APPLIED
GEOSTATISTICS
SERIES

GEOSTATISTICAL RESERVOIR MODELING

Clayton V. Deutsch

GEOSTATISTICAL RESERVOIR MODELING

APPLIED GEOSTATISTICS SERIES

General Editor
André G. Journel

Clayton V. Deutsch, *Geostatistical Reservoir Modeling*

Jean-Laurent Mallet, *Geomodeling*

Pierre Goovaerts, *Geostatistics for Natural Resources Evaluation*

Clayton V. Deutsch and André G. Journel, *GSLIB: Geostatistical Software Library and User's Guide, Second Edition*

Edward H. Isaak and R. Mohan Srivastava, *An Introduction to Applied Geostatistics*

کتابخانه اکتشاف

GEOSTATISTICAL RESERVOIR MODELING

Clayton V. Deutsch

شرکت ملی نفت ایران - کتابخانه اکتشاف
شماره مسلسل: 7493
شماره کتاب: TN870/571 D48

OXFORD
UNIVERSITY PRESS
2002

OXFORD
UNIVERSITY PRESS

Oxford New York
Auckland Bangkok Buenos Aires Cape Town Chennai
Dar es Salaam Delhi Hong Kong Istanbul Karachi Kolkata
Kuala Lumpur Madrid Melbourne Mexico City Mumbai Nairobi
São Paulo Shanghai Singapore Taipei Tokyo Toronto

and an associated company in Berlin

Copyright © 2002 by Oxford University Press, Inc.

Published by Oxford University Press, Inc.
198 Madison Avenue, New York, New York 10016

www.oup.com

Oxford is a registered trademark of Oxford University Press.

All rights reserved. No part of this publication may be reproduced,
stored in a retrieval system, or transmitted, in any form or by any means,
electronic, mechanical, photocopying, recording, or otherwise,
without the prior permission of Oxford University Press.

Library of Congress Cataloging-in-Publication Data
Deutsch, Clayton V.
Geostatistical reservoir modeling / Clayton V. Deutsch.
p. cm. — (Applied geostatistics series)
Includes bibliographical references and index.
ISBN 0-19-513806-6
1. Hydrocarbon reservoirs—Mathematical models. 2. Petroleum—Geology—Statistical methods.
I. Title. II. Series.
TN870.57 .D48 2001
622'.1828'015118—dc21 2001016405

9 8 7 6 5 4 3 2 1

Printed in the United States of America
on acid-free paper

Preface

This book presents geostatistical tools for building numerical geological models of petroleum reservoirs. The use of geostatistical tools for modeling reservoir heterogeneities and assessing uncertainty in reservoir forecasting has increased significantly since the 1990s. There are many technical papers that track this growth in theory and applications; however, few collected works exist to bring together the practice of petroleum geostatistics into a coherent framework.

Building a suitable reservoir model calls upon the skills of many people. Geologists provide critical input on the sedimentology and stratigraphy of the subsurface. Geophysicists provide valuable information on the geometry of the reservoir and the internal distribution of reservoir properties in the interwell region. Engineers, with their knowledge of flow processes and production data, provide critical information on connectivity and major heterogeneities. I do not address any single discipline; the emphasis is on the interdisciplinary interaction necessary to build numerical geological models consistent with all available information sources.

This book is aimed at the *practice* of geostatistical reservoir modeling; it is not intended to be a theoretical reference textbook. I focus on tools, techniques, examples, tradecraft, and providing guidance on the practice of petroleum reservoir modeling. *Geostatistical Reservoir Modeling* would serve as a reference text for the practitioner and be appropriate for an advanced undergraduate or graduate class on reservoir characterization.

Acknowledgments

Many people and organizations contributed to the preparation of this book. I would like to extend special acknowledgment to the industrial affiliates of the Center for Computational Geostatistics (CCG) at the University of Alberta.

André Journel has been a mentor and friend for more than 15 years of geostatistical adventure. His advice, insight, encyclopedic knowledge of geostatistics, analytical and literary skills have been invaluable. André's influence can be seen on virtually every page of this book; however, I must confess to taking short cuts and demonstrating less rigor than he would accept if I were still his student. The maxim "expand or remove" still rings in my ears and I have chosen the latter too often.

I dedicate this book to my family for their love and encouragement. My wife, Pauline, and children, Jared, Rebecca, and Matthew, are a tremendous source of joy and stability.

