wick drains,” have been developed to replace the use of sand drains in applications involving the increase in consolidation rate of soft, saturated fine-grained soils. Geocomposite strip drains actually do not wick moisture, but simply provide a conduit for excess pore water pressure induced flow. They are placed vertically through high water content silts and clays to produce shortened drainage paths and thus increase the rate of consolidation. Other names commonly used for these products are “band shaped drains” and “prefabricated vertical drains.”

Sand drains were originally introduced in the 1930s as a method for improvement of soft soil foundations. The method of rapid consolidation of saturated fine-grained soils using sand drains involves placement of vertical columns of sand (usually 200 to 450 mm in diameter) at spacings of 1.5 to 6.0 m centers throughout the subsurface to be dewatered. Due to the low installed product cost and speed of installation, the use of geocomposite strip drains dominates over the use of sand drains in projects involving dewatering of saturated fine-grained soils. Their lengths are site-specific and extend to the bottom of the soft layer, and, thus, lengths generally will vary along a project. Once installed, a surcharge load is placed on the ground surface to mobilize excess pore water pressures. This surcharge load is placed in incremental lifts, which induce pore water pressure in the underlying soil. The pore water pressure is then dissipated through the vertical drains. Water takes the shortest drainage path (i.e., horizontally radial) to the vertical drain, at which point it flows vertically as the drain has a much higher hydraulic conductivity than the fine-grained soil being consolidated. The rate at which surcharge fill is added is critical in this process.

Most commercially available geocomposite strip drains have adequate capacity to drain the water expelled during consolidation of the fine-grained soils. As their flow capacity is usually adequate, selection of the vertical drains is governed by the consolidation rate required in the project. Hansbo's equation (Hansbo, 1979) is generally used to estimate the time required to achieve a desired percentage of consolidation as a function of the horizontal coefficient of consolidation of the foundation soil, the equivalent diameter of the geocomposite strip drain, and the spacing of the drains. As with geocomposite sheet drains,



FIGURE 37.15 View of expanded geocell (photo courtesy of Presto Product Company).

the primary function of the geotextile covering is filtration. Determining the filtration requirements for the geotextile is an essential element of the design.

Installation of geocomposite strip drains is very rapid and uses lightweight construction equipment fitted with hollow leads (called a “lance” or mandrel) for insertion to the desired depth. The bottom of the lance should be covered by an expendable shoe that keeps soil out of the lance so as not to bind the strip drain within it. The allowable flow rate of geocomposite strip drains is determined by the ASTM D4716 test method. Typical values of ultimate flow rate at a hydraulic gradient of 1.0 under 207 kPa normal stress vary from 1.5 to 3.0 $\text{m}^3/\text{sec.-m}$. This value must then be reduced on the basis of site-specific partial factors of safety. Specifications for geocomposite strip drains are provided by Holtz et al. (1997).

37.3.7 Geocells

Geocells (or cellular confinement systems) are three-dimensional, expandable panels made from strips, typically 50 to 100 mm wide. When expanded during installation, the interconnected strips form the walls of a flexible, three-dimensional cellular structure into which specified infill materials are placed and compacted (Figure 37.15). This creates a system that holds the infill material in place and prevents mass movements by providing tensile reinforcement. Cellular confinement systems improve the structural and functional behavior of soil infill materials.

Geocells were developed in the late 1970s and early 1980s for support of military vehicles on weak subgrade soils. The original type of geocell consists of HDPE strips 200 mm wide and approximately 1.2 mm thick. They are ultrasonically welded along their 200 mm width at approximately 330 mm intervals and are shipped to the job site in a collapsed configuration. At the job site they are placed directly onto the subgrade surface and propped open in an accordion-like fashion with an external stretcher assembly. They are then filled generally with gravel or sand (although other infill materials such as concrete, can be used) and compacted using a vibratory hand-operated plate compactor. Geocell applications include protection and stabilization of steep slope surfaces, protective linings of channels and hydraulic structures, static

and dynamic load support on weak subgrade soils, and multilayered earth-retaining and water-retaining gravity structures.

Geocells have proven very effective in providing a stable foundation over soft soils. The cellular confinement system improves the load-deformation performance of infill materials because cohesionless materials gain considerable shear strength and stiffness under confined conditions. Confining stresses are effectively induced in a geocell by means of the hoop strength developed by the HDPE cell walls. The overall increase in the load carrying performance of the system is provided through a combination of the cell wall strength, the passive resistance of the infill material in adjacent cells, and the frictional interaction between the infill soil and the cell walls. The cellular structure distributes concentrated loads to surrounding cells thus reducing the stress on the subgrade directly beneath the geocell.

Infill selection is primarily governed by the nature and intensity of anticipated working stresses, availability and cost of candidate materials, and aesthetic requirements for a fully vegetated appearance. Aggregates, vegetated topsoil, and concrete constitute typical geocell infill types. A complete cellular confinement system may also include geotextiles, geomembranes, geonets, geogrids, integral polymeric tendons, erosion-control blankets, and a variety of earth anchors.

37.3.8 Erosion Control Products

Erosion-control products represent one of the fastest growing application areas in the geosynthetics industry. Erosion-control products provide protection against sheet and gully erosion on soil slopes either until vegetation is established or for long-term applications. These products can be classified as temporary degradable erosion control blankets, long-term nondegradable erosion control mats, and permanent hard armored systems.

Temporary degradable erosion control blankets are used to enhance the establishment of vegetation. These products are used where vegetation alone would provide sufficient site protection once established after the erosion control product has degraded. Some of these products are completely biodegradable (e.g., straw, hay, jute, and hydraulic mulches), while others are only partially biodegradable (e.g., erosion control meshes and nets).

Long-term nondegradable erosion control mats provide permanent reinforcement of vegetation root structure. They are used in critical erosion-control applications where immediate high-performance erosion protection, followed by the permanent reinforcement of established vegetation is required. These soft armor related products provide erosion control, aid in vegetative growth, and eventually become entangled with the vegetation to provide reinforcement to the root system.

Finally, the permanent hard armored systems include riprap on geotextile filters, modular concrete block over geotextile filter systems, gabions over geotextile filters, geocell products with gravel or concrete infill, vegetated concrete block systems, and fabric-formed revetments.

Figure 37.16 shows an erosion control mat installed to help vegetation establishment on a steep reinforced soil slope. Installation of flexible erosion control products is straightforward. The products are usually placed on a prepared soil surface (e.g., facing of the reinforced embankment in Figure 37.16) by stapling or pinning them to the soil surface. Intimate contact between the blanket or mat and the soil is very important as water flow beneath the material has usually been the cause of poor functioning.

37.3.9 HDPE Vertical Barrier Systems

The use of geomembranes (Section 37.3.2) as horizontal barrier layers has been extended for the case of seepage control in remediation projects, in which vertically deployed geomembranes are used in vertical cutoff trenches. The construction process involves excavation of a trench and placement of a seamed geomembrane in the open trench. This procedure is usually not possible for deep trenches due to the potential collapse of the sidewalls, so the use of slurry to stabilize the trench becomes necessary. The mixture of water and bentonite clay balances the pressures exerted by the *in situ* soils. The geomembrane



FIGURE 37.16 Erosion control mat placed to help establish the vegetation on the face of a 1H:1V reinforced soil slope.

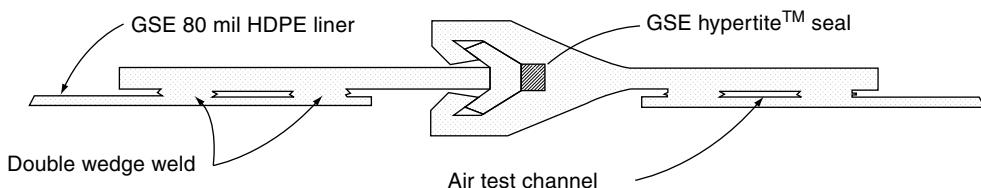


FIGURE 37.17 Interlocking system for HDPE vertical barriers (figure courtesy of GSE Lining Technology, Inc.).

is placed in the slurry after trench excavation to the intended depth. Once the geomembrane is in place the backfill can be introduced, displacing the slurry and forcing the geomembrane to the side of the trench.

As installation of vertically deployed conventional geomembranes is difficult, other systems have become available. These systems involve the use of thick HDPE or nonplasticized PVC geomembranes in the form of tongue-and-groove sheeting. Sealing of the interlocks is often achieved by using chloroprene-based, hydrophilic seals. Figure 37.17 shows one type of interlocking system with a hydrophilic seal. The seal is an extruded profile, typically 8 mm in diameter, which can expand up to eight times its original volume when exposed to water. These interlocking HDPE vertical barrier systems are being increasingly used as an alternative to soil-bentonite slurry walls, especially in projects involving areas of limited access, high disposal costs, depths where performance of a slurry wall is questionable and high concentrations of saline and chemicals.

An additional advantage of the HDPE vertical barrier system is that it can work both as containment and a collection system (e.g., a composite sheet drain) and can be constructed in one trench. A recently developed method utilizes a biopolymer, or biodegradable, slurry. These slurries allow the HDPE panels and collection system (e.g., geocomposite sheet drains) to be installed in the same trench. Unlike bentonite, these slurries will either biodegrade or can be reversed to allow the collection system to drain clear and free of fines.

Another method of achieving construction of a containment and collection system in the same trench has been developed which utilizes a trenchless, vibratory method for installation of the HDPE panels and geocomposite sheet drains. First, a collection trench is constructed to the required depth. This is followed with the installation of the geomembrane panel using modified pile-driving techniques. Panel widths ranging typically between 0.91 to 1.83 m are driven to depths up to 12 m. This construction method is most often reserved for sites on which excavation and disposal costs are high, access is limited, or the barrier is too close to a body of water. A case history in which this installation method was used for placement of an HDPE vertical barrier system is described in Section 37.5.1.

A recent development in the installation of these systems is the use of a “one-pass” deep trencher. Installation of HDPE vertical barrier systems using this technology has proven to be fast and safe. A special trencher equipment can install a vertical geomembrane wall with a collection system consisting of a HDPE pipe and gravel fill in one trench, in one pass. Figure 37.18 shows the placement of an HDPE panel using this one-pass deep trencher.



FIGURE 37.18 One pass trencher for installation of an HDPE vertical barrier system. (Photo courtesy of Groundwater Control, Inc.)

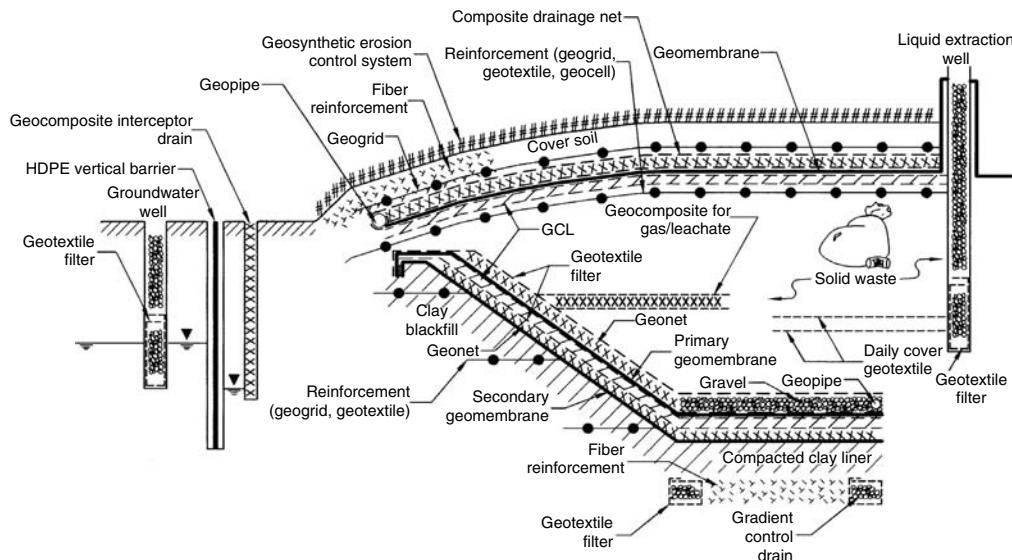


FIGURE 37.19 Multiple use of geosynthetics in landfill design.

37.4 Geosynthetic Applications in Landfill Design

The multiple uses of geosynthetics in the design of modern municipal solid waste landfills is a good illustration of an application in which different geosynthetics are used to perform all the functions discussed in Section 37.2. Virtually all the different types of geosynthetics discussed in Section 37.3 have been used in the design of both base and cover liner systems of landfill facilities. The extensive use of geosynthetics in modern landfills has been triggered by the economical and technical advantages that geosynthetics offer in relation to traditional liner systems. A geomembrane infiltration barrier and geocomposite sheet drain collection layers of a few millimeters in thickness can provide equivalent performance as a soil infiltration barrier with a gravel collection layer and graded granular filter layer of up to several meters in thickness.

Landfill base liners are placed below the waste to minimize the release of liquids from the waste (i.e., leachate). Leachate is the main source of contamination of the soil underlying the landfill and, most importantly, the groundwater. Landfill cover liners are placed above the final waste configuration to prevent water, usually from rain or snow, from percolating into the waste and producing leachate. Waste containment systems employ geosynthetics to varying degrees. Figure 37.19 illustrates the extensive multiple uses of geosynthetics in both the cover and the base liner systems of a modern landfill facility.

The base liner system illustrated in Figure 37.19 is a double composite liner system. Double composite liner systems are used in some instances for containment of municipal solid waste and are frequently used for landfills designed to contain hazardous waste. The base liner system shown in this figure includes a geomembrane/GCL composite as the primary liner system and a geomembrane/compacted clay liner composite as the secondary system. The leak detection system, located between the primary and secondary liners, is a geotextile/geonet composite. The leachate collection system overlying the primary liner on the bottom of the liner system consists of gravel with a network of perforated pipes. A geotextile protection layer beneath the gravel provides a cushion to protect the primary geomembrane from puncture by stones in the overlying gravel. The leachate collection system overlying the primary liner on the side slopes of the liner system is a geocomposite sheet drain (geotextile/geonet composite) merging into the gravel on the base. A geotextile filter covers the entire footprint of the landfill and prevents clogging of the leachate collection and removal system. The groundwater level may be controlled at the bottom of the landfill

by gradient control drains built using geotextile filters. Different types of geosynthetics (e.g., geogrids, geotextiles, fibers) can be selected for stabilization of the foundation soils.

The cover system of the landfill illustrated in Figure 37.19 contains a composite geomembrane/GCL barrier layer. The drainage layer overlying the geomembrane is a geocomposite sheet drain (composite geotextile/geonet). In addition, the soil cover system may include geogrid, geotextile, or geocell reinforcements below the infiltration barrier system. This layer of reinforcements may be used to minimize the strains that could be induced in the barrier layers by differential settlements of the refuse or by a future vertical expansion of the landfill. In addition, the cover system could include a geogrid or geotextile reinforcement above the infiltration barrier to provide stability to the vegetative cover soil. Fiber reinforcement may also be used for stabilization of the steep portion of the vegetative cover soil. A geocomposite erosion control system above the vegetative cover soil is indicated in the figure and provides protection against sheet and gully erosion.

Figure 37.19 also illustrates the use of geosynthetics within the waste mass, which are used to facilitate waste placement during landfilling. Specifically, the figure illustrates the use of geotextiles as daily cover layers and of geocomposites within the waste mass for collection of gas and leachate. Geotextile filters are also extensively used for leachate collection and detection blanket drains and around the gravel in leachate collection trenches at the base of the landfill. Geosynthetics can also be used as part of the groundwater and leachate collection well system. The use of geotextiles as filters in groundwater and leachate extraction wells is also illustrated in the figure. Finally, the figure shows the use of a HDPE vertical barrier system and a geocomposite interceptor drain along the perimeter of the landfill facility. Although not all of the components shown in Figure 37.19 would normally be needed at any one landfill facility, the figure illustrates the many geosynthetic applications that can be considered in landfill design.

Bouazza et al. (2002) provide an update on the use of geosynthetics in the design of waste containment facilities. A case history involving the design of a liner system with multiple uses of geosynthetics is described in Section 37.5.2.

37.5 Case Histories

37.5.1 Case History of Vertical Barrier System

Although the use of geosynthetics in many geotechnical and environmental projects is related to groundwater applications (e.g., landfill liners, which prevent groundwater contamination), a geosynthetic application directly related to groundwater remediation and control is the use of HDPE panels as vertical barrier systems. A case history is presented herein to illustrate the use of HDPE panels as part of a remediation plan for a site contaminated with coal tar (Burson et al., 1997).

The site was a defunct manufactured gas plant located in York, Pennsylvania. The site is surrounded by commercial and residential areas and a creek (Codorus Creek) borders the site for a distance of approximately 305 m. During years of operation and the subsequent closure of the manufactured gas plant, some process residuals migrated to the subsurface soils and groundwater. Over time, the presence of a coal tar-like material in the form of a dense non-aqueous phase liquid (DNAPL), was observed seeping from the bank of the Codorus Creek. DNAPL was also noted in some on-site monitoring wells.

Several remediation scenarios were evaluated with the purpose of intercepting the tar-like material migrating through the soil and into groundwater, encountered approximately 5.0 m below ground surface. A system consisting of a combination of soil improvement by jet grouting, a vertical barrier using HDPE panels, and a network of recovery wells was finally selected.

The use of vertical HDPE panels and trenchless technology allowed placement of the barrier as close as 3 m from the bank of the Codorus Creek, which was difficult to achieve with conventional slurry wall technology. The HDPE barrier system selected for this project was a 2 mm thick geomembrane, which allowed for the vibratory, trenchless installation. Sealing of the interlocks was achieved with a chloroprene-based, hydrophilic seal (see Figure 37.17). HDPE panels were keyed into the soil improved by jet grouting,



FIGURE 37.20 Installation of HDPE barrier wall utilizing conventional pile driving equipment. (Photo courtesy of Groundwater Control Inc.)

as discussed below. The panels were installed using conventional vibratory pile driving equipment, without a trench, thus reducing the amount of contaminated spoils to be disposed (Figure 37.20).

To complete closure of the contaminated material, jet grouting was used to provide a seal to control DNAPL migration between the bottom of the HDPE panels and the irregular bedrock contact. Jet grouting consists of high pressure injection of cement and bentonite slurry, horizontally into the soil strata to improve its mechanical and hydraulic properties. The containment wall was approximately 290 m in length. The soil along the alignment of the barrier system consisted of granular fills, with large amounts of cinder material. Also mixed into the fill were varying amounts of rubble and debris. These highly permeable soils were underlain by the competent bedrock. Holes were pre-drilled down to bedrock, and the jet grouting improvement was done by injecting the grout horizontally from the competent rock up to an elevation approximately 6 m below the ground surface.

A groundwater recovery system was implemented once the barrier was completed. Since its installation in the fall of 1995, the HDPE panel jet grout barrier system has performed as intended.

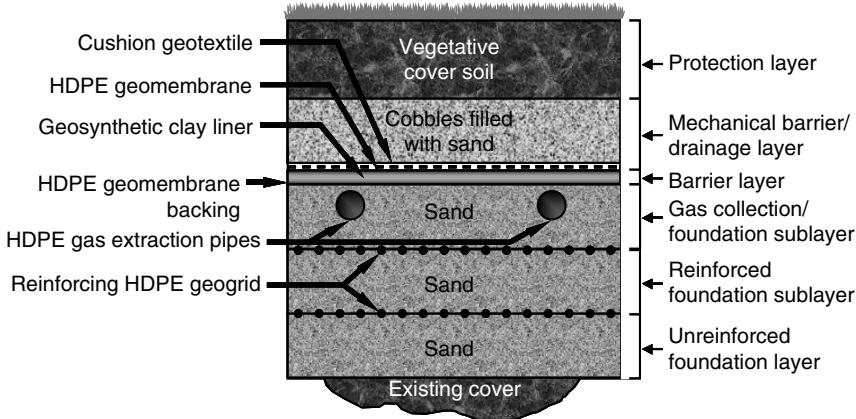


FIGURE 37.21 Cover system reinforced using geogrids.

37.5.2 Case History of Multiple Use of Geosynthetics in Landfill Cover Design

The cover system at the McColl Superfund Site, Fullerton, California is a good example of a site where multiple systems of soil reinforcement were used for stabilization of the final cover system. One of these uses involves placement of geogrids along the cover system. The project also included the construction of conventional reinforced structures (Collins et al., 1998; Hendrickson et al., 1998).

The site has 12 pits containing petroleum sludge and oil-based drilling muds. The sludge was generated by the production of high-octane aviation fuel and was placed into the pits between 1942 and 1946. Between 1952 and 1964, the site was used for disposal of oil-based drilling muds. These wastes and their reaction products and byproducts were found as liquid, gas, and solid phases within the pits. At the time of deposition, essentially all of the waste materials were mobile. Over time, much of the waste had hardened. The drilling mud is a thixotropic semi-solid sludge, which can behave as a very viscous fluid.

Key considerations for the selection of the final remedy were to: (1) provide a cover system that includes a barrier layer and a gas collection and treatment system over the pits to minimize infiltration of water and release of hazardous or malodorous gas emissions; (2) provide a subsurface vertical barrier around the pits to minimize outward lateral migration of mobile waste or waste byproducts and inward lateral migration of subsurface liquid; and (3) provide slope stability improvements for unstable slopes at the site.

The geogrid reinforcement for the cover system over the more stable pits was constructed with two layers of uniaxial reinforcement placed orthogonally to one another. Connections at the end of each geogrid roll were provided by Bodkin joints. Adjacent geogrid panels did not have any permanent mechanical connections. This proved to be somewhat problematic, as additional care was required during placement of the overlying gas collection sand to minimize geogrid separation. Details of the cover system involving geogrid reinforcement are shown in Figure 37.21.

A geocell reinforcement layer was constructed over the pits containing high percentages of drilling mud. While the construction of this reinforcement layer proceeded at a slower pace than the geogrid reinforcement, it did provide an immediate platform to support the load. As the bearing capacity of the underlying drilling mud was quite low, the geocell provided load distribution, increasing the overall bearing capacity of the cover system. Details of the cover system involving geogrid reinforcement are shown in Figure 37.22.

In addition to reinforced covers, three conventional reinforced soil structures were constructed at the site. One of the structures was necessary to provide a working pad for construction of the subsurface vertical barrier. This reinforced earth structure had to support the excavator with a gross operating weight of 1100 kN that was used to dig the soil-bentonite cutoff wall. Another reinforced earth structure at the site

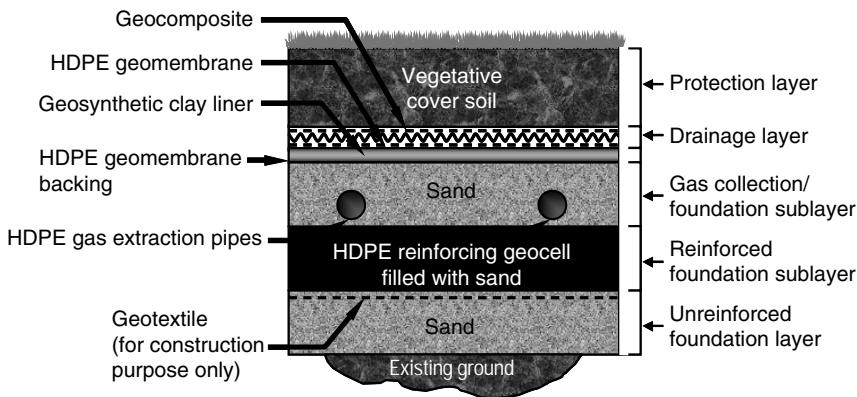


FIGURE 37.22 Cover system reinforced using geocells.

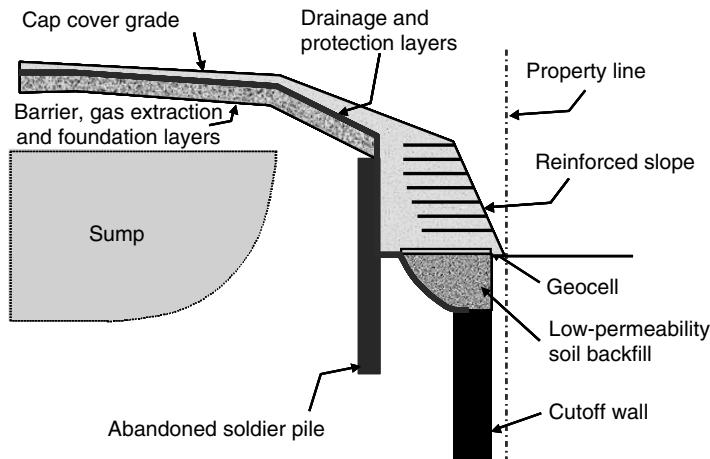


FIGURE 37.23 Buttressing reinforced slope at McColl Superfund site.

had to span a portion of the completed cutoff wall. Due to concerns that the stress of the reinforced earth structure on the underlying soil-bentonite cutoff wall would lead to excessive deformation of the wall due to consolidation of the cutoff wall backfill, a flexible wall fascia was selected. As shown in Figure 37.23, a soldier pile wall was constructed to provide stability of the system during construction. The use of geosynthetic alternatives in this project was more suitable and cost effective than their conventional counterparts.