Edmonton, Alberta C.V.D
July 2001

Contents

1 Introduction 3
- 1.1 Plan for the Book 4
- 1.2 Key Concepts 6
- 1.3 Motivation for Reservoir Models 10
- 1.4 Data for Reservoir Modeling 12
- 1.5 An Introductory Example 14
- 1.6 Work Flow 21

2 Preliminary Statistical Concepts 31
- 2.1 Geological Populations and Stationarity 32
- 2.2 Notation and Definitions 33
- 2.3 Bivariate Distributions 41
- 2.4 Q-Q Plots and Data Transformation 44
- 2.5 Declustering 50
- 2.6 Histogram and Cross Plot Smoothing 64
- 2.7 Monte Carlo Methods and The Bootstrap 65
- 2.8 Work Flow 70

3 Gridding Reservoir Layers 79
- 3.1 Gridding for Geologic Modeling 80
- 3.2 Stratigraphic Correlation/Coordinates 85
- 3.3 Faults 93
- 3.4 Uncertainty in Reservoir Geometry 96
- 3.5 Work Flow 96

4 Quantifying Spatial Correlation 101
- 4.1 The Random Function Concept 102
- 4.2 Calculating Experimental Variograms 105
- 4.3 Interpreting Experimental Variograms 114
- 4.4 Horizontal Variograms 124
- 4.5 Variogram Modeling 131
- 4.6 Cross Variograms 139
- 4.7 Work Flow 148

5 Preliminary Mapping Concepts 153
5.1 Kriging 154
5.2 Sequential Gaussian Simulation 162
5.3 Direct Sequential Simulation 167
5.4 Indicator Formalism 168
5.5 P-Field Methods 175
5.6 Accounting for Trends 178
5.7 Work Flow 187

6 Cell-Based Facies Modeling 193
6.1 Choosing the Appropriate Method 193
6.2 Sequential Indicator Simulation 196
6.3 Truncated Gaussian Simulation 204
6.4 Cleaning Cell-Based Facies Realizations 210
6.5 Work Flow 216

7 Object-Based Facies Modeling 223
7.1 Background 223
7.2 Stochastic Shales 229
7.3 Fluvial Modeling 230
7.4 Non-Fluvial Depositional Systems 238
7.5 Work Flow 241

8 Porosity and Permeability Modeling 245
8.1 Background 246
8.2 Gaussian Techniques for Porosity 249
8.3 Porosity/Permeability Transforms 254
8.4 Gaussian Techniques for Permeability 260
8.5 Indicator Technique for Permeability 262
8.6 Work Flow 268

9 Simulated Annealing for Geostatistics 275
9.1 Background 275
9.2 Steps in Annealing 277
9.3 Problem Areas 291
9.4 Place of Simulated Annealing/Work Flow 293

10 Uncertainty Management 295
10.1 Models of Uncertainty 296
10.2 Cross Validation and the Jackknife 298
10.3 Checking Distributions of Uncertainty 299
10.4 How Many Realizations? 309
10.5 Ranking Realizations 313
10.6 Decision Making with Uncertainty 316
10.7 Work Flow 320

11 Special Topics 325
11.1 Scale Up from Core to Modeling Cell 326
11.2 Surface-Based Modeling 328
11.3 Multiple Point Statistics 330
11.4 Dynamic Data 331
11.5 Input to Flow Simulation 332
11.6 Final Thoughts 334

A Glossary and Notation 337
A.1 Glossary 337
A.2 Notation 347

Bibliography 351

Index 372

GEOSTATISTICAL RESERVOIR MODELING

Chapter 1

Introduction

Geostatistical tools are being increasingly used for the modeling of petroleum reservoirs and yet the theoretical descriptions of those tools, case studies, and tradecraft details are widely dispersed in a variety of technical journals, books, and software not all related to petroleum geostatistics. This work aims at collecting that dispersed information into a handbook for production geologists, reservoir geophysicists, and reservoir engineers involved in building and using numerical reservoir models.

No software is included in this book. Algorithms are described with flow charts that illuminate the methodology. Source code for most algorithms can be found in the public domain GSLIB software [62]. There are a number of commercial software packages available, [28, 168, 209, 224, 241] among others.

This book is written for undergraduate to graduate students and the practicing geologist, geophysicist, and engineer desiring to use geostatistical tools. The essential theory will be developed in the context of modeling petroleum reservoirs. Most theoretical sections could be skipped by readers with a "high-level" understanding of geostatistics. Those who seek to major in petroleum geostatistics, however, should master the modest statistics and mathematics presented here. Original work and papers presenting greater detail are cited for those seeking a solid background.

This book is not designed around the modeling of a single reservoir. Instead, a number of different reservoir data sets will be used to illustrate specific techniques. This makes it easier to present the diversity of useful techniques and to demonstrate the relative strengths and weaknesses of alternative approaches.

Various methods are presented. Guidance is given on when to use each particular method. A fascinating, but often frustrating, aspect of geostatistics is that there is a large combination of tools that can be brought to bear on any problem. No attempt has been made to present an exhaustive list of tools; however, citations to methods not developed have been included. Omissions are the responsibility of the author.