Glossary

Alloys, polymeric a blend of two or more polymers (e.g., a rubber and plastic) to improve a given property (e.g., impact strength).

Apparent opening size (AOS), O₉₅ for geotextile, a property which indicates the diameter of the approximate largest particle that would effectively pass through the geotextile. At least 95% of the openings apparently have that diameter or are smaller as measured by the dry sieve test.

Chemical stability stability of a geosynthetic; ability to resist degradation from chemicals, such as acids, bases, solvents, oils and oxidation agents; and chemical reactions, including those catalyzed by light.

Chlorosulfonated polyethylene (CSPE) family of polymers that is produced by polyethylene reacting with chlorine and sulfur dioxide. Present CSPEs contain 25 to 43% chlorine and 1.0 to 1.4% sulfur.

Clogging movement by mechanical action or hydraulic flow of soil particles into the voids of fabric and retention therein, thereby reducing the hydraulic conductivity of the geotextile.

Cross-machine direction the axis within the plane of a fabric perpendicular to the predominant axis of the direction of production.

Cross-plane The direction of a geosynthetic that is perpendicular to the plane of its manufactured direction. Referred to in hydraulic situations.

Fiber basic element of fabrics and other textile structures, characterized by having a length at least 100 times its diameter or width that can be spun into a yarn or otherwise made into a fabric.

Filament yarn the yarn made from continuous filament fibers.

Filtration in geotextiles; the process of retaining soil in place while allowing water to pass from soil. In chemistry, removal of particles from a fluid stream.

Geocell a three-dimensional structure filled with soil, thereby forming a mattress for increased stability when used with loose or compressible subsoils.

Geocomposite a manufactured material using geotextiles, geogrids, and geomembranes in laminated or composite form. May or may not include natural materials.

Geogrid open grid structure of orthogonal filaments and strands of polymeric material used primarily for tensile reinforcement.

Geomembrane very low hydraulic conductivity synthetic membrane liners or barriers used with any geotechnical engineering related material so as to control fluid migration in a man-made project, structure, or system.

Geonet a geosynthetic consisting of integrally connected parallel sets of ribs overlying similar sets at various angles for planar drainage of liquids or gases.

Geopipe any plastic pipe used with foundation, soil, rock, earth, or any other subsurface related material as an integral part of a human-made project, structure, or system.

Geosynthetic a planar product manufactured from polymeric material used with soil, rock, earth, or other geotechnical engineering related material as an integral part of a constructed project, structure, or system.

Geosynthetic clay liner (GCL) factory-manufactured hydraulic barriers consisting of a layer of bentonite clay or other very low permeability material supported by geotextiles and geomembranes, and mechanically held together by needling, stitching, or chemical adhesives.

Geotextile any permeable textile used with foundation, soil, rock, earth, or any other geotechnical engineering-related material as an integral part of a constructed project, structure, or system.

Grab test in fabric testing, a tension test in which only a part of the width of the specimen is gripped in the clamps.

Gradient ratio the ratio of the average hydraulic gradient across the fabric and the 25 mm of soil immediately next to the fabric to the average hydraulic gradient across the 50 mm of soil between 25 and 75 mm above the fabric, as measured in a constant head permeability test.

Heat bonded thermally bonded by melting the fibers to form weld points.

Hot wedge common method of heat seaming thermoplastic geomembranes by a fusing process wherein heat is delivered by a hot wedge passing between the opposing surfaces to be bonded.

Hydraulic transmissivity for a geotextile or related product, the volumetric flow rate of water per unit width of specimen per unit gradient in a direction parallel to the plane of the specimen.

Index test a test procedure that may contain a known bias but which may be used to establish an order for a set of specimens with respect to the property of interest.

In-plane the direction of a geosynthetic that is parallel to its longitudinal, manufactured, or machine direction. Referred to in hydraulic situations.

Leachate liquid that has percolated through or drained from solid waste or other human-emplaced materials and contains soluble, partially soluble, or miscible components removed from such waste.

Liner a layer of emplaced materials beneath a surface impoundment or landfill that serves to restrict the escape of waste or its constituents from the impoundment or landfill.

Machine direction the direction in the plane of the fabric parallel to the direction of manufacture.

Mass per unit area the proper term to represent and compare the amount of material per unit area (units are or g/m²) of a geosynthetic.

Monofilament a single filament of a fiber (normally synthetic).

Mullen burst hydraulic bursting strength of textiles.

Multifilament a yarn consisting of many continuous filaments or strands.

Needlepunched in geotextiles, mechanical bonding of staple or filament fibers with barbed needles to form a compact fabric.

Nonwoven fabric a textile structure produced by bonding or interlocking of fibers, or both, accomplished by mechanical, thermal, or chemical means.

Permittivity of geotextiles and related products, the volumetric flow rate of water per unit cross sectional area per unit head under laminar flow conditions, in the normal direction through a geotextile.

Plasticizer a plasticizer is a material, frequently solvent-like, incorporated in a plastic or a rubber to increase its ease of workability, its flexibility, or distensibility.

Polyester fiber generic name for a manufactured fiber in which the fiber-forming substance is any long-chain synthetic polymer composed of an ester of a dihydric alcohol and terephthalic acid.

Polyethylene a polyolefin formed by bulk polymerization (for low density) or solution polymerization (for high density) where the ethylene monomer is placed in a reactor under high pressure and temperature.

Polymer a macromolecular material formed by the chemical combination of monomers having either the same or different chemical composition. Plastics, rubbers, and textile fibers are all high-molecular-weight polymers.

Polyolefin a family of polymeric materials that includes polypropylene and polyethylene, the former being very common in geotextiles and the latter in geomembranes.

Polyvinyl chloride (PVC) a synthetic thermoplastic polymer prepared from vinyl chloride.

Quality assurance (QA) a planned system of activities whose purpose is to provide a continuing evaluation of the quality control program, initiating corrective action where necessary. It is applicable to both the manufactured product and its field installation.

Quality control (QC) actions that provide a means of controlling and measuring the characteristics of (both) the manufactured and the field installed product.

Separation the function of geosynthetics as a partition between two adjacent materials (usually dissimilar) to prevent mixing of the two materials.

Specification a precise statement of a set of requirements to be satisfied by a material, product, system, or service that indicates the procedures for determining whether each of the requirements is satisfied.

Spun-bonded fabrics fabric formed by continuous filaments that have been spun (extruded), drawn, laid into a web and bonded (chemical, mechanical, or thermal bonding) together in one continuous process.

Staple fibers fibers of short lengths; frequently used to make needlepunched nonwoven fabrics.

Subgrade intrusion localized aggregate penetration of a soft cohesive subgrade and resulting displacement of the subgrade into the cohesionless material.

Subgrade pumping the displacement of cohesive or low-cohesion fines from a saturated subgrade into overlying aggregate as the result of hydraulic forces created by transmittal of wheel-load stresses to the subgrade.

Survivability the ability of a geosynthetic to be placed and to perform its intended function without undergoing degradation.

Tensile strength the maximum resistance to deformation developed for a specific material when subjected to tension by an external force.

Trapezoid Tear Test test method used to measure the tearing strength of geotextiles.

Transmissivity for a geosynthetic, the volumetric flow rate per unit thickness under laminar flow conditions, in the in-plane direction of the fabric or geocomposite.

Ultraviolet degradation the breakdown of polymeric structure when exposed to natural light.

Woven geotextilea planar geotextile structure produced by interlacing two or more sets of elements such as yarns, fibers, rovings of filaments where the elements pass each other usually at right angles and one set of elements are parallel to the fabric axis.

Wide-width strip tensile test a uniaxial tensile test in which the entire width of a 200 mm wide specimen is gripped in the clamps and the gauge length is 100 mm.

Yarn a generic term for continuous strand strands (1 or more) of textile filaments, monofilaments, or slit form suitable for knitting, weaving, or otherwise intertwining or bonding to form a textile fabric.

Sources for these and other definitions of terms can be found in ASTM (1997) and Koerner (2005).

References

- ASTM. 1995. *ASTM Standards on Geosynthetics*. Sponsored by ASTM Committee D-35 on Geosynthetics, Fourth Edition, 178p.
- Bhatia, S.K. and Smith, J.L. 1996a. "Geotextile Characterization and Pore-Size Distribution: Part I. A Review of Manufacturing Processes," *Geosynthetics International*, 3, 85–105.
- Bhatia, S.K. and Smith, J.L. 1996b. "Geotextile Characterization and Pore-Size Distribution: Part II. A Review of Test Methods and Results," *Geosynthetics International*, 3, 155–180.
- Bhatia, S.K., Smith, J.L., and Christopher, B.R. 1996. "Geotextile Characterization and Pore-Size Distribution: Part III. Comparison of methods and Application to Design," *Geosynthetics International*, 3, 301–328.
- Bouazza, A. 2002. Geosynthetic clay liners. *Geotextiles and Geomembranes*, 20, 1–17.
- Bouazza, A., Zornberg, J.G., and Adam, D. 2002. "Geosynthetics in Waste Containment Facilities: Recent Advances." State-of-the-Art keynote paper, *Proceedings of the Seventh International Conference on Geosynthetics*, Nice, France, September 22–27, A.A. Balkema, Vol. 2, pp. 445–510.
- Burson, B., Baker, A.C., Jones, B., and Shailer, J. 1997. "Development and Installation of an Innovative Vertical Containment System," *Proceedings of the Geosynthetics '97 Conference*, Long Beach, California, March, Vol. 1, pp. 467–480.
- Christopher, B.R. and Dahlstrand, T.K. 1989. *Geosynthetics in Dams — Design and Use*, Association of State Dam Safety Officials, Lexington, KY, 311 p.
- Christopher, B.R. and Fisher, G.R. 1992. "Geotextile Filtration Principles, Practices and Problems," *Journal of Geotextile and Geomembranes*, Elsevier, 11, 4–6, 337–354.
- Collins, P., Ng, A.S., and Ramanujam, R. 1998. Superfund Success, Superfast. *Civil Engineering December*, 42–45.
- Elias, V., Christopher, B.R., and Berg, R.R. 2001. *Mechanically Stabilized Earth Walls and Reinforced Soil Slopes*. Publication Number FHWA-NH-00-043, NHI-FHWA.
- Giroud, J.P. 1996. "Granular Filters and Geotextile Filters," *Proceedings of Geofilters '96*, Montreal, pp. 565–680.
- Giroud, J.P. with cooperation of Beech, J.F. and Khatami, A. 1993. *Geosynthetics Bibliography. Volume 1*. IGS, Industrial Fabrics Association International (IFAI) Publishers, St. Paul, MN, 781 p.
- Giroud, J.P. with cooperation of Beech, J.F., Khatami, A., and Badu-Tweneboah, K. 1994. *Geosynthetics Bibliography. Volume 2*. IGS, IFAI Publishers, St. Paul, MN, 940 p.
- Giroud, J.P., Zornberg, J.G., and Zhao, A. 2000. "Hydraulic Design of Geosynthetic and Granular Liquid Collection Layers." *Geosynthetics International*, Special Issue on Liquid Collection Systems, 7, 285–380.

- Halse, Y., Koerner, R.M., and Lord, A.E., Jr. 1987. "Filtration Properties of Geotextiles Under Long Term Testing," *Proceedings of ASCE/Penn DOT Conference on Advances in Geotechnical Engineering*, Hershey, PA, pp. 1–13.
- Hansbo, S. 1979. "Consolidation of Clay by Band Shaped Perforated Drains," *Ground Engineering*, July, 16–25.
- Hendricker, A.T., Fredianelli, K.H., Kavazanjian, E., and McKelvey, J.A. 1998. Reinforcement Requirements at a Hazardous Waste Site. *Proceedings of the 6th International Conference on Geosynthetics*. IFAI, Atlanta, Vol. 1, 465–468.
- Holtz, R.D., Christopher, B.R., and Berg, R.R. 1997. *Geosynthetic Engineering*. Bitech Publishers Ltd., Richmond, British Columbia, Canada, 452 p.
- Holtz, R.D., Christopher, B.R., and Berg, R.R. 1998. *Geosynthetic Design and Construction Guidelines*, U.S. Department of Transportation, Federal Highway Administration, Washington DC, Report No. HI-95-038, 396 p.
- Industrial Fabrics Association International. 1996. North American Market for Geosynthetics. 97 p.
- Koerner, R.M. 2005. *Designing with geosynthetics*, 5th ed. Prentice Hall, 783 p.
- Narejo, D., Koerner, R.M., and Wilson-Fahmy, R.F. 1996. "Puncture Protection of Geomembranes Part II: Experimental," *Geosynthetics International*, 3, 629–653.
- Palmeira, E. and Fannin, J. 2002. "Soil-Geotextile Compatibility in Filtration." State-of-the-Art keynote paper, *Proceedings of the Seventh International Conference on Geosynthetics*, Nice, France, September 22–27, A.A. Balkema, Vol. 3, pp. 853–872.
- Scuero, A. M., Vaschetti, G., Sembelli, P., Aguiar Gonzalez, E., Bartek, P., Blanco Fernandez, M., Brezina, P., Brunold, H., Cazzuffi, D., Girard, H., Lefranc, M., do Vale, M., Massaro, C., Millmore, J., and Schewe, L. (2005). Geomembrane Sealing Systems for Dams. International Commission on Large Dams.
- Vidal, H. 1969. "La Terre Armée." *Annales de l'Institut Technique du Bâtiment et des Travaux Publics*, Series: Materials (38), No. 259–260, pp.1–59.
- Wilson-Fahmy, R.F., Narejo, D., and Koerner, R.M. 1996. "Puncture Protection of Geomembranes Part I: Theory," *Geosynthetics International*, 3, 605–628.
- Zornberg, J.G. 2002. "Discrete Framework for Limit Equilibrium Analysis of Fibre-Reinforced Soil." *Géotechnique*, 52, 593–604.
- Zornberg, J.G. (2005). "Advances on the Use of Geosynthetics in Hydraulic Systems." *Proceedings of the Nineteenth Geosynthetic Research Institute Conference* (GRI-19), Geosynthetics Institute, Las Vegas, NV, December 14–16, pp. 1–17 (CD-ROM).
- Zornberg, J.G. and Weber, C.T. 2003. "Geosynthetics Research Needs for Hydraulic Structures" *Proceedings of the Seventeenth Geosynthetic Research Institute Conference*, Hot Topics in Geosynthetics IV, December 15–16, Las Vegas, Nevada, pp. 183–197.
- Zornberg, J.G., Barrows, R.J., Christopher, B.R., and Wayne, M.H. 1995. "Construction and Instrumentation of a Highway Slope Reinforced with High Strength Geotextiles." *Proceedings of the Geosynthetics '95 Conference*, Nashville, TN, February, Vol. 1, pp.13–27.
- Zornberg, J.G., Sitar, N., and Mitchell, J.K. 1998. "Performance of Geosynthetic Reinforced Slopes at Failure." *Journal of Geotechnical and Geoenvironmental Engineering*, ASCE, 124, 670–683.
- Zornberg, J.G., McCartney, J.S., and Swan, R.H. 2005. "Analysis of a Large Database of GCL Internal Shear Strength Results," *Journal of Geotechnical and Geotechnical Engineering*, ASCE, 131, 367–380.

Further Information

- Koerner (2005) provides an excellent, well-illustrated overview of the different types of geosynthetics and their applications.
- Holtz, Christopher and Berg (1997, 1998) provide well-documented practical design and construction information on the different uses of geosynthetic products.

Giroud et al. (1993, 1994) provide a two-volume comprehensive database on technical literature relative to geosynthetics, including technical papers from conferences, technical papers from journals, books, theses, and research reports.

Technical advances on geosynthetics are also published in the two official journals of the IGS: Geosynthetics International, and Geotextiles and Geomembranes. Similarly, the Geotechnical Fabrics Report (GFR), published by the Industrial Fabrics Association International (IFAI) provides updated information, including the annual Specifier's Guide, which offers a summary of the properties of products available in the geosynthetics market.

The ASTM Standards on Geosynthetics, sponsored by ASTM Committee D-35 on Geosynthetics (ASTM, 1995), provides information on the standard test procedures for the different types of geosynthetics.

The proceedings of the International Conferences on Geosynthetics, organized by the International Geosynthetic Society (IGS), offer a relevant source of information on the different topics related to geosynthetics. These international conferences are organized every four years. Equally relevant are the proceedings of conferences organized by the regional chapters of IGS. Finally, the proceedings of the series of conferences organized by the Geosynthetics Research Institute (GRI) provide information on specific topics relevant to geosynthetic design.

Geosynthetic manufacturers' literature is also a valuable source of information, which provides product-specific properties, suggested design methods, and recommended factors of safety.

Index

A

- Above-ground fill, landfill, 33-6
Absolute dominion doctrine, 32-4–32-5
Absolute pressure, 3-2–3-3
Absolute Thermodynamic Temperature Scale,
 see Kelvin scale
Absorption, 17-10
Acetone, 31-8
Acid and chemicals
 in well development, 11-27–11-28
Acid–base
 behavior, 17-5–17-6
 diagram, 17-12
 reaction, 17-4, 17-8
Acid dissociation constants, 17-6
Acid mine drainage, 17-18–17-19
Acidophilic bacteria, 31-6
Acoustic logs, 14-19–14-20
Adenosine triphosphate (ATP), 31-3
Adhesion, 2-12, 2-14
Adsorbed contaminant, 3-29
Adsorption, 3-37, 17-10
Advection, 22-10
 versus convection, 13-6
Advection–dispersion equation (ADE), 18-4
 linear and nonlinear sorption, 22-11–22-12
 transport equations, 22-11
 volatilization, 22-13
Advection–dispersion model (ADM), 26-6, 26-8–26-9
Aeolian deposits, 2-9
Aerobic chlorinated aliphatic hydrocarbon
 bioremediation, 31-21
Aerobic petroleum hydrocarbon bioremediation, 31-21
Aerobic respiring organisms, 31-3
Aerosols, 17-16
Air entry pressure, *see* Bubbling pressure
Air potential, 6-6
Air rotary, 11-25, 35-19
Air sparging (AS), 36-36–36-37
Air surging, 11-27
Airport safety, 33-4
Air-temperature changes, annual, 13-9
Akaike's Information Criterion (AIC), 14-44
Alaska, 2-49
Aliphatic compounds, halogenated, 31-8–31-10
Alkanes, 31-7
Allogenic recharge, 21-2
Alluvium, 2-4, 2-9
Alpha particles, 17-25
Alternate concentration limit (ACL), 35-3
Amargosa Desert, 2-38
Amchitka model, uncertainty reduction,
 24-18–24-22
 background, 24-19–24-20
 Bayesian framework, 24-20–24-21
 Milrow site, 24-21–24-22
American Society for Testing and Materials
 (ASTM), 2-10
Amphoterism, 17-4
Amsterdam, 1-9
Anaerobic bacteria, 31-7
Anaerobic respiration, 31-3
Analysis of Contaminant Transport (ACT), 22-35
Analytical model, 23-2
Anchorage, 2-49
Anilines, 31-10
Anisotropic barriers, 34-12–34-13
Anisotropic media, 3-10
Anisotropy, 4-9–4-10, 15-7
 and heterogeneity, 2-22–2-25
ANOVA, 35-42
Anthropogenic changes, in recharge areas,
 2-51–2-52
Aperture, 20-31
Appalachians, 2-9, 2-50–2-51
Approximate mobility index (AMI), 2-49
Aqueous solubility, affecting factors
 $p\epsilon$, 17-11–17-13
 pH and chemical composition, 17-3–17-10
 sorption, to solids, 17-10–17-11
 temperature, 17-13
Aquiclude, 2-27

- Aquifer, 1-15, 1-27, 2-25–2-28
 artesian, 2-25, 2-26
 basin and range, dynamics of, 27-6–27-8
 coastal, 3-28
 management, 12-5, 12-23–12-26
 confined, 1-13, 2-25–2-26, 3-23, 3-25, 10-2–10-3
 distance–drawdown analyses, 10-27–10-28
 flow equation, 4-19–4-20
 time–drawdown analyses, 10-17–10-19
 well-flow equations, 10-5–10-6
 dynamics and models, 27-5–27-6
 exploration, 1-14–1-16
 heterogeneous interpretation, 10-30–10-31
 heterogeneity accounting
 in solute transport modeling, MADE site case study, 26-1–26-14
 hydrogeologic parameters, 35-8
 karstic, 20-3, 21-2
 leaky, 3-14–3-15, 10-2, 10-3, 10-28–10-30
 losses, 11-5–11-6, 11-16
 well efficiency, 11-7
 multilayered, 10-2
 perched, 2-26–2-27
 phreatic, 4-20–4-24
 physical properties, 10-2–10-4
 pumped, 10-6–10-9
 reactivity, 19-5–19-6
 restoration time, 36-26–36-28
 semi-confined, 2-27, 2-28
 shallow unconfined, flow and transport, 23-31–23-35
 surficial, 23-32
 tests, performance, 10-12–10-17
 data processing, 10-15–10-17
 date interpretation, 10-17
 measurements, 10-13–10-15
 pumping test duration, 10-15
 3D modeling, 23-42
 types, 10-2
 unconfined, 2-26, 3-23, 3-25, 10-2–10-3
 Aquitard, 2-27, 4-31
 flow equation, 4-24
 ArcGIS, 30-2
 Archie's relationship, 14-6
 Archimedian screw, 1-3
 ArcIMS, 30-2, 30-5
 ArcView, 30-2
 Area fill, landfill, 33-6
 Arizona, 2-38
 Arkansas, 2-38
 Aromatic compounds, halogenated, 31-11
 Aromatic rings, 31-5
 Artesian aquifer, 2-25–2-26
 Artesian pressure and surface effects, 2-50
 Artesian wells, 1-3, 1-5, 1-8
 Artificial drainage effects, 2-51
 Artificial protective barriers, 2-45
 Artois, 1-3, 2-7, 2-25
 Arya–Paris model, 6-25–6-28
 drawbacks, 6-27–6-28
 ASCE Hydraulics Division Committee on Hydrologic Transport Dispersion, 31-14
 Ash Meadows Springs, 2-38
 AspMap, 30-5
 Assessment monitoring programs, groundwater monitoring, 35-3–35-4
 AT123D, 22-25
 Atlantic Coastal Plain, 2-48
 Atlas, 30-2
 Atmospheric discharge and seepage face, 2-32–2-33
 Attachment/detachment models, 22-15–22-16
 Attenuation, 15-10, 15-14
 difference tomography, 15-24–15-25
 tomography, 15-18
 Auger-hole test, 3-12
 Aurora, 2-43
 Automated parameter estimation techniques, 23-23
 Automatic history matching, 23-23
 Autotrophs, 31-2
 Average pore diameter, 2-34
- B**
- B. subtilis*, 31-12
 BA model, 7-9
 Backfill materials, 35-21
 Bacteria, 31-2–31-4, 31-6
 Balkan Peninsula, 2-7, 2-9
 Bank storage, 2-42
 Barrier function, geosynthetic, 37-9–37-10
 Barrier layer, 33-15, 33-22–33-23
 Base flow/runoff relations, 21-28
 Baseflow recession constant, 2-36
 Basin and range aquifer, dynamics of, 27-6–27-8
 Basin and range areas, 2-38
 Basin-area method, 2-37
 Batch-flush model, 36-27–36-28
 Bayesian approach, 30-11
 geostatistics, 15-19
 inversion approach, 15-20
 for uncertainty reduction, 24-20–24-21
 Bayesian Information Criterion (BIC), 14-44
 Bedrock
 granitic, 14-5
 surface, 21-4
 Beerkan method, 6-22
 Benchmarking, 23-19
 Bentonite, 2-45, 35-14, 35-21, 37-10, 37-18
 Benzene, 31-8
 Benzoates, 31-10
 Berms, 33-15–33-16
 Bernoulli equation, 3-13, 5-3, 5-4
 for conservation of energy, 1-5
 Bessel equation, modified, 4-32
 Beta distribution, 5-20–5-21
 Beta particles, 17-25
 Big Sioux River valley, 2-9
 Big Springs, Texas, 2-38
 Bi-modal behavior, 6-12, 6-14
 Bioavailability and mass transfer limitations, 31-11
 Biodegradation
 bioremediation field applications
 case histories, 31-21
 engineered bioremediation, 31-21
 intrinsic bioremediation, 31-20
 technologies, emerging, 31-22
 environmental conditions, influencing
 biotransformations, 31-6–31-10
 limits, 31-10–31-11