1.1 Plan for the Book

This book consists of 11 chapters addressing the application of geostatistics to reservoir modeling. A petroleum-specific glossary of geostatistical terms and a list of geostatistics notation is given at the end. An extensive bibliography provides a resource for newcomers to the field. Acknowledging that few readers can pick up a book and read it cover to cover, a comprehensive index has been prepared. This book attempts to answer five questions:

1. What is reservoir modeling and what is the role of geostatistics?
2. What essential background do we need before we engage in geostatistical reservoir modeling? (Chapters 2 through 5)
3. What are the main steps in reservoir modeling and where/how does geostatistics intervene? (Chapters 6 through 9)
4. What do we do with the models once we have them? (Chapter 10)
5. What promising new developments might we foresee in the future? (Chapter 11)

A more specific description of each chapter.

This first chapter presents an overview of the book, the steps and methodologies involved in typical reservoir modeling, and how the different disciplines of geology, geophysics, geostatistics, and petroleum engineering work together. Flow charts are presented throughout the book to illustrate the flow of data, the geostatistical operations, and the decisions to be taken during reservoir modeling. High-level flow charts are presented in this chapter; detailed flow charts for specific tasks are presented in later chapters.

Preliminary statistical concepts related to reservoir modeling are covered in Chapter 2. Each core plug is unique and yet we must pool samples together before calculating any statistics or building a reservoir model. Considerations for choosing geological populations are presented. There are a number of classical statistical tools that can help with the selection and description of geological populations. Introductory geostatistical tools, such as declustering and correcting for non-representative data, are also presented in Chapter 2.

Chapter 3 discusses gridding of large-scale reservoir layers. Geostatistical tools are particularly well suited to work in convenient rectangular Cartesian coordinate systems; however, most reservoirs are complicated by structural deformation, erosion, and faulting. This chapter reviews coordinate transformations to represent or grid complex real-world reservoirs in Cartesian coordinates.

Chapter 4 presents the variogram as the main geostatistical tool for quantifying spatial variability. Variogram calculation, interpretation, and modeling are covered with numerous examples from real reservoirs. Indicator variograms are presented as more sophisticated measures of spatial correlation.

Chapter 5 presents the background for kriging-based geostatistical techniques. The use of kriging and cokriging is widespread in geostatistical modeling. The background of such techniques and their practical application is discussed. The p-field simulation technique is presented. The use of locally varying directions of continuity and trend models is presented. This chapter collects essential geostatistics theory together with small illustrative examples. More extensive case studies are presented in later chapters.

The distribution of facies must be established prior to modeling of continuous properties such as porosity and permeability. Chapter 6 presents cell-based facies modeling techniques, that is, methods that assign facies on a cell-by-cell basis with statistical control. Sequential indicator simulation and truncated Gaussian simulation, with their variants, are the important cell-based techniques. Procedures to remove unwanted short-scale variations (noise) from the resulting realizations are also discussed.

Object-based facies modeling is described in Chapter 7. Hierarchical, object-based modeling schemes are considered appropriate in specific geological environments where the facies are organized in clear geometric units. Fluvial channels, crevasse splays, levees, and turbidite lobes are some examples of facies units suitable for object-based modeling.

Details of building 3-D models of continuous variables such as porosity and permeability are presented in Chapter 8. The alternative approaches will be presented together with practical aspects of uncertainty assessment and accounting for secondary data such as seismic data. Techniques to capture the continuity of extreme permeability values and correlation with porosity are discussed in this chapter.

The use of simulated annealing methods for geostatistical applications has been hindered by excessive CPU times and the need to set many delicate tuning parameters. Nevertheless, the capability of annealing for difficult data integration problems is yet unsurpassed. Chapter 9 presents the background and implementation details of simulated annealing for geostatistical modeling. Issues such as the artifacts and "space of uncertainty" of simulated annealing is discussed.

The important subjects of model verification and assessing uncertainty are covered in Chapter 10. How do we judge the suitability of our geostatistical reservoir models? What do we do with multiple realizations? How can decisions be made in the face of uncertainty? These questions and others related to checking model *goodness* and managing uncertainty will be addressed.

Finally, Chapter 11 reviews a number of special topics in geostatistical reservoir modeling. These topics deserve more complete treatment, but are beyond the scope of this book. Some of the special topics include (1) techniques to account for the *missing scale* from the data scale to the geological modeling scale, (2) techniques to model stochastic geological bounding surfaces for increased geologic realism they provide, (3) multiple point statistics, (4) integration of dynamic data such as time lapse seismic data, permanent

pressure gauge data, and multiphase historical production data, and (5) input to flow simulation.

The jargon of geostatistics and reservoir modeling can be overwhelming. Different words are often used to describe the same activity. For example, *variography* and *structural analysis* both refer to the calculation, interpretation, and model fitting of measures of spatial correlation. Moreover, the same words are often used to describe different activities: *simulation* can refer to the stochastic simulation of petrophysical properties or to the subsequent flow simulation. The *Geostatistical Glossary and Multilingual Dictionary* [197] has been used as a starting point for most definitions. The notations developed in GSLIB [62] and expanded upon in the book by Goovaerts [110] have been used as much as possible. For completeness a glossary and list of notation are included in the Glossary.