- reactive transport and biodegradation, quantitative description
 microbial reaction kinetics, 31-12–31-15
 microbial reactions and sorption, 31-18–31-20
 sorption, as reactive kinetic process, 31-15–31-18
 subsurface microbiology and geochemistry
 metabolism, 31-3–31-6
 microbiology, 31-1–31-3
- Biological remediation**
 in *ex situ* treatment, 36-41
 in *in situ* treatment, 36-32–36-34
- Biomass**, 19-19
- Bioreactors**, 36-41
 landfills, 33-19–33-21
- Bioremediation field applications**, 36-32–36-34
 case histories, 31-21
 engineered bioremediation, 31-21
 intrinsic bioremediation, 31-20
- Biosparging**, 36-36
- Bioventing**, 36-36
- Bittern brines**, 2-36
- Black-centered grid**, 23-11
- Black Hills**, 2-35
- Blanks**, QC check, 35-30
- BLUE method**, 16-10, 16-13, 16-14
- Bog iron ore**, 2-51
- Boise, Idaho**, 15-24
- Book of Genesis**, 2-8
- Borden Aquifer**, 15-3, 15-24, 18-14
- Borehole**, 20-10, 35-18
 acoustic televiewer, 20-12
 flow meters, 15-3, 15-5
 geophysical methods, 14-16–14-20, 20-12
 acoustic logs, 14-19–14-20
 caliper logs, 14-17
 electric logs, 14-17–14-19
 nuclear logs, 14-19
 instrumentation, 20-23–20-25
 video camera (BVC), 20-12
- Borrow areas**, 33-10
- Boulders**, 2-8
- Boundary conditions**, 4-13–4-15
- Boussinesq equation**, 4-23
- Branching**, 31-7
- Breakthrough curve**, 3-34, 19-19, 23-24
 boundary layer effect on, 9-12–9-14
- British Columbia**, 15-11
- British Thermal Unit (BTU)**, 13-3
- Bromide**, 26-4–26-5
- Brooks–Corey formula**, 2-31, 6-13–6-15
- Brundtland Commission**, 27-1
- Brussels World Exhibition**, 1-22
- BTTEX 31-5, 31-7, 31-11, 31-14**
 bioremediation, 31-21
- Bubbling pressure**, 2-29
- Bucket auger**, 35-19
- Buoyancy induced dispersion**, 18-3
- C**
- C/N ratio**, 29-26
- C-14 dating**, 1-3
- Ca/Mg ratio**, 21-21
- Ca²⁺**, 17-7
- Cable tool**, 11-25, 35-19
- Calcite**, 2-9
 dissolution, 21-18
- Calibration**
 checks, QC check, 35-31
 criteria, 23-26–23-27
 of deterministic groundwater model, 23-22
 of model, 23-22–23-23
- California**, 1-16, 1-21, 2-45, 2-50, 3-28, 15-7, 15-17, 15-19
- Caliper logs**, 14-17
- Calories**, 13-3
- Canada**, 2-3, 2-38–2-39, 2-52, 18-14
- Cape Cod Aquifer**, 2-9, 15-3–15-4, 15-11, 18-14
- Capillarity**, 2-12–2-13, 2-50
 barrier, 33-22, 34-11–34-12
 fringe, 2-28, 2-32, 2-50, 15-12
 potential, 1-13–1-14
 water, 2-14
- Capture**, 27-2
- Capture zone**, 2-45
 around pumping well, 2-47
- Carbon**, 17-27
 adsorption, 36-43
 cycle, in soil-plant system, 29-25
- Carbon–nutrient–water cycle**, 29-1
- Carbonate**, 2-9
 dissolution chemistry, 21-17–21-24
 system, 17-15
- Carbonic acid**, 17-15, 21-17–21-18
- Cartesian coordinates**, 4-10
- Cartesian tensor**, 23-4
- Cation exchange**, 22-20
- Cauchy boundary condition**, 22-22
- Caves**, 21-14
- Cellular confinement system**, 37-23–37-24
- Celsius scale**, 13-2–13-3
- Centrifugation**, 34-7
- Centrifuge permeameter for unsaturated soils (CPUS)**, 34-9
- CFITIM**, 22-24–22-26
- CFITM**, 22-24–22-25
- CHAIN code**, 22-26
- Channel model**, 20-27
- Charcoal**, as dye receptor, 21-29
- Chebyshev's inequality**, 5-14
- Chemical and isotope method**, 2-37
- Chemical changes and surface-water interactions**, 2-51
- Chemical encrustation**, 11-29
- Chemical fences**, 1-27
- Chemical hardness**, 17-17
- Chemical nonequilibrium**
 attachment/detachment models, 22-15–22-16
 kinetic sorption models, 22-15
- Chemical precipitates**, 2-9
- Chemical transformations**, microbiologically mediated, 17-15
- Chemisorption**, 17-10
- Chemographs**, of karst springs, 21-34–21-37
- China**, 1-3, 2-7, 2-10
- Chlorinated congeners**, 31-10
- Chlorine**, 11-29
- Chlorobenzene**, 31-10
- Chlorofluorocarbons (CFC)**, 23-32
- Chlorosulfonated polyethylene (CSPE)**, 37-12

- Chromatography, 17-11, 35-36
- Churn drill, 2-7
- Classical deterministic models, 18-4
- Clay, 2-8-2-9
- Clean Air Act (CAA), 17-24, 33-2-33-3
- Clean Water Act (CWA), 32-7-32-10, 33-2
- legislative background
- general provisions, 32-8-32-9
 - goals and objectives, 32-8
 - historical perspective, 32-7
 - issues and outlook, 32-9
- regulations, 32-9
- Cleanup Standards, 32-24-32-25
- Climate change, impact of, 2-46-2-48, 28-1, 28-3-28-5
- climate change, 28-2-28-3
- Grand River watershed, case study, 28-20
- conclusions, 28-33-28-34
 - input data, 28-21-28-25
 - methodology, 28-25-28-26
 - results, 28-26-28-33
- groundwater recharge, 28-5-28-7
- recharge, affecting factors
- infiltration and flow, in unsaturated zone, 28-10
 - land use and cover, 28-8
 - overland flow, 28-10
 - precipitation, 28-7-28-8
 - soil properties, 28-10
 - temperature, 28-11
 - urbanization, 28-9
 - vegetation, 28-8-28-9
- recharge, in groundwater modeling, 28-15-28-20
- coupled models, 28-16-28-18
 - integrated models, 28-18-28-20
- recharge estimation methods, 28-11-28-14
- combined modeling, 28-14
 - Darcian approaches, 28-13
 - empirical methods, 28-12-28-13
 - field techniques, 28-12
 - inverse groundwater modeling, 28-14
 - tracers, 28-13
 - water budget methods, 28-13
- Climatic zonation, 1-26
- Clogging, 37-7
- Closed depressions, 21-3
- CO₂, 1-12, 2-47, 2-51, 17-5, 17-15
- Coastal aquifer, 3-28
- management, 12-5, 12-23-12-26
 - water withdrawal, 12-25
- Coastal collector, *see* Upconing
- Cokriging, 15-19, 16-13
- Colloidal-facilitated solute transport, 22-16-22-17
- Colophon, 1-4
- Colorado, 2-38, 2-43, 2-52, 15-21, 30-9
- Cometabolism, 31-7
- Common midpoint (CMP), 14-10, 14-12
- gathers, 15-9, 15-11
- Complexation, 22-20
- Comprehensive Environmental Response, Compensation and Liability Act (CERCLA), 3-39, 32-23-32-28, 36-2
- legislative background
- general provisions, 32-23-32-24
 - goals and objectives, 32-23
- historical perspective, 32-23
- issues and outlook, 32-24
- regulations
- guidance, 32-27
 - monitored natural attenuation, 32-27
 - National Contingency Plan, 32-25-32-26
 - Natural Resource Damage Assessments, 32-26-32-27
 - presumptive remedies, 32-27
 - technical impracticability, 32-28
 - vapor intrusion, 32-28
- Computed axial tomography (CAT) 15-14
- Computer
- code, 12-15
 - modeling, 1-29
- Concentration moment analysis, 15-4
- Conceptual errors, 23-23
- Conceptual model, 23-2
- Condensation, 1-4, 1-13
- Conditional indicator simulation, 15-21
- Conditional realization
- function simulation, 16-6
 - point simulation, 16-5-16-6
- Conduction, of heat, 13-5
- Conduits, 21-14
- Cone of depression, 3-18
- Cone penetrometer testing (CPT), 35-23, 35-30
- Confined aquifer, 1-13, 2-25, 2-26, 3-23, 3-25, 10-2, 10-3
- distance-drawdown analyses, 10-27-10-28
 - flow equation, 4-19-4-20
 - time-drawdown analyses, 10-17-10-19
 - well-flow equations, 10-5-10-6
- Conformal mapping technique, 1-10
- Congaree Bottomland Hardwood Swamp, *see* Congaree National Park
- Congaree National Park, 2-48
- Congaree River, 2-48
- Conjunctive management, 32-6
- Connate water, 2-36
- Connectivity function, in GIS spatial analysis, 30-5
- Conservation of energy, for steady flow, 3-12
- Conservation of mass, 23-3-23-4
- Conservation tracer, 18-14
- Conservative solutes, 3-29
- Consortium for Risk Evaluation with Stakeholder Participation II (CRESP II), 24-19
- Constant-head injection tests (CHIT), 20-13-20-16
- Constant head permeameter, 3-11
- Constant pressure tests, *see* Constant-head injection tests (CHIT)
- Construction, groundwater influence, 2-52
- Contaminant
- adsorbed, 3-29
 - advection, 3-32
 - aqueous solubility, 17-3-17-13
 - diffusion, 3-32-3-33
 - dispersion, 3-34-3-37
 - dissolved, transport mechanisms of, 3-32-3-39
 - immiscible, 3-29
 - inorganic constituents, 17-14-17-19
 - maximum contaminant level (MCL), 35-3
 - miscible, 3-29
 - organic constituents, 17-19-17-24

- particulates, 17-29–17-30
plumes, dynamic imaging, 15-24–15-25
radioactive decay, 17-24–17-25
radioactive waste disposal, 17-28–17-29
radioactivity decay and degradation, 3-38–3-39
radionuclides, 17-24–17-29
remediation, 17-29
sorption, 3-37–3-38
sources, 17-26–17-28
units, 17-26
water physicochemical characteristics, 17-2–17-3
- Contaminated groundwater, remediation of
extraction systems design, 36-21–36-30
 aquifer restoration time, 36-26–36-28
- Contaminated groundwater (continued)
 extraction trench systems, 36-25–36-26
 extraction well systems, 36-23–36-25
 other approaches, 36-28–36-30
- fundamentals, 36-1–36-6
 associated risks, 36-4–36-6
 remediation goals, 36-2–36-4
- hydraulic containment, 36-13–36-21
 alternative approaches, 36-20–36-21
 hydraulic barriers, 36-17–36-20
 physical barriers, 36-14–36-17
- performance monitoring, of remediation system, 36-43–36-44
remediation system design
 candidate remediation, screening, 36-8–36-12
 design effectiveness, confirmation, 36-13
 design implementation, 36-13
 detailed design, preparation, 36-12–36-13
 problem defining, 36-6–36-7
 remediation goals, defining, 36-7–36-8
- treatment techniques, 36-30
 ex situ treatment, 36-41–36-43
 in situ treatment, 36-31–36-41
- Contaminated transport, in unsaturated zone, 22-1
waterflow, variably saturated, 22-3–22-7
 mass balance equation, 22-4
 preferential flow, 22-5–22-7
 uniform flow, 22-4–22-5
- analytical models
 Fourier transformation, 22-23–22-24
 Laplace transformation, 22-23
 method of moments, 22-24
- multi-dimensional models, 22-25–22-27
- numerical models, 22-27–22-36
 finite differences, 22-28–22-29
 finite elements, 22-29–22-31
 geochemical transport models, 22-35–22-36
 single-species solute transport models, 22-32–22-35
 stability and oscillations, 22-31–22-32
- one-dimensional models, 22-24–22-25
- solute transport
 advection-dispersion equation, 22-11–22-13
 initial and boundary conditions, 22-21–22-22
 multicomponent reactive solute transport, 22-19–22-21
 multiphase flow and transport, 22-21
 nonequilibrium transport, 22-13–22-17
 stochastic models, 22-17–22-18
- transport processes, 22-8–22-10
Continuity equation, 3-12, 4-2–4-5
Continuity principle, 1-10
Continuous hierarchy, 18-2
Continuous Time Random Walk (CTRW), 18-5–18-7
Contributing area, of well, 23-35
Convection, of heat, 13-5–13-6
Convective-dispersive equation (CDE), 9-2, 9-4
Convergent flow tracer tests, 15-7
Convolution-Fickian equation, 18-5
Core drilling, 1-3, 2-7
Corrective action monitoring programs, groundwater monitoring, 35-4
Correlation, 5-27–5-30
Correlative rights doctrine, 32-5–32-6
Coteau Des Prairies, 2-38
COUP, 22-32
Coupled continuum pipe flow models, for karst aquifer, 21-38
Coupled system of equations, 22-20–22-21
Courant number (C_r^e), 22-32
Covariance, 5-28
Covariance matrix, 15-20
Crater Lake, 2-38
Creeping motion, 4-8
Cretaceous Period, 2-9, 2-26
Crosshole tomographic GPR techniques, 14-10
Cross-validation, 14-42–14-43
Cross-well radar tomography, 15-24–15-25
Cross-well tomography, 15-14–15-18
 attenuation tomography, 15-18
 travel time tomography, 15-16–15-18
- Crusted soil
 infiltration rate, 7-13
- Cryptosporidium*, 31-2
- Cryptosporidium parvum*, 17-30
- Crystalline rocks, 20-3, 20-29
- Cubic law, 2-47–2-48, 20-3
- Cubic packing, 2-11
- Cumulative infiltration, 6-33, 7-4
- Curve number method, 28-10
- Curve tangent method, 2-37
- CXTFIT, 22-18, 22-24–22-26
- D**
- DAISY, 22-32
- Dakota aquifer, 2-26, 2-35
- Dakota sandstone, 1-15, 1-16
- Damming, 2-42
- Dams, 37-9
- Darcian velocity, 2-21
 see also Specific discharge
- Darcy-Buckingham equation, 22-4
- Darcy water flux density, 7-3
- Darcy-Weisbach equation, 3-13, 21-10
- Darcy's law, 1-20, 2-20, 4-8–4-9, 5-3–5-5, 18-4, 20-34, 21-8, 23-4, 30-13, 3-7–3-12
 extensions, 4-9–4-10
 hydraulic conductivity
 field measurement, 3-11–3-12
 laboratory measurement, 3-10–3-11
 limitations, 3-9–3-10

- Data quality objectives (DQOs), groundwater monitoring, 35-9–35-10
- Data transformation function in GIS spatial analysis, 30-4
- DBCP (1,2-dibromo-3-chloropropane), 25-12–25-14
- DC resistivity methods, 14-3–14-7
dipole–dipole array, 14-4
Schlumberger array, 14-4
Wenner array, 14-3
- De Glee formula, 1-15
- Death Valley, 2-38
- Deep percolation model (DPM), 28-3–28-4
- Deep surface pit mining, 2-50
- Deep tunnel mining, 2-50
- Degradation, mass transformation, 19-19
- Degree of saturation, 2-12
- Delayed yield principle, 1-17
- Delmarva Peninsula, 23-31
- DEM (Digital Elevation Model), 30-2
- Dendro-chronologica dating, 1-3
- Dense NAPL, 17-21
- Depression springs, 2-38
- Desalination, 1-3
- Design optimization, of well fields, 11-22
- Desorption, 3-37
- Detection monitoring programs, groundwater monitoring, 35-2–35-3
- Detergents, 17-3
- Deterministic groundwater-simulation models, 23-2, 23-22
- Deterministic nonequilibrium sorption, 19-7–19-8
- Deterministic part, 16-2–16-3
- Deuterium, 1-27–1-28
- Diagenesis, 2-9
- 1,2-dibromo-3-chloropropane, *see* DBCP
- Dichloroethane (DCE), 31-8, 31-10
- Dielectric permeativity, 15-12
- Diffusion, 22-9
of heat, 13-6
- Diffusion and back-diffusion modeling, 9-18–9-22
- Diffusion driven dispersion, 18-3
- Diffusive flux, 23-6
- Digital computer, 2-2
simulation models, 1-29
- Dijon, 2-20, 3-7
- Dilution index, 18-12–18-13
- Dimensional and inspectional analyses, 6-41, 6-42
- 1,4-dioxane, 31-8
- Dipmeter, 14-17
- Diprotic acid, 17-5
- Direct-measurement method, 2-37
- Direct methods, for algebraic equations, 23-17
- Direct push technology (DPT) sampler, 35-23, 35-30
- Direct recharge, 28-6
- Direct rotary drilling method, 11-25
- Dirichlet boundary condition, 22-22
- Discharge phenomena, 1-22
- Discharge point, 36-25
- Discrete fracture networks, flow in, 20-33–20-35
- Discrete hierarchy, 18-2
- Discrete interval sampling, groundwater sampling, 35-29
- Discrete matrix potential regression methods, 6-23
- Discretization, 23-2
- Dispersion, 18-3
- Dispersion tensor, 18-6
- Dispersivity, 22-10, 23-6–23-7
flow-direction-dependent, 23-6
- Dissolved chemical transport, 9-4
- Dissolved chloride, 19-2
- Distance–drawdown analyses, 10-26–10-30
confined and unconfined aquifer, 10-27–10-28
leaky aquifers, 10-28–10-30
- DNAPLs, 21-40
- Dolomite, 2-9
dissolution, 21-18
- Dolostone, 2-9
- Dose equivalent, 17-26
- Double indexing, 23-11
- Double porosity, 1-18
- Downhole video camera, 21-31
- Downscaling, 25-7–25-9, 25-10
- Drainage function, geosynthetic, 37-8–37-9
- Drainage layer, 33-15, 33-22–33-23
- DRASTIC, 30-11
Index, 30-10
- Drawdown, 3-14, 10-1
- Drift, 2-8
- Drilling, 21-31, 35-16–35-17, 35-19
QA/QC, 35-22
- Dry bulk density, 6-5
- Dual-domain mass transfer model (DDMT), 26-8, 26-11–26-12
- Dual-permeability model, 22-5–22-7, 22-14
- Dual-porosity and mobile-immobile water models, 22-14
- Dual-porosity model, 22-5–22-7
with distribution layer at soil surface, 9-22–9-26
for gradual change at fracture–matrix interface, 9-9–9-22
boundary layer effect on BTC, at fracture outlet, 9-12–9-14
diffusion and back-diffusion modeling, 9-18–9-22
transition layer effect on concentration distribution, in matrix, 9-14–9-18
for sharp change at fracture–matrix interface, 9-3–9-9
first-order approximation, 9-8–9-9
zeroth-order approximation, 9-7–9-8
- Dug wells, 1-3
- Dune, 1-10, 1-19, 2-10
- Dupuit discharge formula, 4-24–4-27
- Dupuit–Forchheimer equation, 1-12
recharge basins, 3-17–3-18
seepage, from open channels, 3-16–3-17
steady flow
in confined and unconfined aquifers, 3-18–3-20
over horizontal aquiclude, 3-15–3-16
- Dynamic imaging, of tracer and contaminant plumes, 15-24–15-25
- Dynamic plant water stress, 29-23
- Dynamic viscosity, 2-18, 3-8, 13-4–13-5
- DYNAMIX, 22-35
- E**
- E. Coli*, 31-12
- Earth materials

- minerals and rocks, 2-7–2-8
unconsolidated materials, 2-8–2-10
- Ecohydrology, 29-1
photosynthesis and transpiration, atmospheric controls
soil moisture deficit and plant water stress, 29-19–29-24
soil nutrient dynamics, 29-24–29-34
soil water balance, in vadose zone, 29-13–29-19
- Effective porosity, 2-10, 3-8
- Effective stress, 1-16
- Effective velocity, 18-6
- Egypt, 1-3, 2-7
- Electric logs, 14-17–14-19
- Electric Network Analogen for Groundwater Flow (ELNAG), 1-19
- Electrical double-layer, 17-9
- Electrical/mechanical control panel, 36-25
- Electrical resistance methods, 6-17
- Electrical resistance tomography (ERT), 14-5
- Electroacoustical decontamination, 36-29
- Electromagnetic (EM) induction, for hydrogeological characterization, 14-7–14-9
- Electron, 17-25
- Electron acceptors, 31-4–31-5, 31-8
nitrate, 31-5–31-6, 31-21
- Electron transport phosphorylation, 31-3–31-4
- Electrostatic forces of attraction and water states, in pores, 2-14
- Elementary isotopes, 17-24
- Elevation head, 3-3, 3-6, 3-8
- Engineered bioremediation, 31-21
- England, 3-28
- Enhanced *in situ* bioremediation (EISB), 36-32
- Entamoeba histolytica*, 31-2
- Envelope potential, 6-6
- Environmental conditions, influencing biotransformations, 31-6–31-10
limits, 31-10–31-11
- Environmental monitoring, in landfill projects, 33-25–33-26
- Environmental problems, modern
anthropogenic changes, in recharge areas, 2-51–2-52
artesian pressure and surface effects, 2-50
artificial drainage effects, 2-51
chemical changes and surface-water interactions, 2-51
climatic change, possible effects, 2-46–2-48
construction, groundwater influence, 2-52
excess evaporation and salt build up, 2-50
groundwater overdraft and consequences, 2-50
irrigation effects, 2-52–2-53
land subsidence, 2-49
mining effects, 2-50–2-51
soil liquefaction, 2-49
wetlands, 2-48–2-49
- Environmental tracers, 1-28
- EPA, 17-24, 30-10, 36-2
National Primary Drinking Water Standards, 32-11–32-12
- Equilibrium flow, *see* Steady-state flow
- Equipotential lines, 3-20–3-21, 5-6–5-9
- Equivalent porous media (EPM), 20-2
models, 20-37
for karst aquifer, 21-38
- ERDAS, 30-2
- Ergodicity, 19-4–19-5
- Erkelenz, 1-3
- Erosion control products, 37-24
- Erosion layer, 33-14
- Ethanol, 31-8, 31-11
- Eulerian and Lagrangian perspectives, 18-3–18-4
- Euphrates, 1-2
- Europe, 31-9
- Evapotranspiration, 1-24, 1-26, 2-13, 2-38–2-39, 2-48, 22-2, 28-8
- Evapotranspirative cover systems, for waste containment, 34-1
case studies, 34-19–34-27
cover types
anisotropic barriers, 34-12–34-13
capillary barriers, 34-11–34-12
monolithic covers, 34-9–34-11
- design
numerical modeling issues, 34-16
performance criteria and regulatory issues, 34-13–34-14
strategies, 34-13
variables, 34-14–34-16
- performance monitoring
compliance demonstration, regulatory issues, 34-17
lysimetry, 34-17
meteorological variables and overland runoff, 34-18–34-19
moisture and suction profiles monitoring, 34-17–34-18
strategies, 34-16
- unsaturated soils
hydraulic concepts, 34-3–34-5
hydraulic conductivity functions, 34-7–34-9
hydraulic properties, 34-5–34-7
- Everglades, 3-25
- Evolutionary models, for karst aquifer, 21-37–21-38
- Ex situ* treatment
biological processes, 36-41
chemical and physical processes, 36-42–36-43
volatilization processes, 36-41–36-42
- Excess evaporation and salt build up, 2-50
- Exit gradient, 5-46–5-47
- Exoenzyme, 29-34
- Expected spatial average infiltration rates, 8-7
- Exploitation wells, 11-1
- Extraction systems design, 36-21–36-30
aquifer restoration time, 36-26–36-28
extraction trench systems, 36-25–36-26
extraction well systems, 36-23–36-25
other approaches, 36-28–36-30
- Extraction trench systems, 36-25–36-26
- Extraction well systems, 36-23–36-25
- Extrapolation, 25-8, 25-12
- F**
- Fahrenheit scale, 13-2
- Falling head permeameter, 3-11
- FAO, 1-27
- Fault, 20-3, 33-4
- FEAS code, 12-15
- Federal and Indian groundwater rights, 32-6–32-7

- Federal Insecticide, Fungicide, and Rodenticide Act (FIFRA), 32-29
- Fermentation, 31-3
- Ferralsols, 6-23–6-24
- Fertigation, 2-51–2-52
- Fiber optic chemical sensors, 35-30
- Fiber reinforcement, 37-6–37-7
- Fick's law, 18-4, 20-8, 22-9, 23-6, 3-9
- Field capacity, 1-14, 2-18
- Field chemical analyses, 35-34
- Field coefficient, of permeability, 2-20
- Field scale infiltration, 8-3, 8-11
- Field techniques, 20-9
- borehole
 - geophysical methods, 20-12
 - instrumentation, 20-23–20-25
 - core analysis, 20-10–20-12
 - fracture mapping, 20-10
 - hydraulic testing methods, 20-13
 - multi-well hydraulic tests, 20-17–20-18
 - point dilution method, 20-18–20-22
 - single-well hydraulic tests, 20-13–20-16
 - tracer experiments, 20-22–20-23
- Field-scale dispersion, 23-7
- Filter pack, 35-16, 35-21–35-22
- Filtration function, geosynthetic, 37-7–37-8
- Filtration theory, 22-16
- Fingering, 9-2, 18-3, 22-5, 28-10
- Finite-difference methods, 23-8–23-10, 23-10–23-13, 30-12
- Finite Element Aquifer Simulator, *see* FEAS code
- Finite-element methods (FEM), 23-8–23-10, 23-13–23-15, 30-12
- First-order approximation, 9-8–9-9
- Fission fragments, 17-28–17-29
- Fixed-head packers, 20-13
- Flanders, 1-3
- Flexible polypropylene (fPP), 37-12, 37-14
- Flood plains, 2-9, 33-4
- Florida, 1-19, 2-38, 3-25, 3-28
- Flow analysis and aquifer tests, 1-16–1-18
- Flow and transport, modeling, 20-31–20-38
- Flow equation
- aquitard, 4-24
 - confined aquifer, 4-19–4-20
 - phreatic aquifer, 4-20–4-24
- Flow mechanism, 8-2
- Flow meter, 20-14
- Flow model, of sea water intrusion, 12-12–12-13
- Flow net, 3-20–3-22, 5-21–5-25
- Flow systems, topography driven, 1-22
- Flow work, 3-7
- Flowing artesian well, 2-25
- Fluid mechanics, 3-3, 4-1
- Fluorescein, 21-30
- Fluorescent dyes, 21-29
- Flushing, 36-29
- FLUTE[®], 20-24–20-25
- Flux, 20-34
- boundary condition, 4-13
- Flux-concentration relation, 6-36
- Fokker-Planck equation, 6-32
- Force potential, 1-20, 3-20
- Forced gradient methods, 15-7
- Forward modeling, with parameter error, 14-31–14-32
- Fossil water, 2-36
- Fourier transformation, 3-9, 18-8–18-10, 22-23–22-24
- Fractional Advection-Dispersion (FADE), 18-7–18-8, 18-14
- Fracture, 20-2
- Fracture flow, 21-9
- Fracture mapping, 20-10
- Fracture-matrix interface
- dual-porosity models
 - for gradual change, 9-9–9-23
 - for sharp change, 9-3–9-9
- Fracture network, 20-28–20-31
- Fracture zones, 20-3
- Fractured media, groundwater flow and solute transport, 20-1
- conceptual models, 20-25–20-31
- discrete fracture networks, flow in, 20-33–20-35
 - flow and transport, modeling, 20-31–20-38
 - fracture network, 20-28–20-31
 - single fracture, 20-26–20-28
 - solute transport, in discrete fracture network, 20-35–20-38
- field techniques, 20-9
- borehole geophysical methods, 20-12
 - borehole instrumentation, 20-23–20-25
 - core analysis, 20-10–20-12
 - fracture mapping, 20-10
 - hydraulic testing methods, 20-13
 - multi-well hydraulic tests, 20-17–20-18
 - point dilution method, 20-18–20-22
 - single-well hydraulic tests, 20-13–20-16
 - tracer experiments, 20-22–20-23
 - single fracture, groundwater flow, 20-4–20-7
 - solute transport, in single fracture, 20-7–20-9
 - structural geology, of fractured rock, 20-2–20-4
- France, 2-7, 2-9, 2-20, 2-25, 2-38
- Freeze-thaw process, 28-11
- Fresh-salt water interface problems, 1-18–1-19
- Fresno Case Study, 25-12–25-14
- Friction head loss, 3-13
- Frontinus, 1-4
- Fuel oxygenates, 17-24
- Fulvic acid, 17-20
- Function simulation, 16-6
- Functional groups, oxygenated, 17-20
- Functional hierarchy, 18-2
- Functional normalization, 6-41
- Functional regression methods, 6-23–6-25
- Fungi, 31-2
- Funnel and gate reaction barrier, 36-38
- Funnel flow, 9-3
- Furans, 31-8
- G**
- Gage pressure, 3-2–3-3
- Gaining streams, 2-40–2-42
- Galerkin method, 12-11, 22-31
- Gamma-gamma log, 14-19
- Gamma log, 14-19
- Gamma particles, 17-25
- Gamma rays attenuation method, 6-17
- Gardner soil, 7-11
- Gas management system, landfill gas, 33-17–33-19

- Gas monitoring, in landfill projects, 33-25
Gaseous-phase analyses, 35-36
Genasys, 30-2
Generalized linear estimation equations
 BLUE, 16-14
 model, 16-13
 parameter estimation, 16-14
Generalized Master Equation (GME), 18-6-18-7
Generic model, 23-2
Geocells, 37-23-37-24
Geochemical setting, 35-8-35-9
Geochemical transport models
 heavy metals leaching, from soil column, 22-36
 HP1, 22-35-22-36
Geocomposite sheet drains, 37-20-37-22
Geocomposite strip (wick) drains, 37-22-37-23
Geographical information system (GIS), in groundwater
 engineering, 25-7, 30-1
 analysis capabilities, 30-3-30-4
 application
 groundwater movement modeling, 30-12-30-13
 groundwater occurrence mapping, 30-6
 policy implementation, involving spatial
 dimensions, 30-13-30-14
 spatial data analyses, using spatial statistics,
 30-7-30-8
 spatial data aspects management, of groundwater
 projects, 30-6
 surface fitting and interpolation, 30-8-30-9
 visualization, 30-9
 water vulnerability modeling, using spatial data,
 30-9-30-12
components
 data, 30-3
 hardware, 30-2-30-3
 software, 30-2
data representation, 30-3
 web services, 30-5-30-6
Geogrids, 33-23, 37-16-37-18
Geologic information, 35-9
Geological aspects, 1-12-1-13
Geological contact springs, 2-37
Geological occurrence, of groundwater, 2-2
 anisotropy and heterogeneity, 2-22-2-25
 aquifers, 2-25-2-28
 capillarity, 2-12-2-13
Earth materials
 minerals and rocks, 2-7-2-8
 unconsolidated materials, 2-8-2-10
electrostatic forces of attraction
 and water states, in pores, 2-14
environmental problems, modern
 anthropogenic changes, in recharge areas,
 2-51-2-52
 artesian pressure and surface effects, 2-50
 artificial drainage effects, 2-51
 chemical changes and surface-water interactions,
 2-51
 climatic change, possible effects, 2-46-2-48
 construction, groundwater influence, 2-52
 excess evaporation and salt build up, 2-50
 groundwater overdraft and consequences, 2-50
 irrigation effects, 2-52-2-53
 land subsidence, 2-49
mining effects, 2-50-2-51
soil liquefaction, 2-49
wetlands, 2-48-2-49
groundwater importance, 2-3-2-7
groundwater supplies protection, 2-42-2-46
 artificial protective barriers, 2-45
 natural protective barriers and waste
 disposal, 2-44
 water laws, 2-43-2-44
 well-head protection programs, 2-45-2-46
hydraulic gradient, 2-20
hydraulic head, 2-16
intrinsic permeability, 2-18
moisture content, 2-12
porosity, 2-10-2-12
saturated media, hydraulic conductivity,
 2-20-2-22
solid earth materials compressibility, 2-15
storage, 2-16-2-18
viscosity, 2-18-2-19
water, in saturated zone, 2-33
 bank storage, 2-42
 residence time, 2-35-2-36
 streams, gaining and losing, 2-40-2-42
 surface discharge, 2-36-2-40
water, in unsaturated zone
 atmospheric discharge and seepage face,
 2-32-2-33
 hydraulic conductivity and specific discharge,
 2-31-2-32
 moisture content vs depth, 2-28-2-29
 recharge and infiltration capacity, 2-30-2-31
 residence time, 2-32
 subsurface stormflow, 2-32
 water compressibility, 2-14-2-15
GeoMedia, 30-2
Geomembrane, 34-17, 36-21, 37-12-37-16
Geophysical and tracer data, combining,
 15-18-15-24
 geophysical and hydrogeologic properties relation,
 estimating, 15-23-15-24
 geostatistical methods, 15-19-15-21
 zonal inversion methods, 15-21-15-23
Geophysical methods, 35-13
 borehole geophysical methods, 14-16-14-20
 acoustic logs, 14-19-14-20
 caliper logs, 14-17
 electric logs, 14-17-14-19
 nuclear logs, 14-19
 to environmental problems
 surface seismic reflection and refraction surveys,
 15-7-15-10
forward modeling, with parameter error,
 14-31-14-32
for hydrogeological characterization
 DC resistivity methods, 14-3-14-7
 electromagnetic induction, 14-7-14-9
 ground penetrating radar, 14-9-14-11
 seismic methods, 14-11-14-15
hydrogeophysical concepts, 14-20-14-29
 hydrogeological mapping, 14-21-14-25
 hydrogeological parameter estimation,
 14-25-14-29
 hydrological process monitoring, 14-29

- Geophysical methods (*Continued*)
 inverse modeling, 14-32–14-38
 Bayesian estimators, 14-32–14-33
 maximum likelihood method, 14-34–14-35
 maximum a posteriori approach, 14-35–14-38
 prior probability, defining, 14-33
 model selection, 14-42–14-44
 point estimation, with known parameters
 Bayesian point estimators, 14-39–14-42
 potential field methods
 gravitational methods, 14-15–14-16
 magnetic methods, 14-16
- GeoServ, 30-5
- Geostatistics, 16-1, 25-7
 cokriging, 16-13
 generalized linear estimation equations
 BLUE, 16-14
 model, 16-13
 parameter estimation, 16-14
 inverse problems, 16-14–16-17
 derivative computation, 16-17
 methodology, 16-15–16-16
 kriging variants, 16-11–16-12
 ordinary kriging
 conditional realization, 16-5–16-6
 function estimate, 16-4–16-5
 interpolation, 16-3–16-4
- preliminaries
 least squares, 16-2–16-3
 model, 16-2
 stochastic approach, 16-2
- variogram
 calibration, 16-11
 experimental variogram, 16-8–16-9
 model, 16-6–16-7
 nonnegativity, 16-11
 residuals, 16-9
 selection, 16-7–16-8
 testing, 16-10–16-11
- Geosynthetic, 37-1
 applications, in landfill designs, 37-27–37-28
 case histories
 multiple use, in landfill cover design
 vertical barrier system, 37-28–37-30
- functions
 barrier function, 37-9–37-10
 design by, 37-3
 drainage function, 37-8–37-9
 filtration function, 37-7–37-8
 protection function, 37-10
 reinforcement function, 37-4–37-7
 separation function, 37-3–37-4
- types
 erosion control products, 37-24
 geocells, 37-23–37-24
 geocomposite sheet drains, 37-20–37-22
 geocomposite strip (wick) drains, 37-22–37-23
 geogrids, 33-23, 37-16–37-18
 geomembranes, 34-17, 36-21, 37-12–37-16
 geosynthetic clay liners, 33-13, 37-18–37-20
 geotextiles, 33-24, 33-26, 34-17, 37-7, 37-10–37-12
 HDPE vertical barrier systems, 37-24–37-27
- Geosynthetic clay liners, 33-13, 37-18–37-20
- Geotechnical factors, 33-5, 33-26
- Geotextiles, 33-24, 33-26, 34-17, 37-7, 37-10–37-12
 cross machine direction, 37-12
 machine direction, 37-11–37-12
 manufacturing, 37-11
 selvage, 37-12
- Geothermal energy, for electric power, 13-7–13-9
- GeoTools, 30-5
- Ghyben–Hertzberg lens, 3-28
- Ghijben–Herzberg principle, 1-9
- Ghyben–Herzberg rule, 12-3
- Giardia*, 31-2
- Giardian lambla*, 17-30
- GIS-MODFLOW, 30-12
- Glacial drift, 2-8
- Glacial ice, 2-8
- Glacial till, 2-8
- GLEAMS (Groundwater Loading Effects of Agricultural Management Systems), 30-10–30-11
- Global neighborhood approach, 16-4
- Global positioning system (GPS) 15-11, 30-3
- GMS software, 30-9
- Gobi Desert, 2-10
- Golden, 15-21
- Good Friday earthquake, 2-49
- Grafton County, 15-25
- Grand Island, 1-16
- Grand River watershed, case study, 28-20
 conclusions, 28-33–28-34
 input data, 28-21–28-25
 land use and land cover data, 28-22
 soil data, 28-22–28-23
 vegetation data, 28-23–28-24
 weather data, 28-24–28-25
 methodology, 28-25–28-26
 results, 28-26–28-33
- GRASS, 30-2
- Gravel, 2-10
 deposits, 2-8–2-9
 pack, 11-19–11-20
- Gravimetric method, 6-17
- Gravitational methods, for hydrogeological characterization, 14-15–14-16
- Gravitational water, 2-14
- Gray, 17-26
- Great Basin, 2-38
- Great Lakes, 2-35, 2-38
- Greece, 1-3, 2-7
- Green and Ampt soils, 6-38–6-39
 and Gardner soil, dynamical similarity, 6-45, 6-48
- Greenhouse effect, 2-47
- Greenhouse gas, 2-47
- Green River, 21-6
- Green's function, 18-8–18-10
- Grid design, 23-21
- Ground penetrating radar (GPR), 14-9–14-11, 14-34, 15-10, 15-14, 15-24
- Groundwater allocation and management
 federal and Indian groundwater rights, 32-6–32-7
 property takings issues, 32-7
 regulation, need for, 32-3–32-4
 state groundwater rights systems
 absolute dominion, 32-4–32-5
 conjunctive management, 32-6
 correlative rights, 32-5–32-6

- prescriptive rights, 32-6
- prior appropriation, 32-6
- reasonable use, 32-5
- torts restatement, 32-5
- Groundwater and seepage**
 - flow net, 5-21–5-25
 - groundwater fundamentals, 5-2–5-9
 - Bernoulli's equation, 5-3
 - Darcy's law, 5-3–5-5
 - homogeneity and isotropy, 5-5
 - Reynolds number, 5-5
 - streamlines and equipotential lines, 5-6–5-9
 - introduction, 5-1–5-2
 - layered systems, flow in, 5-37–5-43
 - method of fragments, 5-31–5-37
 - piping, 5-43–5-46
- Point Estimate Method (PEM)**
 - one random variable, 5-25–5-27
 - regression and correlation, 5-27–5-30
 - several random variable, 5-30–5-31
- probabilistic fundamentals**, 9
 - moments, 5-11–5-17
 - probability distribution, 5-17–5-21
- Groundwater circulation well (GCW) technologies**, 36-40
- Groundwater contaminants**
 - aqueous solubility, affecting factors $p\epsilon$, 17-11–17-13
 - pH and chemical composition, 17-3–17-10
 - sorption, to solids, 17-10–17-11
 - temperature, 17-13
- inorganic constituents
 - source water, soil interaction, 17-17–17-19
 - source water composition, 17-14–17-17
- organic constituents
 - background organics, 17-19–17-20
 - fuel oxygenates, 17-24
 - high-profile organic contaminants, 17-21–17-24
 - nonaqueous phase liquids (NAPLs), 17-20–17-21
- particulates, 17-29–17-30
- radionuclides
 - radioactive decay, 17-24–17-25
 - radioactive waste disposal, 17-28–17-29
 - remediation, 17-29
 - sources, 17-26–17-28
 - units, 17-26
 - water physicochemical characteristics, 17-2–17-3
- Groundwater development, quantitative**
 - maintaining, 27-2
- Groundwater dynamics**, 1-6–1-12
- Groundwater flow equations**, 4-10–4-15, 23-4–23-5
 - boundary conditions, 4-13–4-15
- Groundwater fundamentals**, 5-2–5-9
 - Bernoulli's equation, 5-3
 - Darcy's law, 5-3–5-5
 - homogeneity and isotropy, 5-5
 - Reynolds number, 5-5
 - streamlines and equipotential lines, 5-6–5-9
- Groundwater mining**, 2-42, 23-1
 - case histories
 - contributing areas delineation, 23-35–23-39
 - sea water intrusion, in coastal aquifer, 23-39–23-44
 - shallow unconfined aquifer, flow and transport, 23-31–23-35
- equations**, 23-3–23-7
 - groundwater-flow equation, 23-4–23-5
 - seepage velocity, 23-5
 - solute-transport equation, 23-5–23-7
- flow and transport processes**, 23-3
- groundwater software**, 23-44–23-45
- model design and application**, 23-19–23-28
 - calibration criteria, 23-26–23-27
 - grid design, 23-21
 - mass balance, 23-25
 - model calibration, 23-22–23-23
 - model error, 23-23–23-25
 - model validation, 23-27–23-28
 - predictions and postaudits, 23-27
 - sensitivity tests, 23-25–23-26
- models**, 23-2–23-3
 - MODFLOW-2000, 23-28–23-30
 - numerical methods, 23-7
 - boundary and initial conditions, 23-17–23-18
 - finite-difference methods, 23-10–23-13
 - finite-element methods (FEM), 23-13–23-15
 - generic model verification, 23-19
 - matrix solution techniques, 23-17
 - method-of-characteristics methods (MOC), 23-15–23-17
 - SUTRA, 23-30–23-31
- Groundwater model validation**, 24-1
 - Amchitka model, uncertainty reduction, 24-18–24-22
 - background, 24-19–24-20
 - Bayesian framework, 24-20–24-21
 - Milrow site, 24-21–24-22
 - practical views, 24-2–24-3
 - Project Shoal Area (PSA) model, 24-6–24-18
 - acceptance criteria, 24-7–24-8
 - multiple validation targets, 24-12–24-18
 - single validation target, 24-8–24-12
 - stochastic numerical models, 24-3–24-6
- Groundwater monitoring**, 35-1
 - data evaluation, 35-37
 - contaminant indications, responding to, 35-42–35-43
 - groundwater depths, 35-37–35-39
 - groundwater elevation measurement, 35-37–35-39
 - quality data, 35-39–35-42
 - devices, design and installation
 - cone penetrometer testing, 35-23
 - direct push technology sampler, 35-23
 - groundwater discharge, 35-23
 - lysimeters, 35-24
 - piezometers, 35-22–35-23
 - soil gas probes, 35-24
 - wells, 35-14–35-22
 - groundwater sampling
 - documentation, 35-34
 - field chemical analyses, 35-34
 - general approach, 35-24–35-25
 - in situ* monitoring, 35-30
 - laboratory chemical analyses, 35-34–35-37
 - QA/QC samples, 35-30–35-31
 - sample handling and preservation, 35-31–35-34
 - techniques, 35-25–35-30

- Groundwater monitoring (*Continued*)
 in landfill projects, 33-25
 plans
 developing, 35-7–35-13
 purpose, 35-2
 assessment monitoring programs, 35-3–35-4
 corrective action monitoring programs, 35-4
 detection monitoring programs, 35-2–35-3
 performance monitoring programs, 35-4
 regulatory requirements for, 35-4–35-5
 statistical methods, 35-41
 strategy considerations, 35-5–35-7
- Groundwater overdraft and consequences, 2-50
- Groundwater quality, 3-29–3-31
- Groundwater whirl, 1-17
- Gulf of Thailand, 15-9
- H**
- H_2CO_3^* , 17-5
- H_2S , 17-5
- Hagen–Poiseulle equation, 21-10
- Halogenation
 compounds, 31-8
- Halophilic bacteria, 31-6
- Halorespiration, 31-9
- Hantush leaky well flow equation, 4-31–4-33
- Harris–Galveston area, 3-25
- Haverkamp–Parlange model, 6-29
- Hazardous waste, RCRA regulations, 32-17–32-21
 corrective action program, 32-18–32-21
 facility standards, 32-18
 land disposal restrictions, 32-18
- HDPE, *see* High density polyethylene
- Head loss, 2-35
- Headwater, 5-22
- Heat dissipation units (HDUs), 34-18
- Heat flow, 13-1
 advection versus convection, 13-6
 applications
 different aquifers, distinguishing, 13-13
 formation water expansion, 13-13
 heat distributions, 13-11
 hot and cold water storage, 13-13
 temperature/depth profiles, in water wells, 13-11–13-13
- basic thermodynamics
 heat as energy, 13-2
 heat units, 13-3
 specific heat of water, 13-3–13-4
 temperature, 13-2–13-3
 thermal and mechanical energy, relationship, 13-5
 thermal conductivity, 13-3
 thermal inertia, 13-4
 thermal resistivity, 13-3
 3D hexagon structures, of water, 13-4
 viscosity and heat, 13-4–13-5
- conduction, of heat, 13-5
- convection, of heat, 13-5–13-6
- different aquifers, distinguishing, 13-13
- diffusion, of heat, 13-6
- formation water expansion, 13-13
- heat distributions, 13-11
- heat flux, 13-5
 hot and cold water storage, 13-13
- into soil, 29-19
- sources to and from Earth's surface
 air-temperature changes, annual, 13-9
 geothermal energy, for electric power, 13-8–13-9
 geothermal gradient, 13-7
 heat groundwater effects, upon hydrogeochemistry, 13-10–13-11
 hydrothermal activity, 13-8
 permafrost, 13-9–13-10
 solar radiation, 13-6–13-7
 surface water and groundwater, heat exchanges, 13-10
- temperature/depth profiles, in water wells, 13-11–13-13
- thermal gradients, 13-5
- Heat flux, 13-5
- Heat groundwater effects, upon hydrogeochemistry, 13-10–13-11
- Heavy metals leaching, from soil column, 22-36
- Hellenic–Roman civilization, 1-3–1-4
- HEM, 19-6–19-7
- Henry's law, 17-4–17-5, 17-14–17-15
- Heterogeneity, 19-4–19-5, 22-18, 23-2, 23-8
 hierarchical, 18-2
 in reactive properties, 19-20–19-21
- Heterogeneity driven dispersion, 18-3
- Heterogeneous aquifers, interpretation, 10-30–10-31
- Heterotrophs, 31-2
- High density polyethylene (HDPE), 37-12, 37-14
 vertical barrier systems, 37-24–37-27
 case history, 37-28–37-30
- High-level waste (HLW), 17-28
- High Plains Aquifer, 2-4, 2-52
- High-profile organic contaminants, 17-21–17-24
- High-velocity jetting, 11-27
- History, of groundwater hydrology, 1-1
 Hellenic–Roman civilization, 1-3–1-4
 hydraulics and hydrology, 1-5
 hydrochemistry, 1-25–1-27
 isotope hydrology, 1-27–1-28
 regional groundwater systems
 basic-scale analysis and evaluation, 1-20–1-21
 flow systems, topography driven, 1-22
 land drainage and groundwater discharge, 1-22–1-24
 runoff generation, on sloping areas, 1-24–1-25
- subsurface hydrology, developments in
 aquifer exploration, 1-14–1-16
 flow analysis and aquifer tests, 1-16–1-18
 fresh–salt water interface problems, 1-18–1-19
 geological aspects, 1-12–1-13
 groundwater dynamics, 1-6–1-12
 hydrogeology consolidation, 1-19–1-20
 soil water physics, 1-13–1-14
- water management, in ancient societies, 1-2–1-3
 water science evolution, during renaissance, 1-4–1-5
- History matching, 16-14
- HO model, 7-9–7-10
- Hollow stem auger, 35-19
- Homogeneity and isotropy, 5-5
- Horizontal drains, 36-20–36-21
- Hortonian runoff, 8-2–8-3
- Hot Springs, Arkansas, 2-38