1.2 Key Concepts

Predicting rock properties at unsampled locations and forecasting the future flow behavior of complex geological and engineering systems is difficult. A number of concepts and assumptions make it possible to address this difficult problem.

Petrophysical Properties

Techniques presented in later chapters are concerned with constructing high-resolution 3-D models of facies types, porosity, and permeability. The definition of these three variables is often ambiguous and the related modeling choice subject to debate. *Facies* are typically based on an assessment of geological variables such as grain size or mineralization, for example, limestone and dolomite in a carbonate setting and channel sandstone and shale in a siliciclastic setting. We model facies first because it narrows the range of possible porosity and permeability (see definition below) and because multiphase fluid flow properties can vary significantly between facies types. There are different definitions and types of facies: lithofacies, electrofacies, flow facies, depositional facies and so on. The choice of facies or rock types is usually evident for a particular reservoir modeling problem. The issue of facies definition is addressed in greater detail in Chapter 6.

Porosity ϕ is the fraction of void space in the rock that may contain fluid. We are interested in the "effective" porosity that contributes to fluid flow rather than the "total" porosity, which includes small isolated pores. The spatial distribution of porosity and the total volume of in-place hydrocarbon is of significant interest to reservoir modelers.

Permeability K is a measure of the ease with which the rock allows fluids to flow through it. Permeability depends not only on the direction of flow but on the local pressure conditions, also called boundary conditions. The commonly used "no flow" boundary conditions are particular to laboratory measurement

Table 1.1: Variables considered in this book, generalizations, and analogous variables in other disciplines.

Variable	Generalization	Analogous Variables
Facies	Categorical variable	Rock or soil type Predominant biological species
Porosity	Volumetric concentration	Mineral grades Contaminant concentration Fracture density Pest concentration Cumulative annual rainfall
Permeability	Dynamic property	Conductivity Dispersivity

devices and many scale-up algorithms; however, actual boundary conditions in the reservoir will certainly be different. Moreover, permeability is a tensor; a pressure gradient in the horizontal direction can induce vertical flow through K_{XZ} permeability terms.

It is necessary to have *hard* truth measurements at some scale. Most often, these *hard* data are the facies assignments, porosity, and permeability observations taken from core measurements. In the absence of direct core measurements, well log data may be the hard data. All other data types including well logs and seismic are called *soft* data and must be calibrated to the hard data.

Although this book focuses on facies, porosity, and permeability in the context of reservoir modeling, these petrophysical properties are also characteristic of many earth science-related spatial variables; see Table 1.1.

Modeling Scale

It is neither possible nor optimal to model the reservoir properties at the resolution of the *hard* core data. The core data must be scaled up (averaged) to some intermediate resolution. Models are generated at that intermediate geological modeling scale, and then usually scaled to an even coarser resolution for flow simulation. An important case- specific issue is to determine the appropriate intermediate geological modeling scale. A too-small choice leads to large and inefficient computer use, which restricts the number of alternative scenarios and sensitivity runs that can be considered. A too-large choice could lead to incorrect flow results due to inadequate representation of important subgrid geological heterogeneities.

Geostatistical reservoir models are built for specific goals and the level of detail should be suitable for the stated goals. In practice, the goals tend to

change and the models are used for more than originally intended. For this reason, it is appropriate to err on the side of building in too much detail even when not explicitly called for by the present modeling objectives.

All data must be related one-to-each-other at the intermediate geological modeling scale. Core and well log derived measurements, which represent a smaller scale, must be scaled up. The larger scale of seismic and production-derived data must be accounted for (essentially scaled down to the modeling resolution) during the construction of the geostatistical model. Dealing with data of different scales that measure petrophysical properties with varying levels of precision is a central problem addressed in subsequent chapters.

Numerical Modeling

At any point in geological time, there is a single true distribution of petrophysical properties in each reservoir. This true distribution is the result of a complex succession of physical, chemical, and biological processes. Although the physics of these depositional and diagenetic processes may be understood quite well, we do not completely understand all of the processes and their interaction, and could never have access to the initial and boundary conditions in sufficient detail to provide the unique true distribution of dynamic properties within the reservoir. We can only hope to create numerical models that mimic the physically significant features.

In general, the large-scale features are the most critical for prediction of reservoir flow performance. Small-scale geological details are summarized by effective properties at a large scale.

We will strive to make our numerical models consistent with the available data, including historical flow results. It is understood that the true distribution of petrophysical properties will not follow *any* of our relatively simplistic mathematical models. It is also understood that visual appearance alone is inadequate to judge the acceptability of a numerical model. The suitability of a numerical model is judged by its ability to accurately predict future flow performance under different boundary conditions; "we are not interested in pretty pictures."

Uncertainty

As indicated above, it is not possible to establish the unique true distribution of facies, porosity, and permeability between widely spaced wells. All numerical models would be found in error if we were to excavate that interwell volume and take exhaustive measurements: There is uncertainty.