- HP1, 22-35–22-36
H-sigma bounds, 5-14, 5-16
Humic substances, 17-19–17-20
Hunt River Basin, 23-35
Hvorslev slug test equation, 4-27–4-28
Hydraulic approach, to groundwater flow
 aquitard, flow equation, 4-24
 confined aquifer, flow equation, 4-19–4-20
 motion equation, 4-15–4-19
 phreatic aquifer, flow equation, 4-20–4-24
Hydraulic barriers, for hydraulic containment, 36-17–36-20
 collection trenches, 36-17–36-18
 extraction wells, 36-18–36-20
Hydraulic civilization, 1-2
Hydraulic conductivity, 6-9, 6-19–6-21, 21-9
 functions, 6-15–6-16
 instantaneous profile method, 6-19–6-20
 pressure ring infiltrometer, 6-21
 and specific discharge, 2-31–2-32
 and suction, relationship, 34-7–34-9
 tension disk infiltrometer, 6-20–6-21
Hydraulic containment, 36-13–36-21
 alternative approaches, 36-20–36-21
 hydraulic barriers, 36-17–36-20
 physical barriers, 36-14–36-17
Hydraulic diameter, 20-6
Hydraulic gradient, 2-20, 3-5, 5-3
Hydraulic head, 2-16, 3-6, 3-8
 measurement, 15-3
Hydraulic potential, 6-6
Hydraulic properties, of soil, 6-5–6-10
 hydraulic conductivity, 6-9
 soil water content, 6-5–6-6
 soil water pressure, 6-6–6-7
 water retention characteristic, 6-7–6-9
Hydraulic rotary method, 2-7
Hydraulic testing methods, 20-13
Hydraulics and hydrology, 1-5
Hydrocarbons, 31-5, 31-7–31-8
Hydrochemistry, 1-25–1-27
HYDROCOIN Project, 23-20
Hydrodynamic dispersion, 3-34, 23-6
Hydrofacies, 15-21
Hydrogen bonds, 17-2
Hydrogen peroxide, 31-5, 31-21
HYDROGEOCHEM, 22-35
Hydrogeologic factors, 33-5
Hydrogeologic setting, 35-8–35-9
Hydrogeology, analytical methods, 27-2–27-3
Hydrogeology consolidation, 1-19–1-20
Hydrogeophysical concepts, 14-20–14-29
 hydrogeological mapping, 14-21–14-25
 hydrogeological parameter estimation, 14-25–14-29
 hydrological process monitoring, 14-29
Hydrographs, of karst springs, 21-31–21-34
Hydrology
 spatial scale, 25-2
 temporal scale, 25-2
Hydrophobicity, 31-10
Hydrothermal activity, 13-8
HYDRUS software package, 22-33
 HYDRUS-1D, 22-32–22-34, 34-16
 HYDRUS-2D, 22-29, 22-32
Hygroscopic coefficient, 2-14
Hygroscopic water, 2-14
Hypochlorite, 11-29
Hypothetical solubility, 31-15
Hysteresis, 1-14, 2-29, 6-7, 6-9
- I**
- Idaho, 2-38
IDRISI, 30-2
Igneous rocks
 extrusives, 2-7
Immiscible contaminant, 3-29
In situ treatment, 36-31–36-41
 biological remediation, 36-32–36-34
 chemical and physical processes, 36-38
 innovative treatment approaches, 36-40–36-41
 monitored natural attenuation (MNA), 36-38–36-40
 thermal treatment, 36-37–36-38
 volatilization processes, 36-34–36-38
India, 1-3
Indiana, 2-8, 2-30, 2-38, 2-44, 2-49, 2-52, 30-6
Indicator seismic velocity thresholds, 15-21
Indirect recharge, 28-6
Infiltrability-depth approximation (IDA), 7-4
Infiltration, 1-24, 2-30, 3-2, 6-33–6-34, 7-1, 28-6, 28-10
 capacity, 2-30–2-31
 infiltration equations
 empirical equations, 7-9–7-10
 Parlange et al. model (PA), 7-8–7-9
 Philip infiltration equation (PH), 7-7–7-8
 Swartzendruber (SW) model, 7-9
layer, 33-14
measurement methods
 permeameter method, 7-11
 ring infiltrometer method, 7-10–7-11
 sprinkler method, 7-10
rate prediction, 7-4
into sealed or crusted soils, 7-12–7-14
soil properties and process of infiltration, 7-2–7-7
spatial variability considerations, 7-11–7-12
Injection-withdrawal format, 20-22
Innovative *in situ* treatment approaches, 36-40–36-41
 groundwater circulation well (GCW) technologies, 36-40
 nanoscale zero-valent iron (nZVI), 36-40
 surfactant enhanced aquifer remediation (SEAR), 36-40–36-41
Inorganic contaminants
 source water
 composition, 17-14–17-17
 soil interaction, 17-17–17-19
Input-output models, for karst aquifer, 21-38–21-39
Instantaneous profile method, 6-19–6-20
Institute of Land and Water Management Research (ICW), 1-22
Integrated Groundwater Surface Water Model (IGSM), 28-19
Interface
 in confined aquifer, 12-4–12-5
 in phreatic coastal aquifer, 1-5
Interference tests, 20-17
Intergovernmental Panel on Climate Change (IPCC), 28-2, 28-29
 global climate predictions, 28-3

- Intermediate layer, 33-15
 International Association of Hydrogeologists (IAH), 1-20, I-28
 International Association of Hydrological Sciences (IAHS), 1-20
 International Association of Scientific Hydrology (IASH), 1-14, 1-20
 International Atomic Energy Agency (IAEA), 1-28
 International Hydrological Decade (IHD), 1-20
 International Soil Science Society, 6-3
 Interpolation and surface fitting, 30-8–30-9
 Interpolation, *see* Geostatistics
 Intrinsic bioremediation, 31-20–31-21
 Intrinsic permeability, 1-20, 2-18, 3-8
 Inverse distance weighting, 30-12
 interpolation, 30-8–30-9
 Inverse flow modeling, 15-21
 Inverse modeling, 14-32–14-38, 23-22
 Bayesian estimators, 14-32–14-33
 maximum a posteriori approach, 14-35–14-38
 maximum likelihood method, 14-34–14-35
 prior probability, defining, 14-33
 Inverse problem, 3-20, 16-14–16-17
 derivative computation, 16-17
 methodology, 16-15–16-16
 Ion exchange, 17-10
 Ionizing radiation, 17-25–17-26
 Ion-selective probes, 35-30
 Iowa, 2-10
 Iran, 1-3
 Iron ore, 2-51
 Irrigation effects, 2-52–2-53
 Island aquifer, 27-3, 27-5
 Isothermal compressibility, of water, 2-14–2-15
 Isotope hydrology, 1-27–1-28
 Isotropic media, 3-10
 Isotropic tracers, 1-27
 Israel, 1-19, 3-28
 Iterative methods, for algebraic equation, 23-17
- J**
 Jacob's method
 in step-drawdown tests, 11-8–11-11
 Japan, 17-28
 Jetting, 35-19
 Joints, 20-3
 JP-4 jet fuel, 31-21
 Jshape, 30-5
- K**
 K-function, 34-7–34-9
 Kankakee River Valley, 2-49
 Kansas, 2-52, 3-25
 Karst, 2-7, 3-25
 Karst springs, 2-38
 Karst terrane, 2-9
 Karstic aquifers, 20-3
 carbonate rock dissolution
 chemical equilibrium, 21-17–21-20
 chemical parameters, 21-20–21-21
 limestone and dolomite, dissolution kinetics of, 21-21–21-24
 conceptual framework
- allogenic recharge, 21-2–21-3
 conceptual model, 21-7
 dispersed infiltration and epikarst, 21-4–21-5
 geologic boundary conditions, 21-6–21-7
 groundwater basins, 21-2
 groundwater discharge, 21-6
 infiltration pathways, in vadose zone, 21-5
 internal runoff, 21-3–21-4
 subsurface drainage patterns, 21-5–21-6
 evolution
 conduit life history, 21-24–21-25
 enlargement, maturity, and decay, 21-26–21-27
 initiation and critical thresholds, 21-25–21-26
 modeling
 coupled continuum pipe flow models, 21-38
 equivalent porous media models, 21-38
 evolutionary models, 21-37–21-38
 input-output models, 21-38–21-39
 pipe flow models, 21-38
 permeability and flow dynamics
 caves and conduits, 21-14–21-16
 conduit permeability and flow dynamics, 21-9–21-12
 matrix and fracture flow, 21-8–21-9
 paleohydrology, 21-16–21-17
 sediment fluxes and transport mechanisms, 21-12–21-14
 quantitative hydrology
 base flow/runoff relations, 21-28
 chemographs of karst springs, 21-34–21-37
 hydrographs, of karst springs, 21-31–21-34
 tracer studies, 21-28–21-30
 water budget, 21-27–21-28
 well tests and water table mapping, 21-30–21-31
 sinkhole development
 hydrologic factors, 21-42
 as landuse hazard, 21-41
 mechanisms, 21-41
 remediation, 21-42–21-43
 water resource issues
 contaminants and transport, 21-40–21-41
 sinkhole flooding, 21-39–21-40
 water quality monitoring, 21-41
 water supplies, 21-39
 Kelvin scale, 13-2–13-3
 Kentucky, 2-9, 2-38
 Kesterson Aquifer, 15-7, 15-17, 15-19, 15-22–15-24
 Kinematic viscosity, 2-19, 13-4–13-5
 Kinetic energy, 3-6
 Kinetic sorption models, 22-15
 Kostiakov model (KO), 7-10
 Kozeny–Carman equation, 1-8
 Kriging, 14-39, 26-8–26-9, 30-8
 variants, 16-11–16-12
 Kurtosis
 coefficient of kurtosis, 5-17
- L**
 Laboratory chemical analyses, 35-34–35-37
 Lacustrine materials, 2-9
 Lagrangian perturbation approach, 18-10–18-11
 Lagrangian stochastic models, for nonlinear reactions, 19-13–19-21
 heterogeneity, in reactive properties, 19-20–19-21

- reaction system examples, 19-18–19-20
stochastic-convective averaging, 19-16–19-17
streamtube formulation, 19-14–19-16
travel-time distribution function, 19-17–19-18
Lakes and ponds, dynamic role, 2-38
Land drainage and groundwater discharge, 1-22–1-24
Land subsidence, 2-49, 4-12
 calculation, 3-25–3-27
 seepage force, 3-27
Landfills, 3-39, 33-1
 Clean Air Act, 33-2–33-3
 Clean Water Act, 33-2
 conceptual design, 33-7–33-12
 borrow areas, 33-10
 cell layout and filling sequence, 33-11–33-12
 final cover grading, 33-10–33-11
 footprint, 33-8–33-9
 sedimentation/detention basins, 33-10
 subbase grading, 33-9–33-10
 configuration, 33-6–33-7
 detailed design
 alternative final covers, 33-21–33-22
 berms, 33-15
 bioreactor, 33-19–33-21
 final cover system, 33-13–33-15
 gas management system, 33-17–33-19
 leachate collection and removal, 33-16–33-17
 liner system, 33-12–33-13
 mechanical properties of waste, 33-12
 storm water management system, 33-16
elements, performance of, 33-26–33-27
environmental monitoring
 gas monitoring, 33-26
 groundwater monitoring, 33-25
hydrogeological and geotechnical considerations,
 33-5
linear and cover system materials
 geosynthetic materials, 33-23–33-24, 37-27–37-28,
 37-30–37-31
 geotextile, 37-10
 soil materials, 33-22–33-23
Site Master Plan, 33-5–33-6
siting
 federal considerations, 33-4
 state and local considerations, 33-3–33-4
state and local regulations, 33-3
 Subtitle D, 33-2
Laplace equation, 3-21, 4-32, 22-23
Las Vegas, 2-43
Law of continuity, 1-4
Law of hydrostatics, 3-3
Law of Tangents, 2-25
Layered systems, flow in, 5-37–5-43
Leachate, 37-27
 collection system, 33-16, 33-27
 removal system, 33-17, 33-27
LEACHM, 34-16, 34-22
Leaf area index (LAI), 28-9, 28-24, 29-3
Leaking underground fuel tanks (LUFTs), 17-21, 17-24
Leaky aquifer, 3-14–3-15, 10-2, 10-3
 distance-drawdown analyses, 10-28–10-30
 time-drawdown analyses, 10-22–10-24
 well-flow equations, 10-7–10-9
 see also Semi-confined aquifer
Least squares method, 15-2, 16-2–16-3
Legal framework for groundwater protection, in United States
 Clean Water Act (CWA), 32-7–32-10
 legislative background, 32-7–32-9
 regulations, 32-9
 Comprehensive Environmental Response,
 Compensation, and Liability Act (CERCLA),
 32-23–32-28
 legislative background, 32-23–32-25
 regulations, 32-25–32-28
 Federal Insecticide, Fungicide, and Rodenticide Act
 (FIFRA), 32-29
groundwater allocation and management
 federal and Indian groundwater rights, 32-6–32-7
 property takings issues, 32-7
 regulation, need for, 32-3–32-4
 state groundwater rights systems, 32-3–32-6
laws and regulations, relationship, 32-2–32-3
Resource Conservation and Recovery Act (RCRA),
 32-16–32-22
 legislative background, 32-16–32-17
 regulations, 32-17–32-22
Safe Drinking Water Act (SDWA), 32-10–32-16
 legislative background, 32-10–32-14
 regulations, 32-14
Surface Mining Control and Reclamation Act
 (SMCRA), 32-29
 surface water and groundwater, interaction, 32-3
Leonardo da Vinci, 1-4
Ligands, 17-7
Light NAPL, *see* LNAPL
Limestone, 2-9
Linear and nonlinear sorption, 22-11–22-12
Linear isotherm, 3-37–3-38
Linear law of groundwater flux and hydraulic gradient,
 1-5
Linear reservoir, 9-23
Linear well losses, 11-6
Linearity and nonequilibrium Eulerian stochastic
 models
 deterministic nonequilibrium sorption, 19-7–19-8
 random nonequilibrium sorption, 19-8–19-9
Liner system, 33-12–33-13
Liquid limit, of soil, 2-52
Lithitrophs, 31-2
Lithoautotrophs, 31-6
Lithologic heterogeneity, 15-12, 15-18
LNAPL, 17-21, 21-40, 35-37
Local equilibrium assumption (LEA), 9-4
Local neighborhood approach, 16-4
Local scale model, 8-5–8-6
Localized recharge, 28-6
Loess, 2-10
London, 17-10
Long Island, 1-20, 2-45
Long-term water balance, 29-17–29-19
Los Angeles, 2-45
Losing streams, 2-41
Low-level waste (LLW), 17-28–17-29
Low-stress approach, groundwater sampling, 35-27,
 35-28
L-THIA, 30-5
Luminescence/spectroscopic analyses, 35-36

- Lungeon tests, *see* Constant-head injection tests (CHIT)
 Lwoff A., 31-12
 Lysimeter, 28-12, 35-13, 35-24
 Lysimetry, 34-17
- M**
 MACRO, 22-32
 MacroDispersion Experiment Site (MADE), 15-3-15-4, 18-14
 Macrodispersion Experiment (MADE) site, in Columbus
 advection-dispersion model (ADM), 26-6, 26-10-26-11
 dual-domain mass transfer model (DDMT), 26-8, 26-11-26-12
 natural-gradient tracer tests, 26-4-26-6
 sensitivity analysis, 26-12-26-13
 site description, 26-2-26-4
 tritium tracer test
 hydraulic and transport properties assignment, 26-8-26-9
 model setup and boundary conditions, 26-8
 objectives and approaches, 26-6-26-8
 Magnesite, 2-9
 Magnetic methods, for hydrogeological characterization, 14-16
 Magnetrötelluric (MT) measurements, 24-19-24-20
 Manganese, 2-51
 Manning equation, 21-11
 MapGuide, 30-5
 MapInfo, 30-2
 Maps, 30-6
 MapServer, 30-2, 30-5
 Marion model, 15-23
 Markov chain Monte Carlo (MCMC) method, 14-28, 24-20
 Marl, 2-9
 Mass balance equation, 2-16, 7-3, 9-5-9-8, 9-23, 12-4, 12-7-12-9, 18-4, 22-4, 22-6, 22-8
 Massachusetts, 15-3
 Mass-transfer coefficient, 9-12
 Master Equation, 18-5-18-6
 Mathematical model, 23-2
 of sea water intrusion, 12-6-12-11
 Matrix flow, 21-8
 Matrix retardation coefficient, 20-8
 Matrix solution techniques, 23-17
 Maximum comparative percolation ratio, 34-14
 Maximum contaminant level (MCL), 35-3
 Maximum quantitative percolation value, 34-14
 Mechanical dispersion, 3-34, 23-6
 Mediterranean area, 2-9
 Mendota, 3-25
 Mesophilic bacteria, 31-6
 Mesozoic Era, 2-9
 Metabolism, 31-3-31-6
 Metadata, 30-3
 Metamorphic rock, 2-8
 Meteorological variables, evapotranspirative covers, 34-15
 Methane, 2-47
 Methanol, 31-8
 Method-of-characteristics methods (MOC), 23-15-23-17
- Method of fragments, 5-31-5-37
 form factor, 5-34
 Method of moments, 22-24
 Methyl tert-butyl ether (MTBE), 17-24
 Mexico, 2-3
 Mezencev (ME) model, 7-10
 Mg²⁺, 17-7
 MGE (Modular GIS Environment), 30-2
 Michaelis-Menten equation, 31-13
 Microaerophiles, 17-19
 Microbial dynamics, upscaling, 19-11-19-13
 Microbial reaction
 kinetics, 31-12-31-15
 and sorption, 31-18-31-20
 Microbiology, 31-1-31-3
 Microorganisms, 17-19, 17-29
 Microresistivity logs, 14-17
 Middle East, 2-10
 Milrow model, 24-18-24-19, 24-21-24-22
 Mineralization-immobilization turnover (MIT), 29-26
 Minerals and rocks, 2-7-2-8
 Mini-disk infiltrometer, 7-11
 Minimal drawdown approach, *see* Low-stress approach
 Minimalpurge approach, groundwater sampling, 35-29
 Minimum relative entropy (MRE), 14-33
 Minimum safe discharge, 23-41-23-42
 Mining effects, 2-50-2-51
 Minnesota, 2-26, 2-30, 31-21
 Miscible contaminant, 3-29
 Miscible liquids, 12-2
 Mississippi, 15-3-15-4
 Mississippi River, 2-9-2-10, 2-35, 2-51
 Missouri, 3-25
 Missouri River, 2-9-2-10
 Mobile-immobile model, 9-4, 9-6
 time scale, 9-4-9-5
 MOC3D, 30-9
 Model, definition, 23-2, 23-20
 Model calibration, 23-26
 Model validation, 16-10
 Model Viewer, 23-45, 30-9
 Modena, 1-5
 MODFLOW, 28-4, 28-17, 30-9
 MODFLOW-2000, 23-28-23-30, 23-32, 23-35
 Ground-Water Flow Process
 MODPATH, 23-28-23-29, 23-37
 observation process, 23-29
 parameter-estimation process, 23-29
 sensitivity process, 23-29
 MODFLOW-SURFACT, 22-32
 MODHMS, 23-30
 MODPATH, 30-9
 Modules of elasticity, 3-25-3-26
 Moffett Naval Air Station, 31-21
 Moisture content vs depth, 2-28-2-29
 Moisture content, 2-12
 Molecular attraction, 2-14
 Molecular diffusion, 3-28, 23-6
 Molecular oxygen, 31-7
 Molecular techniques, 31-9
 Moments, 5-11-5-17
 Monitored natural attenuation (MNA), 36-38-36-40
 Monitoring wells, 11-23-11-24

- Monoaromatic hydrocarbons, **31-7–31-8**
 Monofilament, **37-11**
 Monolithic covers, **34-9–34-11**
 Monolithic final cover, **33-21**
 Montmorillonite crystal, **37-18**
 Motile organisms, **31-3**
 Motion equation, **4-6–4-9, 4-15–4-19**
 MT3DMS, **23-30**
 MTBE, **22-13**
 Mud rotary, **35-19**
 Multicomponent reactive solute transport
 cation exchange, **22-20**
 complexation, **22-20**
 components and reversible chemical
 reaction, **22-19**
 coupled system of equations, **22-20–22-21**
 precipitation and dissolution, **22-20**
 Multi-dimensional models, **22-25–22-27**
 Multilayered aquifer, **10-2**
 Multi-parameter probes, **35-30**
 Multi-permeability models, **22-6**
 Multiphase flow, **3-29**
 Multiple boreholes, **20-22**
 Multiple population inversion (MPI), **15-17**
 Multiple-packer strings, **20-13**
 Multiple scale roles, **18-2–18-3**
 Multi-pore group models, **9-26–9-28, 22-6**
 Multistep outflow method, **6-21**
 Multi-well hydraulic tests, **20-17–20-18**
 interference tests, **20-17**
 pulse interference tests, **20-17–20-18**
 wellbore storage, **20-17**
 Murot, **1-5**
 Mutually exclusive outcomes, **5-10**
 MYGRT, **22-25**
- N**
- N3DADE, **22-25**
 Nanoscale zero-valent iron (nZVI), **36-40**
 NAPRA (National Agricultural Pesticide Risk Analysis),
 30-5, 30-10–30-11
 NASIS (National Soil Information System), **30-2, 30-6**
 National Bureau for Drinking Water Supply, **1-15, 1-21**
 National Contingency Plan, **32-24**
 National Research Council (NRC), **36-2, 36-8**
 National Well Water Association, **30-10**
 Natural Attenuation Study (NATS), **26-5**
 Natural gradient methods, **15-3–15-7**
 Natural pixels, **15-17**
 Natural protective barriers and waste disposal, **2-44**
 Natural Resources Conservation Service (NRCS), **7-10,**
 28-10
 “Natural softening”, **1-26**
 Navier–Stokes equation, **1-20, 4-8, 21-9**
 Nebraska, **2-52**
 Negligible run-on
 spatial averages, **8-9–8-10**
 Neighborhood function, in GIS spatial analysis, **30-4**
 Neptunium, **17-27**
 Net system discharge, **23-41**
 The Netherlands, **1-9, 1-14, 1-16–1-18, 1-19, 1-21–1-24,**
 2-46, 3-28
 Neutron thermalization, **6-17**
 Nevada, **2-10, 2-38, 2-43, 2-46, 17-28, 19-6**
 New Hampshire, **15-25**
 New Jersey Geological Survey (NJGS), **28-24**
 New Mexico, **2-9–2-10, 2-52, 3-25**
 New York, **2-45**
 Newton’s law of viscosity, **22-9**
 Newton’s stress law, **21-14**
 NHD (National Hydrography Dataset), **30-2**
 Nitrate, **31-5–31-6, 31-21**
 Nitrogen, **31-6**
 cycle, in soil-plant system, **29-26**
 N-nitrosamines
 N-nitrosodimethylamine (NDMA), **17-23–17-24**
 Noahian Flood, **2-8**
 Noble gas tracers, **21-29**
 Nonaqueous phase liquids (NAPLs), **17-20–17-21**
 Nonconservative solutes, **3-29**
 Nonequilibrium transport, **22-13–22-17**
 chemical nonequilibrium, **22-15–22-16**
 colloid-facilitated solute transport, **22-16–22-17**
 physical nonequilibrium, **22-14–22-15**
 see also Unsteady-state flow
 Nonhazardous waste, RCRA regulations, **32-21**
 Nonlinear well losses, **11-6–11-7, 11-13**
 Non-methane organic compounds (NMOCs),
 33-17–33-19
 Nonmicrobial colloidal particles, **17-30**
 Non-point source (NPS), **30-7, 30-9–30-10**
 Nonreactive contaminant transport, in saturated zone,
 18-1
 classical deterministic models, **18-4**
 Continuous Time Random Walk (CTRW), **18-5–18-7**
 dispersion, **18-3**
 Eulerian and Lagrangian perspectives, **18-3–18-4**
 field experiments, **18-14–18-15**
 Fractional Advection-Dispersion (FADE), **18-7–18-8,**
 18-14
 heterogeneity, hierarchical, **18-2**
 lab experiments, **18-14**
 multiple scale roles, **18-2–18-3**
 spreading vs dilution, **18-12–18-14**
 stationary systems, locally, **18-11–18-12**
 statistical mechanical approach, **18-4–18-5**
 stochastic perturbation
 Fourier transform method and Green’s function
 approaches, **18-8–18-10**
 Lagrangian perturbation approach, **18-10–18-11**
 Noria, **1-3**
 Normal movement correction, **15-9, 15-20**
 Normality, **16-10**
 North Africa, **1-3**
 North America, **2-3, 20-10, 31-9**
 groundwater regions, **2-5**
 Norway, **2-49**
 NO_x, **17-15–17-16**
 Nuclear logs, **14-19**
 Nuclear Waste Policy Act of 1982, **17-28**
 Numerical errors, **23-23**
 Numerical model, **23-2**
 evapotranspirative covers, **34-16**
 of sea water intrusion, **12-11–12-15**
 flow model, **12-12–12-13**
 specific discharge, calculation, **12-14–12-15**
 transport model, **12-13–12-14**

O

Oahu island, 23-39
 Observation wells, 11-24
 Octanol : water partitioning coefficient (K_{ow}), 17-20
 Ogallala Aquifer, 2-52
 Ogallala formation, 2-4
 Ohio River, 2-9
 Ohm's law, 3-9
 OII landfill, evapotranspirative cover system, 34-19-34-23
 Oklahoma, 2-52, 3-25
 One-dimensional hydraulics, 3-12-3-13
 One-dimensional models, 22-24-22-25
 Ontario, 15-3, 20-21
 Open-hole slug tests, 20-16
 Optical brighteners, 21-29
 Optimization problem, 23-29
 Ordinary kriging, 16-3-16-6
 conditional realization, 16-5-16-6
 function estimate, 16-4-16-5
 Oregon, 2-38
 Organic contaminants
 background organics, 17-19-17-20
 fuel oxygenates, 17-24
 high-profile organic contaminants, 17-21-17-24
 nonaqueous phase liquids (NAPLs), 17-20-17-21
 Organotrophs, 31-2
 Orthonormal, 16-9
 Osmotic potential, 6-6
 Outwash
 glacial, 2-8
 sand, 2-8-2-9
 Overdraft, 2-50, 2-52
 Overland flow, 2-31
 Overlay function, in GIS spatial analysis, 30-4
 Overpumping method, 11-27
 Oxbow lakes, 2-38
 Oxygen, 1-27-1-28, 2-51, 17-15, 17-19
 Oxygenase enzymes, 31-5, 31-7
 Oyster, Virginia site, 15-20

P

P_e , 17-11-17-13
 P_e -pC diagram, 17-12
 P_e -pH diagram, 17-12-17-13
 Packers, 20-13
 Paleohydrology, 21-16-21-17
 Paleozoic carbonate aquifer, 21-5
 Paradise Valley, 27-8-27-10
 Parameter-estimation procedures, 23-22
 Paris, 1-5
 Park County, Colorado, 2-43
 Parlange et al. model (PA), 7-8-7-9
 Partial-area, 1-25
 Particle pathlines, 23-16
 Particle-size distribution, 6-11-6-13
 bi-modal behavior, 6-12
 particle-size scale parameter, 6-13
 particle-size shape parameter, 6-12
 Particle size limits, 6-3
 Particle strength exchange method, 23-16
 Particle tracking, 18-5, 20-35-20-36
 Particle Tracking Velocimetry (PTV), 18-14
 Particulates, 17-29-17-30

Passive capillary sampler (PCAPS), 28-12
 Passive sampling, groundwater sampling, 35-29-35-30
 Passive tracer, 19-2
 PC-pH diagram, 17-6-17-8
 PCBs, *see* Polychlorinated biphenyls
 PCE, *see* Perchloroethylene
 Pearl Harbor, 23-39, 23-41-23-42
 Peat, 2-9
 Peclét number (Pe^c), 22-31-22-32
 Pedo-Transfer Functions (PTF), 6-11-6-12, 6-24, 6-31
 Pellicular water, *see* Capillarity
 Pentachlorophenol (PCP), 31-10
 Perched aquifer, 2-26-2-27
 Perchlorate, 31-4
 Perchloroethylene (PCE), 17-21, 31-8, 31-10
 Percolation, 1-5, 28-6, 28-10
 Perennial streams, 2-40
 Performance monitoring programs, groundwater
 monitoring, 35-4
 remediation system, 36-43-36-44
 Permafrost, 13-9-13-10
 Permeability, 1-8, 2-39, 6-12
 coefficient of, 5-5
 conduit, 21-9, 21-17, 21-27
 dual-permeability model, 22-5-22-7, 22-14
 field coefficient of, 2-20
 fracture, 21-17, 21-31
 intrinsic, 1-20, 2-18, 3-8
 karstic aquifers, 20-9
 matrix, 21-31
 multi-permeability models, 22-6
 of natural soil, 33-10
 Permeameter, 3-10-3-11, 7-11, 15-5
 Permian basin, 3-25
 Permittivity, 37-7
 Petroleum
 hydrocarbon, 17-21, 31-11, 35-4
 aerobic bioremediation, 31-21
 Petrophysics, 14-11, 15-23
 pF curve, 1-14
 pH and chemical composition, 17-3-17-10
 Phenols, 31-10
 Philip infiltration equation (PH), 6-37, 7-7-7-8
 Phosphorous, 31-6
 Photoautotrophs, 31-2
 Photoheterotrophs, 31-2
 Photometry, 35-36
 Photosynthesis and transpiration, atmospheric controls, 29-2
 ecosystem fluxes modeling, 29-10
 flow field inside canopies, 29-9-29-10
 leaf energy budget, 29-8-29-9
 radiative transfer, 29-8
 scalar mass balance and turbulent transport, 29-3-29-5
 scalar transfer from leaf stomata and resistance
 parameterization, 29-5-29-6
 stomata sub-cavity space concentration
 parameterization, 29-6-29-8
 transpiration and hydraulic controls, 29-10-29-13
 Phreatic aquifer, 1-17, 4-15, 4-18
 flow equation, 4-20-4-24
 recharge, 1-21
 PHREEQC, 22-35

- PHREEQE, 1-27
- Physical barriers, for hydraulic containment, 36-14-36-17
- cutoff walls, 36-14
 - soil-bentonite slurry walls, 36-14-36-17
- Physical heterogeneity, 19-4
- Physical nonequilibrium
- dual-permeability model, 22-14
 - dual-porosity and mobile-immobile water models, 22-14
 - mass transfer, 22-14-22-15
- Physical properties, of soil, 6-4-6-5
- and hydraulic properties, functional relationships
 - hydraulic conductivity functions, 6-15-6-16
 - particle-size distribution, 6-11-6-13
 - water retention curves, 6-13-6-15
- Physical techniques, soil water retention curve, 34-6
- Physisorption, 17-10
- Picard method, 12-11
- Piecewise approximation, 23-13-23-14
- Piezometer, 1-9, 2-50, 3-3, 3-12, 4-18, 35-13, 35-22-35-23
- Piezometric head (potential), 3-3, 3-5, 3-7, 3-20
- Piezometric surface, 2-25, 2-27
- Pipe flow models, for karst aquifer, 21-38
- Piping, 5-43-5-46
- Plant roots, in recharge process, 28-9
- Plant water stress
- modeling, 29-23
 - soil-climate-vegetation control and plant carbon assimilation, 29-23-29-24
- Plastic limit, of soil, 2-52
- Platte River, 2-52
- Pleistocene, 2-10
- Pleistocene Epoch, 2-8, 2-38, 2-46
- Pleistocene fluvial sand, 1-15
- Pneumatic hydraulic fracturing, 36-28-36-29
- Point dilution method, 20-18-20-22
- Point estimate method (PEM)
- one random variable, 5-25-5-27
 - regression and correlation, 5-27-5-30
 - several random variable, 5-30-5-31
- Point estimation, with known parameters
- Bayesian point estimators, 14-39-14-42
 - normal linear methods, 14-40-14-42
- Point simulation, 16-5-16-6
- Point-water level, 1-19
- Poiseuille's law, 22-9
- Polychlorinated biphenyls (PCBs), 31-10
- Polyethylene, 37-12
- Polyethylene diffusive bag (PDBs), 35-29-35-30
- Polyhalogenated compounds, 31-8
- Polymer, 37-11
- Polynomial regression interpolation, 30-8
- Polyphosphates, 11-27-11-28
- Polyvinyl chloride (PVC), 37-12-37-14
- Ponding, 1-24, 2-38, 6-33, 7-4, 8-1-8-2
- time of, 7-4, 7-10-7-12, 7-14, 8-9
- Pore
- pressure, 3-3
 - scale mixing, 19-17
 - size distribution, 7-3, 29-33, 34-6
 - space, 5-1
 - velocity, 3-8
- Porosity, 2-8, 2-10-2-12, 3-2, 15-23
- double, 1-18
 - dual-porosity models, 22-5-22-7
 - with distribution layer at soil surface, 9-22-9-26
 - for gradual change at fracture-matrix interface, 9-9-9-23
 - for sharp change at fracture-matrix interface, 9-3-9-9
 - effective, 2-10, 3-8
 - inter-aggregated, 9-3
 - primary, 2-10
 - secondary, 2-10
 - single-porosity model, 22-5
- Porous-cup tensiometer, 1-14
- Positrons, 17-25
- Potential boundary condition, 4-13
- Potential flow problem, 4-12
- Potential theory, 5-6
- Potentiometric surface, 2-27, 35-9
- Prandl-Von Karman equation, 21-10
- Precipitation, 2-48, 2-51, 36-43
- and dissolution, 22-20
- Predominance area diagram, 17-8
- Preferential flow
- dual-permeability model, 22-5-22-7
 - dual-porosity models, 22-5-22-7
 - mass transfer, 22-7
- Prescriptive rights, 32-6
- Pressure energy, 3-7
- Pressure head, 3-3, 3-7, 3-8
- Pressure plate test, 34-6
- Pressure ring infiltrometer, 6-21
- Pressure surface, *see* Piezometric surface
- Primary porosity, 2-10
- Prior appropriation, 32-6
- Probabilistic fundamentals, 5-9
- moments, 5-11-5-17
 - expected value, 5-12-5-13
 - probability distribution, 5-17-5-21
 - space, 5-18
- Production casing, 11-15
- Project Shoal Area (PSA) model, 24-6-24-18
- acceptance criteria, 24-7-24-8
 - Corrective Action Investigation Plan (CAIP), 24-6
 - multiple validation targets, 24-12-24-18
 - single validation target, 24-8-24-12
- Property taking issues, 32-7
- Protection function, geosynthetic, 37-10
- Protozoa, 31-2
- Protozoan cysts, 17-30
- Pseudo-steady state, 10-6
- Psychrometer, 6-18
- Psychrophiles, 31-6
- Pulse interference tests, 20-17-20-18
- Pumice, 2-7, 2-10
- Pump
- and motor, 36-23, 36-26
 - in well design, 11-20-11-22
- Pump housing
- length calculation, 11-15-11-17
- Pumped aquifer
- confined and unconfined aquifer, differences, 10-6
 - leaky aquifers, 10-7-10-9
- Pumping, 2-27, 23-40

Pumping tests, 10-1
duration, 10-15
Pyrite oxidation, 17-18
Pyroclastic rock, 2-8

Q

Q1 statistic, 16-10
Q2 statistic, 16-10
Qanat system, 1-2-1-3
Quantitative hydrology, 1-4
Quasi-exact infiltration solution, 6-38-6-39
Quasilinear approach, 16-16
Quicksand, 3-27, 5-43

R

Radar, 6-18
signals, 15-14
Radial-convergent experiment, 20-22
Radial-divergent experiment, 20-22
Radial flow, 1-9
Radiative transfer, 29-8
Radiometers, 6-18
Radionuclides
radioactive decay, 17-24-17-25
radioactive waste disposal, 17-28-17-29
remediation, 17-29
sources, 17-26-17-28
units, 17-26
Radium, 17-26-17-27
“Radius of influence”, 1-9
Radon, 17-26
Random field, 19-4-19-5
generation, 8-4-8-5
Random nonequilibrium sorption, 19-8-19-9
Random part, 16-2
Raster data, 30-3
Rate-limited model, 9-4-9-5
Raw variogram, 16-8
Raymonds number, 21-10
Reaction system examples, 19-18-19-20
Reactive contaminant transport, in saturated zone, 19-2
connection, between approaches, 19-22-19-23
heterogeneity, 19-4-19-5
Lagrangian stochastic models, for nonlinear
reactions, 19-13-19-21
heterogeneity, in reactive properties, 19-20-19-21
reaction system examples, 19-18-19-20
stochastic-convective averaging, 19-16-19-17
streamtube formulation, 19-14-19-16
travel-time distribution function, 19-17-19-18
linearity and nonequilibrium Eulerian stochastic
models
deterministic nonequilibrium sorption, 19-7-19-8
random nonequilibrium sorption, 19-8-19-9
microbial dynamics, upscaling, 19-11-19-13
numerical results, 19-9-19-11
scale of observation, 19-5
upscaling
chemical heterogeneity methods, 19-5-19-7
Reactive transport and biodegradation, quantitative
description
microbial reaction kinetics, 31-12-31-15
microbial reactions and sorption, 31-18-31-20

sorption, as reactive kinetic process, 31-15-31-18

Reactor ratio, 18-12-18-13

Reasonable doctrine, 32-5

Recharge, 28-2, 28-5-28-7

affecting factors

infiltration and flow, in unsaturated zone, 28-10

land use and cover, 28-8

overland flow, 28-10

precipitation, 28-7-28-8

soil properties, 28-10

temperature, 28-11

urbanization, 28-9

vegetation, 28-8-28-9

estimation methods, 28-11-28-14

combined modeling, 28-14

Darcian approaches, 28-13

empirical methods, 28-12-28-13

field techniques, 28-12

inverse groundwater modeling, 28-14

tracers, 28-13

water budget methods, 28-13

in groundwater modeling, 28-15-28-20

coupled models, 28-16-28-18

integrated models, 28-18-28-20

Recharge and infiltration capacity, 2-30-2-31

Recharge basins, 3-17-3-18

Recharge estimation, to water table, 30-6

Recharge spreading layer (RSL), 28-15

Redox potential, 31-6

Redox reaction, 17-12

Redox transformation, 31-3-31-4

REELOGGER™, 13-12

Regional Aquifer System Analysis (RASA), 2-4

Regional groundwater systems

basic-scale analysis and evaluation, 1-20-1-21

flow systems, topography driven, 1-22

land drainage and groundwater discharge, 1-22-1-24

runoff generation, on sloping areas, 1-24-1-25

Regression approach, 5-27-5-30, 15-2, 23-22

Reinforcement function, geosynthetic, 37-4-37-7

Relative roughness, of fracture, 20-6

Remediation goals, 36-2-36-4

advantages and disadvantages, 36-3

Remediation system design, of contaminated

groundwater

candidate remediation, screening, 36-8-36-12

design effectiveness, confirmation, 36-13

design implementation, 36-13

detailed design, preparation, 36-12-36-13

problem, defining, 36-6-36-7

remediation goals, defining, 36-7-36-8

Remediation wells, 11-22-11-23

Remote sensing, 6-18

radars, 6-18

radiometers, 6-18

Replicates, QC check, 35-31

Representative elementary volume (REV), 3-2, 3-34,

6-5, 18-2, 19-6

Residence time, 2-32, 2-35-2-36

Residual drawdown, 10-1

Resource Conservation and Recovery Act (RCRA), 3-39,

32-16-32-22, 36-2

legislative background

general provisions, 32-16-32-17

- goals and objectives, 32-16
- historical perspective, 32-16
- issues and outlook, 32-17
- regulations
 - hazardous waste, 32-17–32-21
 - nonhazardous waste, 32-21
 - underground storage tanks, 32-21–32-22
- Respiring organisms, 31-3
- Response Authorities, 32-24
- Response-matrix method, 23-29
- Restatement doctrine, 32-5
- Reston, Virginia, 2-4
- Retardation factor, 3-38
- Retrieval function, in GIS spatial analysis, 30-4
- Reverse osmosis (RO), 36-43
- Reversible chemical reaction and components, 22-19
- Reynolds number, 2-33, 3-10, 3-13, 5-5, 20-6
- Rhode Island, 23-35
- Rhombohedral packing, 2-11
- Richard's equation, 1-14, 1-18, 20-34, 22-3–22-5, 34-16
- Ring infiltrometers, 7-10–7-11
- Riser pipe, 35-21
- Risk, of contaminated groundwater, 36-4–36-6
- Rock, 2-7–2-8
- Rock flour, 2-10
- Rock physics, *see* Petrophysics
- Rocky Mountain Arsenal, evapotranspirative cover system, 34-23–34-27
- ROMIN, 30-5
- Rorabaugh's method
 - in step-drawdown tests, 11-11–11-13
- Rose diagrams, 20-10
- Rotary drilling technique, 2-7
- Runoff generation, on sloping areas, 1-24–1-25
- Run-on process, 8-2
 - on Hortonian overland flow, 8-2
 - importance, 8-8–8-9
- numerical simulation results
 - discussion, 8-7–8-8
 - expected spatial average infiltration rates, 8-7
 - local scale model, 8-5–8-6
 - run-on importance, 8-8–8-9
 - soil properties and selected cases, 8-6
- S**
- Safe Drinking Water Act (SDWA), 32-10–32-16
 - legislative background
 - general provisions, 32-10–32-13
 - goals and objectives, 32-10
 - historical perspective, 32-10
 - issues and outlook, 32-14
 - regulations, 32-14–32-15
 - Safe yield, 1-21, 2-42–2-43
 - Sahara, 2-10
 - Salinization, 1-18
 - Salt licks, 2-51
 - Salt transport model, 12-13–12-14
 - Salt water interfaces, 3-28–3-29
 - Salt water intrusion, 3-28
 - Salt Water Intrusion Meetings (SWIM), 1-19
 - Sample, in groundwater sampling
 - analysis
 - field chemical analyses, 35-34
 - laboratory chemical analyses, 35-34–35-37
 - collection, 35-31–35-32
 - filtration, 35-32
 - preservation, 35-32–35-33
 - storage, 35-33
- San Francisco Bay, 2-49
- San Joaquin Valley, 2-50, 3-25, 15-17, 15-19
- Sand pumping, 11-29
- Sand trap, 11-20
- Saturated media, hydraulic conductivity, 2-20–2-22
- Saturation indices, 21-21
- Saturation ratio, 2-12
- Scaling, 25-1
 - background, 25-4–25-5
 - catchment scale, spatial variability characterization, 25-10–25-12
 - extrapolation, 25-12
 - literature, 25-5
 - measurement and interpretation, 25-5–25-7
 - of soil water flow, 6-40–6-48
 - scaling evaporation, 6-47–6-48
 - scaling infiltration, 6-41–6-47
 - triplet, 25-7, 27-8
 - upscaling and downscaling, 25-7–25-10, 25-12–25-14
 - downscaling, 25-10
 - steps, 25-8
 - upscaling, 25-9–25-10
- Scarborough, 1-5
- Schofield, 1-14
- Schoolcraft Plume-A Site, 15-5
- Screen section, 11-17–11-18
 - length, 11-17
 - slot size, 11-17
 - well screens, calculation example, 11-18–11-19
- Sea water intrusion, into coastal aquifers, 12-1, 23-39–23-44
 - coastal aquifer management, 12-23–12-26
 - examples
 - in double-layered aquifer, 12-23
 - typical case, 12-15–12-18
 - with upconing, due to pumping, 12-18–12-23
 - mathematical modeling, 12-6–12-11
 - numerical modeling and computer code
 - FEAS code, 12-15
 - numerical model, 12-11–12-15
- Seal formation, 7-12–7-14
 - depositional or sedimentary seals, 7-13
 - structural seals, 7-12–7-13
- Search technique, 30-4
- SEAWAT-2000, 23-30
- Secondary porosity, 2-10
- Sedimentary rock, 2-8
- Sedimentation/detention basins, 33-10
- SEDSPEC, 30-5
- SEEPAGE, 30-10–30-11
- Seepage, from open channels, 3-16–3-17
- Seepage face, 2-32–2-33
- Seepage force, 5-43, 5-45
- Seepage velocity, 3-8, 23-5
- Seeps, 35-13
- Seismic impact zones, 33-4
- Seismic methods, for hydrogeological characterization, 14-11–14-15
 - petrophysics, 14-15
 - reflection methods, 14-12–14-14

- Seismic methods, for hydrogeological characterization
(Continued)
- refraction methods, 14-14
 - tomography, 14-14–14-15
- Semi-confined aquifer, 2-27–2-28
- Semophysical methods
- Arya-Paris model, 6-25–6-28
 - Haverkamp-Parlange model, 6-29
- Sensitivity tests, 23-25–23-26
- Separation function, geosynthetic, 37-3–37-4
- Separation streamline, 12-16
- Sequential Gaussian co-simulation, 15-19
- Sequential imaging, of saline plume, 15-25
- Shaduf, 1-2
- Shallow unconfined aquifer, flow and transport, 23-31–23-35
- Sharp interface approximation, 12-3
- SHAW, 22-32
- Shear fractures, 20-3
- Shear zones, 20-3
- Shearing stress, 2-19
- Sheeting structure, 20-3
- Silica sand, 17-9
- Silt, 2-8–2-10
- Silurian dolostone, 20-21
- Silver springs, 2-38
- Simulation experiments, 15-7
- Simulation/optimization methods, for aquifer property estimation, 15-2–15-7
- tracer tests
 - forced gradient methods, 15-7
 - natural gradient methods, 15-3–15-7
- Simultaneous iterative reconstruction technique, 15-17
- Single fracture, 20-4–20-7, 20-26–20-28
- Single point resistance, 14-17
- Single-porosity model, 22-5
- Single-species solute transport models, 22-32–22-36
- HYDRUS software packages, 22-33–22-34
 - MODFLOW-SURFACT, 22-34–22-35
- Single-well hydraulic tests, 10-24–10-26, 20-13–20-16
- constant-head injection tests (CHIT), 20-13–20-16
 - slug tests, 20-16
- Sinkhole, 3-25, 21-39–21-40
- development
 - hydrologic factors, 21-42
 - mechanisms of, 21-41–21-42
 - lakes, 2-38
 - as land use hazards, 21-41
 - remediation, 21-42–21-43
 - spring, 2-38
- Site Master Plan, 33-5–33-6
- Site remediation, 3-39
- Site-specific model, 23-2
- Skewness
- coefficient of skewness, 1-12, 5-17
- Skibitski, Herb, 1-22, 27-6
- Sliding-head packers, 20-13
- Slit-film filament, 37-11
- Slowness, 15-17
- Sludge, 2-44
- Slug test, 3-12, 20-16
- Slurry-phase treatment, 36-41
- Slurry wall, for contamination migration prevention, 2-45
- Snowpack melting, 28-11
- Sodium chloride, 17-19
- Soft data, 16-8
- Soil and rock, physical properties, 2-12
- Soil-bentonite slurry walls, 36-14–36-17
- Soil characteristics
- estimation techniques, 6-22–6-30
 - hydraulic conductivity relation, 6-29–6-31
 - water retention relation, 6-23–6-29
- measurement, 6-16–6-22
- hydraulic conductivity, 6-19–6-21
 - soil water content, 6-17–6-18
 - soil water pressure head, 6-18–6-19
 - water retention and hydraulic conductivity, combined, 6-21–6-22
- spatial variability, of soil water properties, 6-31
- Soil Conservation Service (SCS) method, 7-10, 28-10
- Soil heterogeneity, 4-20
- Soil infiltration capacity, 7-4–7-5
- Soil liquefaction, 2-49
- Soil moisture
- characteristic curves, 2-29
 - deficit, 29-19–29-21
 - physiological impact, on plants, 29-21–29-23
 - dynamics, 29-24–29-34
 - model structure, 29-26–29-29
 - soil carbon and nitrogen cycles, 29-25–29-26
 - soil organic matter and soil hydraulic properties, 29-33–29-34
 - system behavior, 29-29–29-33
- Soil phases
- gaseous phase, 6-4
 - liquid phase, 6-4
 - solid phase, 6-3–6-4
- Soil physics, 1-24
- Soil properties
- functional relationships, 6-11–6-16
 - hydraulic properties, 6-5–6-10
 - physical properties, 6-4–6-5
 - soil phases, 6-3–6-4
- Soil surface
- dual-porosity models
 - with distribution layer, 9-22–9-26
- Soil taxonomy, 25-12
- Soil tests, 2-52
- Soil vapor extraction (SVE), 36-34–36-35
- Soil water balance, in vadose zone, 29-13–29-19
- heat flow into soil, 29-19
 - long-term water balance, 29-17–29-19
 - soil-moisture dynamics
 - at daily time scale, 29-14–29-17
 - instantaneous equation, 29-13–29-14
- Soil water diffusivity, 6-32
- Soil water physics, 1-13–1-14
- Soil water retention
- curve, 34-5–34-7
 - function, 7-2–7-3
- Solar radiation, 13-6–13-7
- Solid earth materials compressibility, 2-15
- Solid stem auger, 35-19
- Solinist® system, 20-23
- Solute transport
- advection-dispersion equation, 22-11–22-13
 - aquifer heterogeneity accounting
 - MADE site in Columbus, case study, 26-1–26-14