This uncertainty exists because of our ignorance/lack of knowledge. It is not an inherent feature of the reservoir. Notwithstanding the elusiveness of the concept of uncertainty, models of uncertainty will be constructed and care will be taken to ensure that these models honestly represent our state of incomplete knowledge. These assessments of uncertainty are nothing more than models, however, and it will not be possible to rigorously validate them.

Geostatistical techniques allow alternative numerical models, also called realizations, to be generated. The response of these realizations, say time to water breakthrough, could be combined in a histogram as a model of uncertainty. The parameters of the geostatistical modeling technique are also uncertain and they could be made variable, that is, "randomized," to lead to a larger and possibly more realistic assessment of uncertainty. The modeling approach itself could be considered uncertain and alternative approaches or geological scenarios considered. At some point this circular quest for a realistic assessment of uncertainty must be stopped [149].

Uniqueness and Smoothing

Conventional mapping algorithms were devised to create smooth maps to reveal large-scale geologic trends; they are low-pass filters that remove high-frequency property variations. The goal of conventional mapping algorithms such as kriging, splines, inverse distance, and contouring algorithms is *not* to show the full spectrum of patterns or variability of the property being mapped. For fluid flow problems, however, the spatial patterns of extreme high and low values of permeability often have a large effect on the flow response.

Geostatistical simulation techniques, conversely, are devised with the goal of introducing the full variability, creating maps or realizations that are neither unique nor smooth. Although the small-scale variability of these realizations may mask large-scale trends, geostatistical simulation is more appropriate for predicting flow performance and modeling uncertainty.

Analogue Data

There are rarely enough data to provide reliable statistics, especially horizontal measures of continuity. For this reason, data from analogue outcrops and similar more densely drilled reservoirs are used to help infer spatial statistics that are impossible to calculate from the present subsurface reservoir data.

We acknowledge that there are general features of certain geological settings that can be transported to other reservoirs, provided they originate from similar geological processes. Although the use of analogue data is essential in reservoir modeling, it should be critically evaluated and adapted to fit any hard data from the reservoir being studied.

Data Integration

The goal of geostatistical reservoir modeling is the creation of detailed numerical 3-D geologic models that *simultaneously* account for (honor) a wide range of relevant geological, geophysical, and engineering data of varying degrees of resolution, quality, and certainty. This data integration must be accomplished by *construction* rather than by selection. The probability of finding

a realization that happens to match data not used in model construction is essentially zero.

A two-step trial-and-error approach could be taken to account for production data: (1) create many realizations that do not explicitly account for historical production data, and then (2) forward simulate the historical production to identify realizations that just happen to reproduce the production data. This approach is not recommended since no realization will match all of the production data. A priori data integration must be attempted since a posteriori screening of realizations will eliminate all possibilities.

Dynamic Reservoir Changes

Geostatistical modeling provides static descriptions of petrophysical properties. In general, it is not recommended to add time as a fourth dimension in the geostatistical model itself and predict future pressure and saturation changes. Such predictions will not necessarily honor important physics such as the conservation of mass and energy. Time dependent changes in pressure and fluid saturations are best modeled with a flow simulator that encodes these physical laws; see for example [17].

The Place of Geostatistics

Geostatistics was not developed as a theory in search of practical problems. On the contrary, the discipline was gradually developed by engineers and geologists faced with real problems and searching for a consistent set of numerical tools that would help them address those real problems. Reasons for seeking such comprehensive technology included (1) an increasing number of data to deal with, (2) a greater diversity of available data at different scales and levels of precision, (3) a need to address problems with consistent and reproducible methods, (4) an availability of computational and mathematical developments in related scientific disciplines, and (5) an understanding that more responsible and profitable decisions would be made with improved numerical models.

This book describes a set of geostatistical tools that have proven useful in practical reservoir modeling. The application of geostatistical (or any geological modeling approach for that matter) depends on the aforementioned principles.

1.3 Motivation for Reservoir Models

A high-resolution 3-D reservoir model that is consistent with the available data may seem reason enough to consider geostatistical reservoir modeling; reconciling all available hard and soft data in a numerical model has many intangible benefits such as (1) transfer of data between disciplines, (2) a tool to focus attention on critical unknowns, and (3) a vehicle to present

spatial variations that may enhance or work against a particular production strategy. In addition to these intangible benefits, there are specific reasons for constructing high-resolution 3-D models:

- There is a need for reliable estimates of the original volume of hydrocarbon in the reservoir. These in situ volumes are important to (1) determine the economic viability of producing a given reservoir, (2) allocate equity among multiple owners, (3) compare the relative economic merits of alternative reservoirs, and (4) determine the appropriate size of production facilities. Geostatistical reservoir modeling provides the numerical models for such volumetric estimates.