- in discrete fracture network, 20-35–20-38
- equation, 23-5–23-7
- initial and boundary conditions, 22-21–22-22
- multicomponent reactive solute transport, 22-19–22-21
- multiphase flow and transport, 22-21
- nonequilibrium transport, 22-13–22-17
- in single fracture, 20-7–20-9
- stochastic models, 22-17–22-18
- transport processes, 22-8–22-10
- Sorption, 36-27
 - as reactive kinetic process
 - thermodynamics, 31-15–31-17
 - time dependent sorption, 31-17–31-18
 - to solids, 17-10–17-11
- Sorptivity, 6-34–6-35, 6-37, 7-6
- South Dakota, 2-9, 2-38, 2-52
- SO_x, 17-15–17-16
- Spatial moments, 15-3
- Spatial variability, 25-3
 - catchment scale, 25-10–25-12
 - characterization
 - rainfall, 8-4
 - random fields, generation, 8-4–8-5
 - saturated conductivity, 8-3–8-4
 - considerations
 - of infiltration, 7-11–7-12
 - infiltration and run-on
 - numerical simulation results, 8-5–8-9
 - of rainfall, 8-4
 - random fields, generation, 8-4–8-5
 - of saturated conductivity, 8-3–8-4
 - theoretical results, 8-9–8-12
 - of soil water properties, 6-31
- Specific discharge, 2-34, 2-37, 3-8
 - calculation, 12-14–12-15
- Specific heat of water, 13-3–13-4
- Specific retention, 2-17
- Specific storage, 2-17, 4-12
 - coefficient, 3-23
- Specific surface, 1-8
- Specific yield, 1-17, 2-17, 2-27, 3-23
- Spectroscopy, 35-36
- Spent nuclear fuel (SNF), 17-28
- Spikes, QC check, 35-30
- Spline interpolation, 30-8–30-9
- Split inversion method (SIM), 15-21–15-23
- Spontaneous potential (SP) logs, 14-17–14-18
- Spreading vs dilution, 18-12–18-14
- Spring, 2-37–2-38
- Sprinkler method, 7-10
- SSURGO (Soil Survey Geographic), 30-2, 30-6
- Stagnation point, 2-39
- Standard deviation, 5-14
- Standard pump test, 21-31
- STANMOD, 22-25–22-26
- Staple filament, 37-11
- State groundwater rights systems
 - absolute dominion, 32-4–32-5
 - conjunctive management, 32-6
 - corrective rights, 32-5–32-6
 - prescriptive rights, 32-6
 - prior appropriation, 32-6
 - reasonable use, 32-5
- torts restatement, 32-5
- Stationary systems, locally, 18-11–18-12
- Statistical mechanical approach, 18-4–18-5
- Steady flow
 - in confined and unconfined aquifers, 3-18–3-20
 - over horizontal aquiclude, 3-15–3-16
- Steady-state centrifuge (SSC), 34-9
- Steady-state flow, 10-15
- Steady-state laminar flow, 2-34
- Steamboat Springs, Colorado, 2-38
- Steam extraction, 36-29
- Step-drawdown tests, 11-7–11-13
 - Jacob's method, 11-8–11-11
 - Rorabaugh's method, 11-11–11-13
- Stochastic approach, 16-2, 23-3
 - heterogeneity, 22-18
 - stochastic numerical models
 - groundwater model validation, 24-3–24-6
 - stream tube models, 22-17–22-18
 - transfer function models, 22-17
 - Stochastic-conductive averaging, 19-16–19-17
 - Stochastic-conductive reaction (SCR) method, 19-14
- Stochastic perturbation
 - Fourier transform method and Green's function approaches, 18-8–18-10
 - Lagrangian perturbation approach, 18-10–18-11
- Stomatal closure, 29-22
- Storage coefficient, 1-16
- Storativity, 3-23, 4-20
- Storm runoff, 1-24
- Storm water management system, 33-16, 33-26–33-27
- Stratigraphy, 20-28
- Stream beds, 2-9
- Stream flow equation, 1-5
- Stream flow hydrograph, 2-36
- Stream lines, 3-20–3-21
- Stream tube models, 22-17–22-18
- Streamflow-routing (STR) package, 23-35–23-37
- Streamlines, 5-6–5-9
- Streamtube formulation, 19-14–19-16
- Stress-balance diagram, 2-15
- Strong acids, 17-17
- Strontium, 17-28
- Structural geology, of fractured rock, 20-2–20-4
- Structural hierarchy, 18-2
- Structural springs, 2-38
- Student *t*-test, 24-10
- Subbase grading, landfill, 33-9
- Submerged unit weight, 5-43
- Subsidence, *see* Land subsidence
- Substrate level phosphorylation (SLP), 31-3
- Subsurface drainage patterns, karst aquifer, 21-5–21-6
- Subsurface hydrology, 1-2
 - aquifer exploration, 1-14–1-16
 - flow analysis and aquifer tests, 1-16–1-18
 - fresh-salt water interface problems, 1-18–1-19
 - geological aspects, 1-12–1-13
 - groundwater dynamics, 1-6–1-12
 - hydrogeology consolidation, 1-19–1-20
 - soil water physics, 1-13–1-14
- Subsurface microbiology and geochemistry
 - metabolism, 31-3–31-6
 - microbiology, 31-1–31-3
- Subsurface stormflow, 2-32

- Subtitle D, of Resource Conservation and Recovery Act, 33-2, 33-4
 Suction, 3-3, 3-6
 Sulfide ion, 17-5
 Sumerians, 1-2
 Surface charge, 17-9–17-10
 Surface fitting and interpolation, 30-8–30-9
 Surface mining, 2-50
 Surface Mining Control and Reclamation Act (SMCRA), 32-29
 Surface tension, 2-14
 Surface water and groundwater, heat exchanges, 13-10
 Surface-saturated zones, 1-25
 Surfactant, 17-3
 Surfactant enhanced aquifer remediation (SEAR), 36-40–36-41
 Surfactant flushing, 36-29
 Surficial aquifer, 23-32
 Sustainable development, 27-1–27-2
 aquifer dynamics and models, 27-5–27-6
 basin and range aquifer, dynamics of, 27-6–27-8
 groundwater development, quantitative maintaining, 27-2
 hydrogeology, analytical methods, 27-2–27-3
 Paradise Valley, 27-8–27-10
 water budget, 27-3–27-5
 SUTRA, 23-30–23-31, 23-39, 30-9
 Swallets, 21-2
 SWAP, 22-32
 Swartzendruber (SW) model, 7-9
 Synthetic organic compounds, 17-20
- T**
 Tailwater, 5-7, 5-9, 5-22–5-23, 5-37
 Tarpon Springs, Florida, 2-38
 Taylor series expansion, 23-11
 Tectonic forces, 2-38
 Temperature, 13-2–13-3
 Tensiometer, 3-6, 6-18, 34-18
 Tension disk infiltrometer, 6-20–6-21
 Tertiary Ogallala Formation, 2-52
 Texas, 2-38, 2-52, 3-25
 Theis equation, 3-9, 4-28–4-30, 10-5
 Thermal and mechanical energy, relationship, 13-5
 Thermal conductivity, 13-3
 Thermal energy, 17-13
 Thermal gradients, 13-5
 Thermal inertia, 13-4
 Thermal resistivity, 13-3
 Thermodynamics
 heat
 as energy, 13-2
 heat units, 13-3
 soil water retention curve, 34-6
 specific heat of water, 13-3–13-4
 temperature, 13-2–13-3
 thermal and mechanical energy, relationship, 13-5
 thermal conductivity, 13-3
 thermal inertia, 13-4
 thermal resistivity, 13-3
 3D hexagon structures, of water, 13-4
 viscosity and heat, 13-4–13-5
 Thermophilic bacteria, 31-6
 Thiem–Dupuit equation, 10-6
- Thiem formula, 1-16
 Thiessen polygon technique, 30-4–30-5
 Thin plate interpolation, *see* Spline interpolation
Thiobacillus ferrooxidans, 17-19
Thiobacillus thiooxidans, 17-19
 Thixotropic system, 2-49
 3D hexagon structures, of water, 13-4
 3D modeling, of aquifer, 23-42
 3DADE, 22-25–22-26
 Tigris, 1-2
 Time domain reflectometry (TDR) methods, 6-17–6-18, 34-17–34-18
 Time–drawdown analyses
 aquifer test
 in confined aquifer, 10-17–10-19
 in leaky aquifer, 10-22–10-24
 in unconfined aquifer, 10-19–10-21
 single-well tests, 10-24–10-26
 Time-lapse imaging, 14-29
 Tippecanoe County, 2-44, 30-6
 Titrimetric procedures, 17-9
 Top model, 1-25
 Topography, 35-9
 Tortuosity factor, 6-27
 Total dissolved solids (TDS), 2-48, 12-7
 Total head, 3-3
 Total pressure, 3-6
 Total soil water potential, 6-6
 TOUGH codes, 22-33
 TOUGH2, 22-32
 TOUGH-REACT, 22-35
 Toxic environmental conditions, 31-11
 Tracer studies, 3-11, 20-22–20-23, 21-28–21-30
 forced gradient methods, 15-7
 natural gradient methods, 15-3–15-7
 Tracers, 3-29
 Transfer function models, 22-17
 Transition layers, 9-18–9-22
 Transmission piping, 36-25
 Transmissivity, 2-21, 4-17–4-18, 23-4, 23-21, 37-9
 Transnuranic waste (TRU), 17-28
 Transport equations, 22-11
 Transport mechanisms, of dissolved contaminants
 advection, 3-32
 diffusion, 3-32–3-33
 dispersion, 3-34–3-37
 radioactivity decay and degradation, 3-38–3-39
 sorption, 3-37–3-38
 Transport processes, 22-8–22-10
 advection, 22-10
 diffusion, 22-9
 Travel-time distribution function, 19-17–19-18
 Travel-time of the solute, 19-14
 Travel time tomography, 15-16–15-18
 Treatment techniques, of contaminated groundwater, 36-30
 ex situ treatment
 biological processes, 36-41
 chemical and physical processes, 36-42–36-43
 volatilization processes, 36-41–36-42
 in situ treatment, 36-31–36-41
 biological remediation, 36-32–36-34
 chemical and physical processes, 36-38
 innovative treatment approaches, 36-40–36-41

- monitored natural attenuation, 36-38–36-40
volatilization processes, 36-34–36-38
- Trench, 36-26
- Trench fill, 33-6
- Trichloroethane, 31-8, 31-10
- Trichloroethylene, 17-21
- Triple permeability model components, for karst aquifer, 21-8
- Tritium, 1-28, 2-30, 17-27, 23-32–23-34
- Tritium tracer test, MADE site
hydraulic and transport properties assignment, 26-8–26-9
model setup and boundary conditions, 26-8
objectives and approaches, 26-6–26-8
- Truncation error, 23-11
- Turbulent diffusion, 3-28
- Turbulent flow, 2-34–2-35
- Turgor, 29-22
- Turning bands method, 8-4
- Two- and three-dimensional flow, of groundwater
continuity equation, 4-2–4-5
Darcy's law extensions, 4-9–4-10
Dupuit discharge formula, 4-24–4-27
groundwater flow equations, 4-10–4-15
boundary conditions, 4-13–4-15
Hantush leaky well flow equation, 4-31–4-33
Hvorslev slug test equation, 4-27–4-28
hydraulic approach, to groundwater flow
aquitard, flow equation, 4-24
confined aquifer, flow equation, 4-19–4-20
motion equation, 4-15–4-19
phreatic aquifer, flow equation, 4-20–4-24
macroscopic approach, 4-5–4-6
motion equation, 4-6–4-9
Theis well flow equation, 4-28–4-30
- 2-box model, 31-15
- Two-packer system, 20-13
- U**
- UCODE, 23-22
- Uncertainty reduction, using validation data, 24-18–24-22
background, 24-19–24-20
Bayesian framework, 24-20–24-21
Milrow site, in Amchitka model, 24-21–24-22
- Unconfined aquifer, 2-26, 3-23, 3-25, 10-2, 10-3
distance–drawdown analyses, 10-27–10-28
time–drawdown analyses, 10-19–10-21
well-flow equations, 10-6–10-7
- Unconfined compressive strength, of soil, 2-52
- Unconsolidated materials, 2-8–2-10
- Underground storage tanks, 32-21–32-22
groundwater protection, 32-22
- UNESCO, 1-27
- UNESCO-IHP, 1-28
- Uniformity coefficient, 2-11
- UNIX operating system, 30-2
- UNSATCHEM, 22-35
- UNSAT-H, 22-32, 34-16
- Unsaturated flow apparatus (UFA), 34-9
- Unsaturated flow equation, analytical solutions
constant negative pressure head condition
with gravity effects, 6-36–6-38
without gravity effects, 6-34–6-36
- constant positive pressure head condition, 6-38–6-40
- Unsaturated soil, evapotranspirative cover performance
hydraulic properties
conductivity functions, 34-7–34-9
soil water retention curve, 34-5–34-7
water flow, through unsaturated porous media, 34-3–34-5
- Unsaturated soil water flow, conceptual aspects, 6-31–6-34
general flow equations, 6-32–6-33
infiltration, 6-33–6-34
- Unsaturated zone, soil properties and moisture
movement, 6-1
conclusions and future research perspectives, 6-48–6-49
- flow equation, analytical solutions
constant negative pressure head condition, 6-34–6-38
constant positive pressure head condition, 6-38–6-40
- functional relationships, 6-11–6-16
hydraulic conductivity functions, 6-15–6-16
particle-size distribution, 6-11–6-13
water retention curves, 6-13–6-15
- hydraulic properties, 6-5–6-10
hydraulic conductivity, 6-9
soil water content, 6-5–6-6
soil water pressure, 6-6–6-7
water retention characteristic, 6-7–6-10
- physical properties, 6-4–6-5
scaling principles, 6-40–6-48
scaling evaporation, 6-47–6-48
scaling infiltration, 6-41–6-47
- soil characteristics, estimation techniques, 6-22–6-30
hydraulic conductivity relation, 6-29–6-31
water retention relation, 6-23–6-29
- soil characteristics, measurement, 6-16–6-22
hydraulic conductivity, 6-19–6-21
soil water content, 6-17–6-18
soil water pressure head, 6-18–6-19
water retention and hydraulic conductivity, combined, 6-21–6-22
- soil phases, 6-3–6-4
- soil water flow, conceptual aspects, 6-31–6-34
general flow equations, 6-32–6-33
infiltration, 6-33–6-34
spatial variability, of soil water properties, 6-31
- Unstable areas, 33-4
- Unsteady-state flow, 10-15
- Upconing, 3-28–3-29, 12-23, 12-26
- Upscaling, 25-7–25-10
chemical heterogeneity methods, 19-5–19-7
effect of, 25-12–25-14
extrapolation, 25-12
microbial dynamics, 19-11–19-13
spatial variability, at catchment scale, 25-10–25-12
- Uranium, 17-27
- Urbanization, 2-51
- US, 1-13, 1-15, 2-2–2-3, 2-38, 2-43–2-44, 2-46, 2-52, 3-25, 17-17, 17-28, 2-9
groundwater regions, hydraulic characteristics, 2-6
legal framework, for groundwater protection, 32-1
- US Department of Agriculture, *see* USDA
- US Department of Energy, 17-28

US Environmental Security Technology Program, 31-9
 US Geological Survey (USGS), 1-12, 1-15, 2-4, 27-2,
 30-9

National Water-Quality Assessment Program, 23-31
 US Geological Survey Fractured Rock Hydrology
 Research Site, 15-25

US Nuclear Regulatory Commission, 17-28
 USDA, 6-3
 soil textural classification chart, 6-4

USEPA, *see* EPA

User index, 6-14

USSR, 1-26

Utah, 2-38

V

Vadose zone, 22-2-22-4, 29-2

infiltration pathways in, 21-5
 soil water balance, 2-13-29-19
see also Unsaturated zone

Vadose/W, 34-16

Valley fill, 33-6

Van der Waal's forces, 17-10

Variable-density flow process, 23-30

Variable-source area, 1-25

Variance, 5-13

Variation, coefficient of, 5-14

Variogram, 16-3

calibration, 16-11
 experimental variogram, 16-8-16-9
 model, 16-6-16-7
 nonnegativity, 16-11
 residuals, 16-9
 selection, 16-7-16-8
 testing, 16-10-16-11

Vector data, 30-3

Vegetative cover, 34-15-34-16

Vegetative layer, 33-23

Veihmeyer, 1-14

Velocity head, 3-6

Velocity potential, 3-20

Vertical rate of strain, 2-19

Very flexible polyethylene (VFPE), 37-12, 37-14

Vienna, 1-28

Vinyl chloride (VC), 31-8-31-10

Virginia, 2-4

Virus, 17-29, 31-2

Viscosity, 2-18-2-19, 18-3

Visual MODFLOW, 30-9

Void ratio, 2-11

Volatilization processes, 36-34-36-38

in *ex situ* treatment, 36-41-36-42
 in *in situ* treatment, 36-34-36-38

Volatilization, 22-13

Volcanic craters, 2-38

Volume of voids, 2-10

W

Wash River, 2-44

Valley, 2-8

Waste, mechanical behavior, 33-12

Waste containment

evapotranspirative cover systems, 34-1

Waste Isolation Pilot Plant, 15-7

Water, in saturated zone, 2-33

artificial protective barriers, 2-45
 bank storage, 2-42

groundwater supplies protection, 2-42-2-46
 natural protective barriers and waste disposal, 2-44

residence time, 2-35-2-36

streams, gaining and losing, 2-40-2-42

surface discharge, 2-36-2-40

water laws, 2-43-2-44

well-head protection programs, 2-45-2-46

Water, in unsaturated zone

atmospheric discharge and seepage face, 2-32-2-33
 hydraulic conductivity and specific discharge,
 2-31-2-32

moisture content vs depth, 2-28-2-29

recharge and infiltration capacity, 2-30-2-31

residence time, 2-32

subsurface stormflow, 2-32

Water and solute movement modeling, 9-1

dual-porosity models

with distribution layer at soil surface, 9-22-9-26
 for gradual change at fracture-matrix interface,
 9-9-9-23

for sharp change at fracture-matrix interface,
 9-3-9-9

multi-pore group models, 9-26-9-28

Water budget, 27-3-27-5

Water compressibility, 2-14-2-15

Water content reflectometer (WCR) probe, 34-18, 34-26

Water laws, 2-43-2-44

Water management, in ancient societies, 1-2-1-3

Water physicochemical characteristics, 17-2-17-3

Water retention and hydraulic conductivity, combined,
 6-21-6-22

Beerkan method, 6-22

Water retention characteristic, 6-7-6-9

Water retention curves, 6-13-6-15

Water retention shape index, 6-14

Water science evolution, during renaissance, 1-4-1-5

Water table, 2-26-2-27, 4-14, 23-18, 30-6

Water vulnerability modeling, using spatial data,
 30-9-30-12

WATERGEN, 30-5

Well, groundwater monitoring, 35-13

decommissioning, 35-22

design, 35-13-35-16

development techniques, 35-21-35-22

drilling methods, 35-16

QA/QC, 35-22

installation monitoring, 35-16-35-21

QA/QC, 35-22

Well, monitoring, 3-39

Well construction

construction methods, 11-24-11-26

maintenance, 11-28-11-29

well development, 11-26-11-28

Well design, 11-13-11-24

casing section, 11-15

design optimization, 11-22

gravel-pack, 11-19-11-20

groundwater remediation wells, 11-22-11-23

monitoring wells, 11-23-11-24

pump, 11-20-11-22

- pump housing length, calculation example, **11-15–11-17**
sand trap, **11-20**
screen section, **11-17–11-18**
well screen, calculation example, **11-18–11-19**
- Well development, **11-26–11-28**
Well-drilling, **1-3**
Well efficiency, **11-7**
Well-fields
 equations, **11-2–11-5**
 aquifer and well losses, **11-5–11-7**
 partial penetration, **11-5**
 step-drawdown tests, **11-7–11-13**
Well flow equations, **10-4–10-12**
 confined aquifers, **10-5–10-6**
 leaky aquifers, **10-7–10-9**
 partial penetration, **10-9–10-11**
 recovery well-flow equations, **10-11–10-12**
 unconfined aquifers, **10-6–10-7**
Well head protection programs, **3-8**
Well hydraulics and aquifer tests, **10-1**
 aquifer test performance, **10-12–10-17**
 data processing, **10-15–10-17**
 date interpretation, **10-17**
 measurements, **10-13–10-15**
 pumping test duration, **10-15**
 aquifer types, **10-2**
 distance-drawdown analyses, **10-26–10-30**
 heterogeneous aquifers, interpretation, **10-30–10-31**
 physical properties, of aquifers, **10-2–10-4**
 time-drawdown analyses
 aquifer test, in confined aquifer, **10-17–10-19**
 aquifer test, in leaky aquifer, **10-22–10-24**
 aquifer test, in unconfined aquifer, **10-19–10-21**
 single-well tests, **10-24–10-26**
well-flow equations, **10-4–10-12**
 confined aquifers, **10-5–10-6**
 leaky aquifers, **10-7–10-9**
 partial penetration, **10-9–10-11**
 recovery well-flow equations, **10-11–10-12**
 unconfined aquifers, **10-6–10-7**
- Well losses, **11-6**, **11-16**
 linear well losses, **11-6**
- nonlinear well losses, **11-6–11-7**
well efficiency, **11-7**
 see also Head loss
- Well volume approach, groundwater sampling, **35-27–35-29**
- Wellbore skin, **20-16**
Wellbore storage, **20-17**
Wellhead protection area (WHPA), **30-13**
Well-head protection programs, **2-45–2-46**
Wellhead, **36-23**
Wellpoints, **36-21**
Wellscreen, **35-21**
 calculation example, **11-18–11-19**
- West Virginia, **2-38**, **2-51**
Westbay[®] System, **20-23**
Wet chemistry analyses, **35-36**
Wetlands, **2-48–2-49**, **33-4**
WHAT, **30-5**
White Sulfur Springs, W. Virginia, **2-38**
Wick pan lysimeter, *see* Passive capillary sampler (PCAPS)
- Wilting point, **2-13–2-14**
Wind-powered mill, **1-4**
World Commission on Environment and Development,
 see Brundtland Commission
- Wyoming, **2-26**
- X**
Xenophanes, **1-4**
XML format, **30-3**
- Y**
Yarn, **37-11**
Yucca Mountain, **17-28**, **19-6**
Yugoslavia, **2-9**
- Z**
Z-test, **24-10–24-12**
Zero flux plane, **6-19–6-20**, **28-12**
Zeroth-order approximation, **9-7–9-8**
Zonal inversion methods, **15-21–15-23**
Zonal slowness model, **15-17**
Zonation, **1-26**

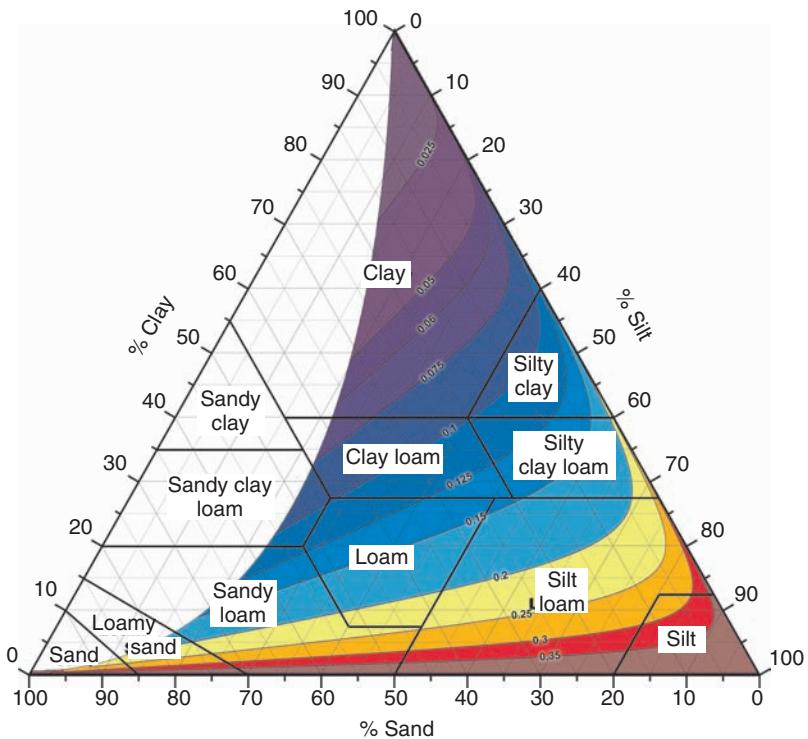


FIGURE 6.7 Contour lines of the particle-size shape parameter M plotted on the basis of the USDA soil textural triangle as a function of the different soil texture classes.

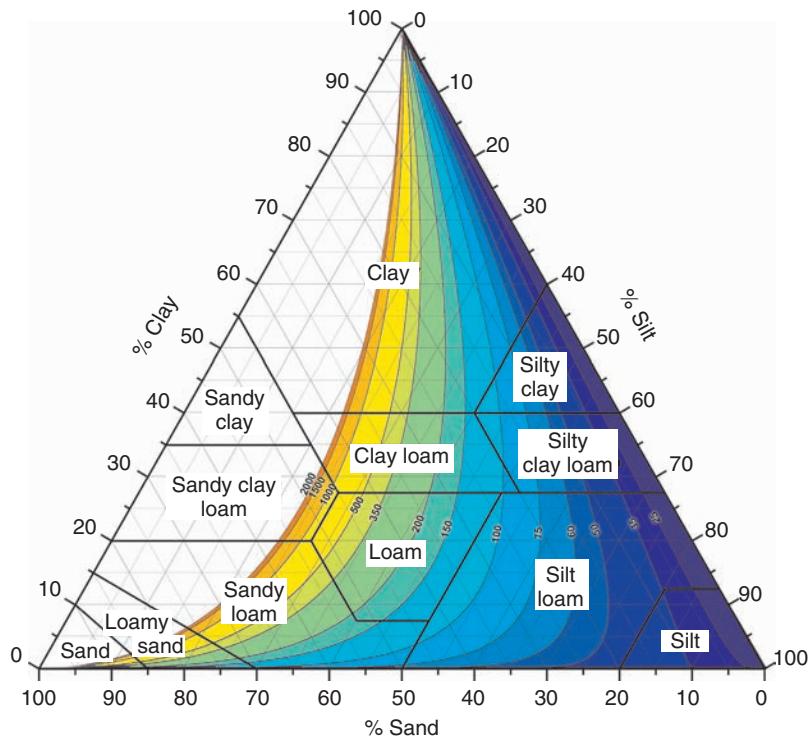


FIGURE 6.8 Contour lines of the particle-size scale parameter D_{pg} (μm) plotted on the basis of the USDA soil textural triangle as a function of the different soil texture classes.

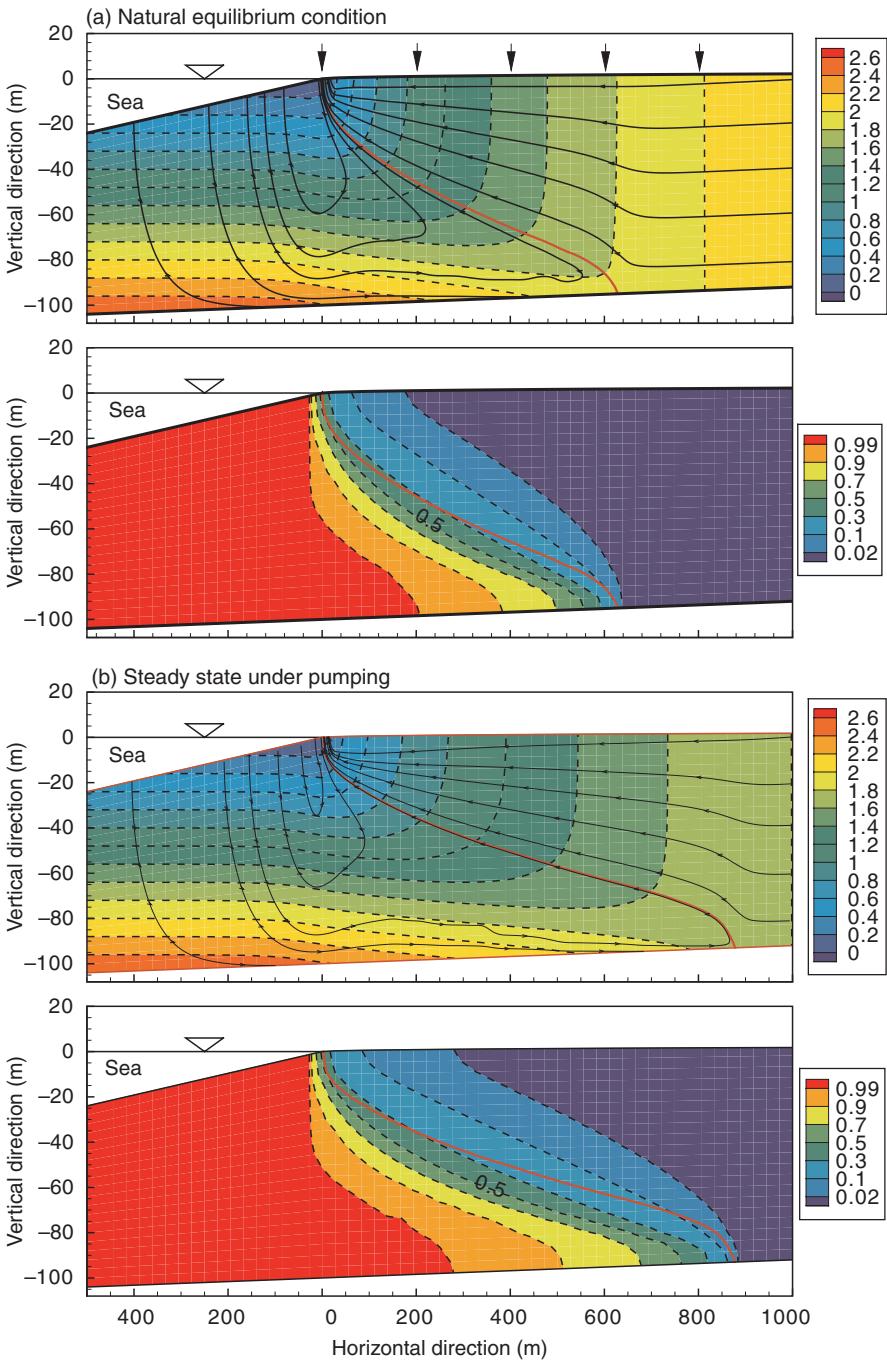


FIGURE 12.5 Numerical model results for (a) a sea water wedge under natural equilibrium conditions (no pumping) and (b) steady-state sea water intrusion under pumping conditions (three pumping wells). In each case, the figure shows the reference head distribution and streamlines in the top figure, and salt mass fraction in the bottom figure, as well as the separation streamline (in red solid line).

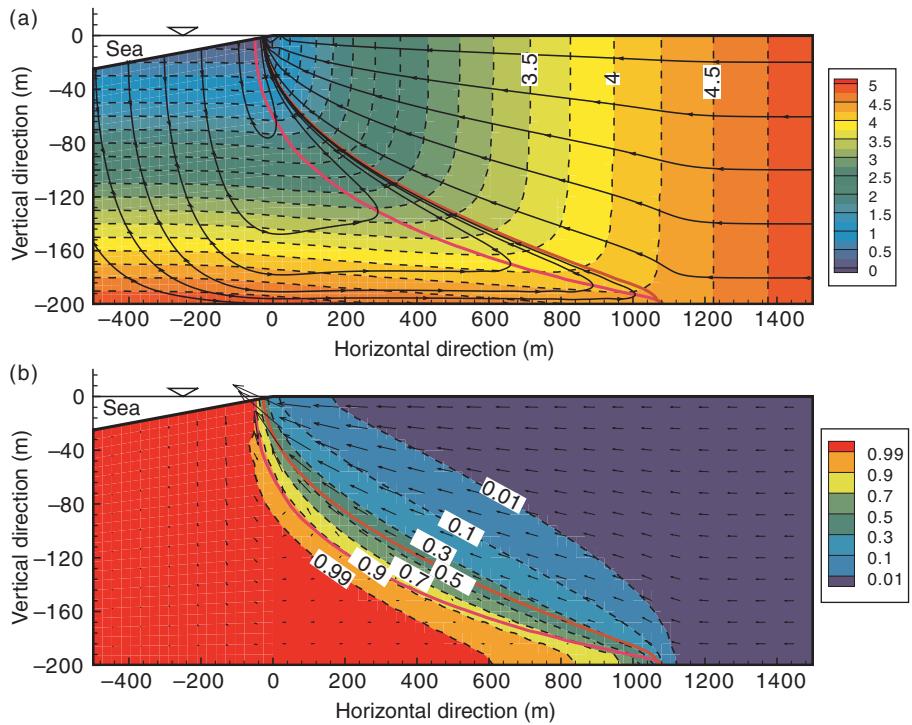


FIGURE 12.7 Steady-state solution of a sea water wedge in a confined aquifer for Example B: (a) the reference hydraulic head distribution (in dashed lines), with the separation streamline (in red line) separating streamlines (in solid lines) originating at the fresh water zone and the sea water zone, and the zero-horizontal-flux line (in pink line), and (b) the mass fraction distribution within the transition zone, and the groundwater flow velocity vector.

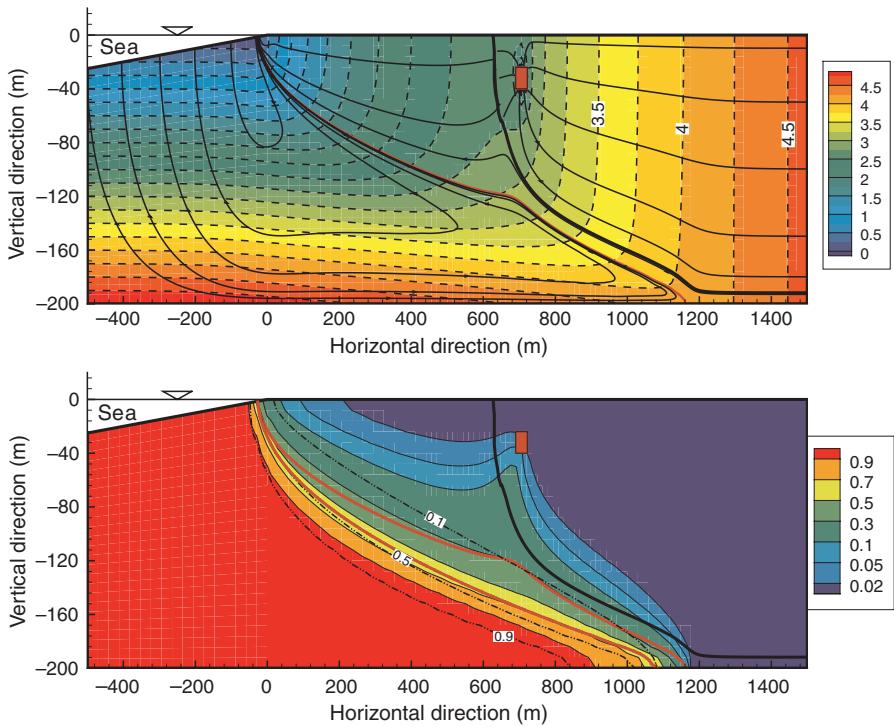


FIGURE 12.8 Transient sea water upconing in a confined aquifer at the end of 8.9 yr in Example B: (a) contours of reference hydraulic head (dashed lines), the separation streamline (red line), and streamlines (solid black lines) originating from the fresh water zone and from the sea water zone, and the capture zone (thicker solid black line), and (b) mass fraction contours (solid lines) with sea water upconing towards the pumping well, in comparison with the initial steady-state transition zone (dashed lines), and the separation streamlines for the upconing and the initial steady-state condition.

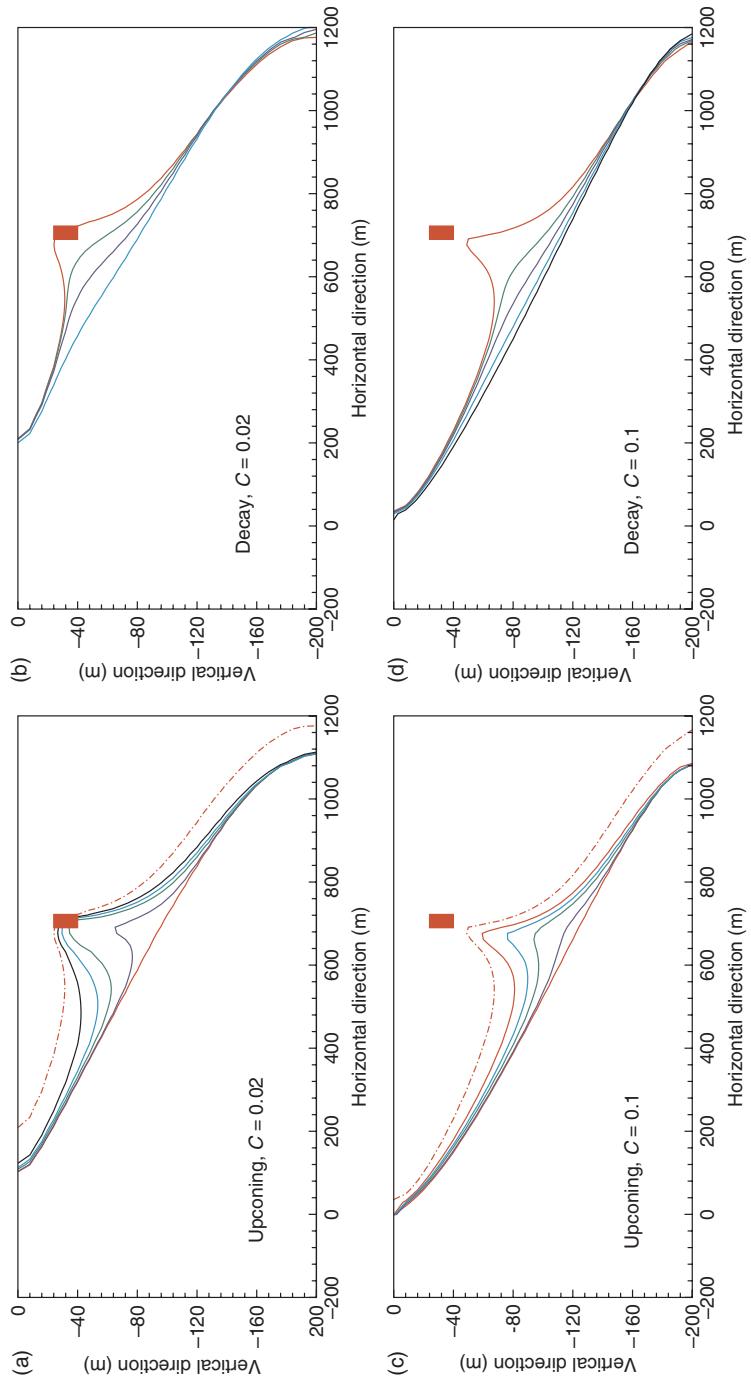


FIGURE 12.9 Transient upconing and decay processes in a confined aquifer (Example B). The times for the upconing process are 53, 239, 439, 839, and 3239 days (well's shutoff), and 3239, 3439, 3639, 4039, and 4639 days for the decay process, all from the beginning of the pumping.

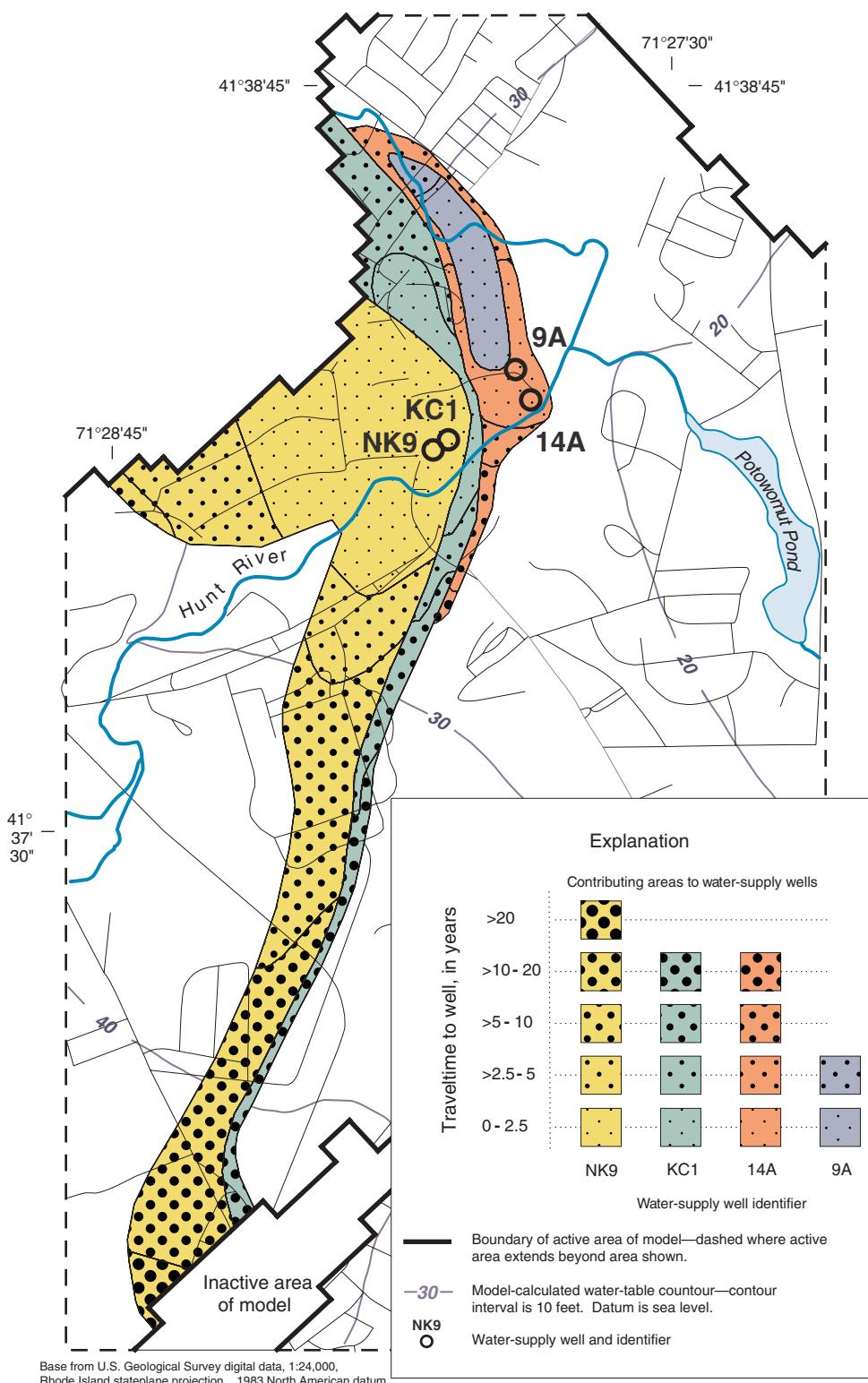


FIGURE 23.15 Model-calculated steady-state contributing areas and traveltimes to water-supply wells in the northern part of the Hunt River Basin, Rhode Island. (Modified from Barlow, P.M. and Dickerman, D.C. (2001) Numerical-simulation and conjunctive-management models of the Hunt-Annquatucket-Pettaquamscutt stream-aquifer system, Rhode Island. U.S. Geol. Survey Prof. Paper 1636.)

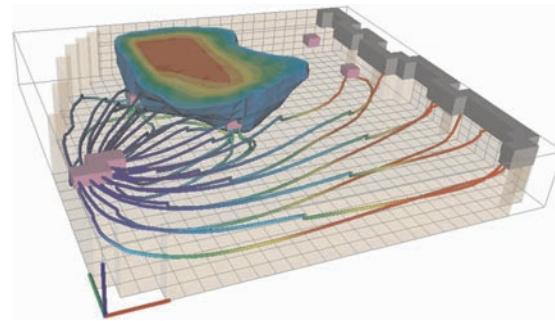


FIGURE 23.23 Example plot generated with Model Viewer software showing the grid shell, selected model features and boundary conditions, particle pathlines (shaded by time of travel), and a solid rendering of a contaminant plume for a 3D MODFLOW simulation. (From Hsieh, P.A. and Winston, R.B. (2002) User's guide to Model Viewer, a program for three-dimensional visualization of ground-water model results, U.S. Geol. Survey Open-File Rept. 02-106.)

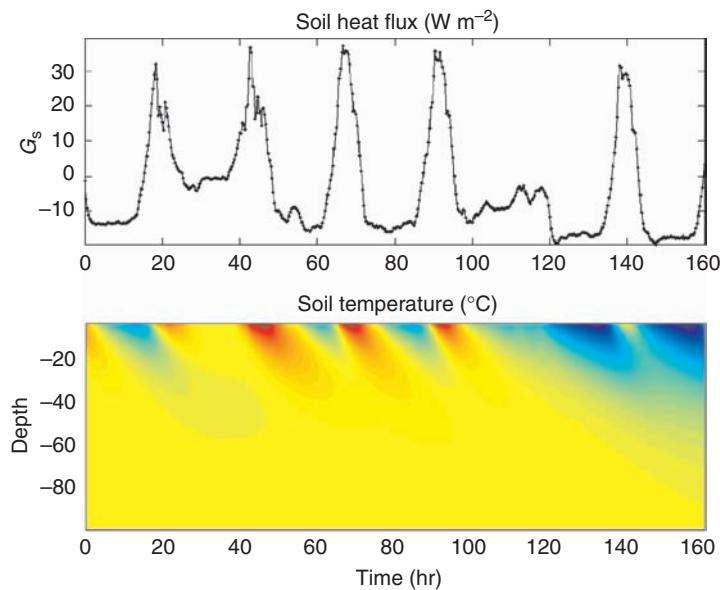


FIGURE 29.11 Measured soil heat flux (G_s) near the land-atmosphere interface and measured soil temperature (ranging from 12, black, to 18°C, white) within a grass-covered surface illustrating the warming and cooling heat waves.

Water Science, Technology, and Engineering

THE HANDBOOK OF

GROUNDWATER ENGINEERING

Second Edition

A complete treatment of the theory and practice of groundwater engineering, *The Handbook of Groundwater Engineering, Second Edition* provides a current and detailed review of how to model the flow of water and the transport of contaminants both in the unsaturated and saturated zones, and covers the production of groundwater and the remediation of contaminated groundwater.

SEE WHAT'S NEW IN THE SECOND EDITION

- New and expanded material on modeling includes groundwater modeling, model validation, modeling salt water intrusions in coastal aquifers, and modeling of contaminant transport in the unsaturated zone
- New chapters explore the history of groundwater hydrology, infiltration and run-on under spatially variable hydrologic properties, sea water intrusion, groundwater and heat flow, scale issues, model validation, ecohydrology, groundwater sustainability, impact of climate change on groundwater, and evaporative covers systems
- Three chapters rewritten by new authors cover infiltration, nonreactive contaminant transport in the saturated zone, and contaminant transportation in the unsaturated zone: theory and modeling
- Revised chapters on soil properties and moisture movement in the unsaturated zone, aquifer characterization, karst aquifers, biodegradation, legal framework for groundwater protection, groundwater monitoring, and remediation of contaminated groundwater

Presenting a multifaceted and up-to-date discussion of groundwater engineering, the book explores groundwater production, well hydraulics, well design and construction, groundwater sustainability, biodegradation techniques, the use of geosynthetics, and the design of landfills and evaporative cover systems from an applied perspective. It addresses legal and regulatory aspects as well as environmental concerns such as ecohydrology and the impact of climate change on groundwater. With contributions from an international group of authors drawn from academic institutions, consulting firms, and government agencies, the book is both the most up-to-date and the most authoritative reference available.



CRC Press

Taylor & Francis Group
an informa business

www.taylorandfrancisgroup.com

6000 Broken Sound Parkway, NW
Suite 300, Boca Raton, FL 33487
270 Madison Avenue
New York, NY 10016
2 Park Square, Milton Park
Abingdon, Oxon OX14 4RN, UK



4316

ISBN 0-8493-4316-X

9 0000



9 780849 343162

www.crcpress.com