- Well locations must be selected to be economically optimal and robust with respect to uncertainty in the reservoir description. Multiple geostatistical realizations make it possible to evaluate the uncertainty underlying well locations and permit selection of "good" well locations. These realizations also make it possible to address related questions such as (1) what type of wells, for example, horizontal, vertical, multilateral should be considered, and (2) how many wells are required?

- There is often a need to reconcile an abundance of soft data (say, from a 3-D seismic survey or years of historical production data) with a limited amount of hard well data. Geostatistical models allow different types of data to be represented in a common format, for example, seismic data can be represented at the scale and with the units of the hard data.

- The 3-D static connectivity of a reservoir can be assessed through simple visualization and various connectivity tools to assess the potential for bypassed oil and the value of infill wells, prior to performing flow simulation.

- Flow simulation allows the prediction of reservoir performance with different production scenarios. The optimal number of wells and operating conditions can be established by comparing the economics of various alternatives. The use of flow simulation was initially hampered by the limited resolution of flow models (primarily a computer hardware constraint) and over simplistic geological input (layercake models). Geostatistical methods now provide the required numerical models of rock properties such as facies, porosity ϕ, and permeability K.

- Major decisions must be made in the presence of significant uncertainty: how many wells? well locations? timing of injection and infill production wells? The goal of many energy producing companies is to make such decisions in a way that maximizes profitable production of the resource. From a geostatistical perspective, these decisions must be robust with respect to the inherent uncertainty in the spatial distribution of reservoir properties. This is illustrated in Chapter 10.

The Benefit of Geostatistics

Ideally, there would be clear documentation of the quantifiable benefits of geostatistics to motivate its use. There are a number of cases studies available in the literature that illustrate the utility of geostatistical reservoir models [5, 6, 39, 47, 203, 215, 252, 275]. In some of these cases studies the combined use of geostatistics and flow simulation lead to a clearly different reservoir management decision than would be obtained with a conventional geological model. One can also argue that the geostatistical models are better since they honor more reservoir-specific information. Notwithstanding the previous arguments, it is difficult to quantify the *worth* or *value* of geostatistics except in simplistic synthetic examples. In some cases, the demonstrated benefit of geostatistical reservoir modeling is in part due to the decision to spend additional effort on reservoir characterization; even conventional methods would do better under such effort.

1.4 Data for Reservoir Modeling

Figure 1.1 schematically illustrates the different types of relevant data for reservoir modeling. Data for reservoir modeling includes reservoir-specific data:

- Core data (ϕ and K by facies) available from a limited number of wells. Although core data or reliable well logs from horizontal wells would be valuable to assess horizontal measures of spatial continuity, such data are still uncommon. Core may be sampled by relatively small core plugs or entire sections of core may be measured for "whole core" data.
- Well log data provide precise information on stratigraphic surfaces and faults, as well as measurements of petrophysical properties such as facies types, ϕ, and, perhaps, soft K measurements.
- Seismic-derived structural interpretation, that is, surface grids and fault locations that define large-scale reservoir geometry.
- Seismic-derived attributes provide information on large-scale variations in facies proportions and porosity.
- Well test and production data yield interpreted permeability thickness, channel widths, connected flow paths, and barriers. The use of inversion methods provide coarse scale ϕ, K models with an associated measure of uncertainty.
- A sedimentologic and chrono (sequence) stratigraphic interpretation provides information on layering and the continuity and trends within each layer of the reservoir.

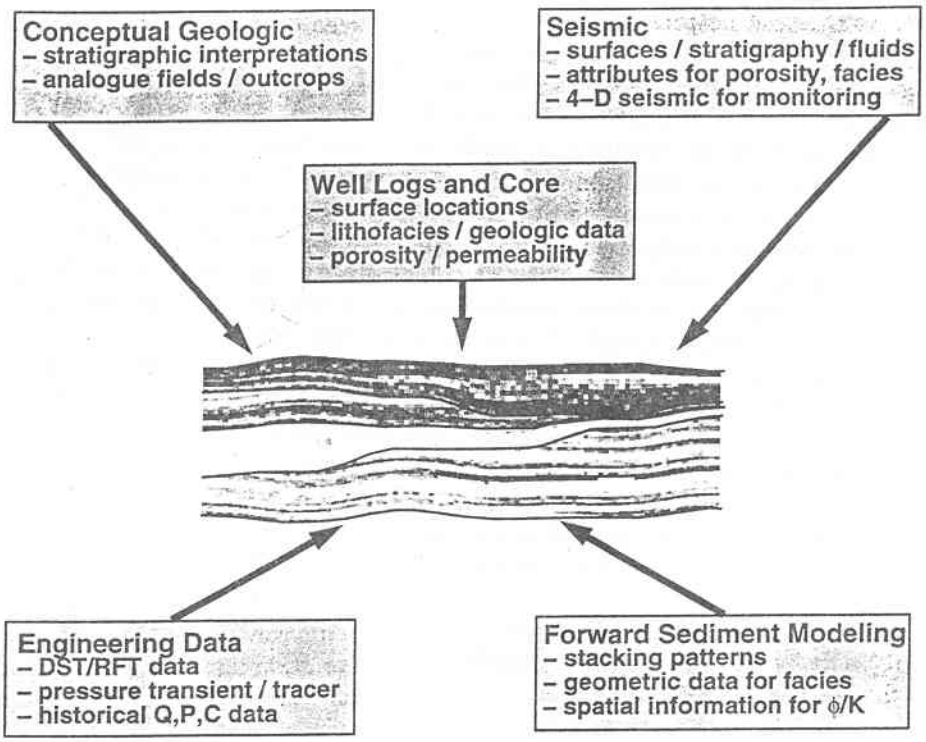


Figure 1.1: A reservoir model that honors all relevant data. The image in the center schematically illustrates a complex 3-D reservoir. The different types of relevant data are described in the surrounding boxes.

In addition to this reservoir-specific data, there is more general knowledge available from analogue data:

- Trends and stacking patterns available from a regional geological interpretation. For example, it may be known that the porosity distribution is fining upward or coarsening upward. Also, the position of the reservoir in the overall regional geological setting could provide soft information on areal trends in the distribution of facies proportions.
- Outcrops or more densely drilled similar fields may provide size distributions, variogram measures of lateral continuity, and other *exportable* spatial statistics.
- Knowledge of geological processes/principles established through widely accepted theories (forward geologic modeling) provides general geologic information of the type described above.

Although there is a large diversity in the data available for reservoir modeling, that data is sparse relative to the vast interwell region that must be modeled.

There can be problems of consistency between data types that must be resolved. For example, core permeability data may imply an average permeability of 50 mD and well test data may imply an average of 300 mD. Perhaps there are fractures or other high-permeability features causing the large-scale effective permeability to be higher than the measured core data; it may be necessary to correct or complete the core data before modeling. Alternatively, an inappropriate interpretive model may have been used in the well test analysis program; reinterpretation may be necessary before using that data. Geostatistical tools, at best, identify inconsistent data. They do not help to resolve any discrepancies or inconsistencies; the reconciliation of the different data types must be done by the various professionals involved in the reservoir modeling. Most often, inconsistencies are reduced by interpreting each source of raw data in light of all other data types.

It is relatively easy to construct a numerical reservoir model that honors only some of the available data, for example; (1) a hand- contoured map of average porosity would honor averaged well data and geologic trends but would not represent the details of the well data nor flow properties from well test and historical production data, (2) a statistical model would honor all local well data and certain statistical measures of heterogeneity but may not represent important geologic features that could impact flow predictions, and (3) a large scale "inversion" from production or seismic data would honor the data from which it was generated but would not represent small scale heterogeneities and other possibly important geologic constraints.

The challenge addressed by geostatistical reservoir modeling, and this book, is to simultaneously account for all relevant geological, geophysical, and engineering data of varying degrees of resolution, quality, and certainty. The idea of a creating a "shared earth model" that represents all relevant data is not new [97].

1.5 An Introductory Example

This introductory example is fashioned after a real problem, and the geological data is based on outcrop observations in a reservoir analogue. The setting of the problem is shown in Figure 1.2; a horizontal well is to be drilled from a vertical well to produce a relatively thin oil column. The goal is to construct a numerical model of porosity and permeability to predict the performance of the horizontal well including (1) oil production rates, (2) timing of gas coning, and (3) water coning.

Figure 1.3 shows information available at the vertical well location. From left to right: (1) stratigraphic layering, (2) facies, (3) the well, (4) vertical profile of porosity, and (5) the fluid contacts. These local data are supplemented by additional data such as the top-of-structure grid. There are four integer-coded facies at the well: 0 – shales, all of which are not laterally extensive, 1 – incised valley fill sandstone, 2 – channel fill sandstone, 3 – lower shoreface sandstone. Table 1.2 gives the permeability characteristics of each

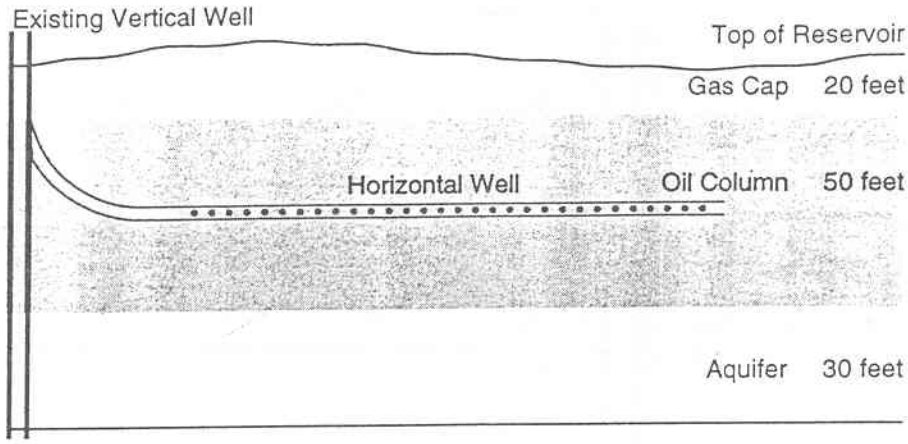


Figure 1.2: Setting of the introductory example: a horizontal well is to be drilled from a vertical well to produce from a thin oil column.

facies type. The coefficient of variation is the average permeability divided by the standard deviation, K_v is the vertical permeability, and K_h is the horizontal permeability. The histograms follow a nearly lognormal distribution, see Figure 1.4.

The sandstone facies have high permeability with correspondingly high variability. Figure 1.4 illustrates the porosity/permeability relationship within each sandstone facies; the permeability histograms are shown to the left. Geostatistical modeling will ensure reproduction of the porosity histograms (not shown), the permeability histograms, and the porosity permeability cross plots.

An important additional constraint is a measure of spatial correlation. Measures of vertical continuity may be derived from a limited number of vertical wells. Establishing the horizontal correlation from limited vertical information and analogue information is an important subject.

Figure 1.5 shows the gridding for flow simulation. For numerical efficiency, the grid is horizontal and aligned with the gas oil and oil water fluid

Table 1.2: Permeability characteristics of each facies in the introductory example.

Code	Facies	Average Perm.(mD)	Coefficient of Variation	K_v:K_h Ratio
0	Coal and shale	1	0.00	0.1
1	Incised valley fill sandstone	1500	1.00	1.0
2	Channel fill sandstone	500	1.50	0.1
3	Lower shoreface sandstone	1000	0.75	0.8

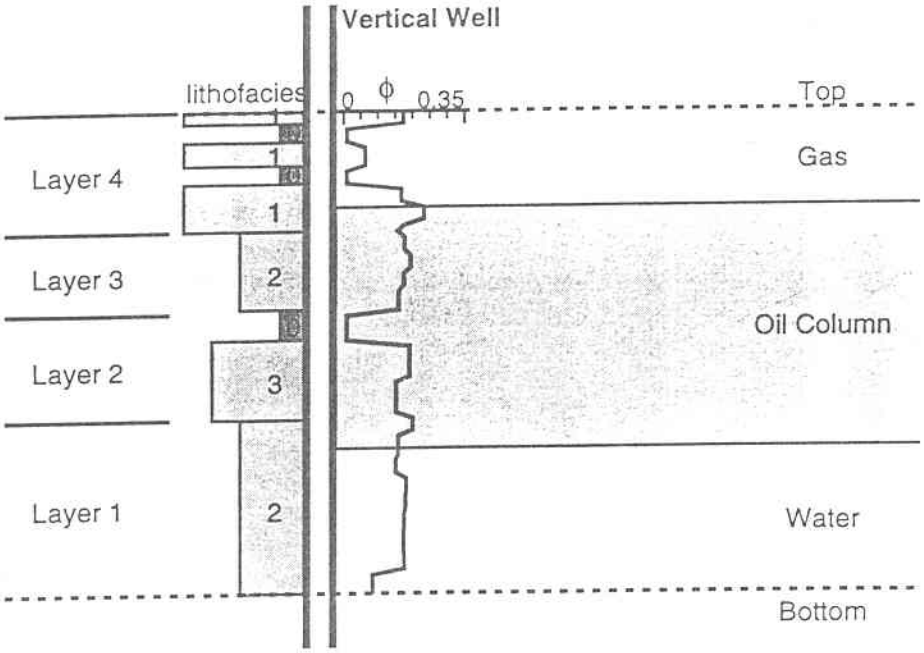


Figure 1.3: Vertical well with geological layering, lithology, porosity, and fluid contacts.

contacts. The black dots on Figure 1.5 illustrate the location of the proposed horizontal well completion. Representative three-phase fluid properties and rock properties such as compressibility have been retained. It would be possible to consider these properties as unknown and build that uncertainty into the model; however, in this introductory example they have been fixed with no uncertainty.

In this example, the two critical inputs to flow simulation provided by geostatistics are numerical models of porosity and permeability. Figure 1.6 shows three simple assignments of porosity using the flow simulation grid (1) a "layercake" model, which simply assigns constant values to each layer, (2) a smooth inverse distance model, and (3) a simple Gaussian-based simulation (geostatistics without geology).

The three models represented in Figure 1.6 were put into the Eclipse [136] flow simulator. The resulting oil production rate, water cut, and gas oil ratio are shown in Figure 1.7. Except for the 2000 cell porosity and permeability values, all other rock and fluid properties were fixed. The smooth model of porosity and permeability (long dashes on Figure 1.7) leads to significantly different results from the layercake and simple geostatistical model.

Figure 1.8(a) shows a geostatistical realization constructed according to a more realistic gridding and facies dependent porosity/permeability (recall the vertical well data in Figure 1.3). The surface grids separating the four

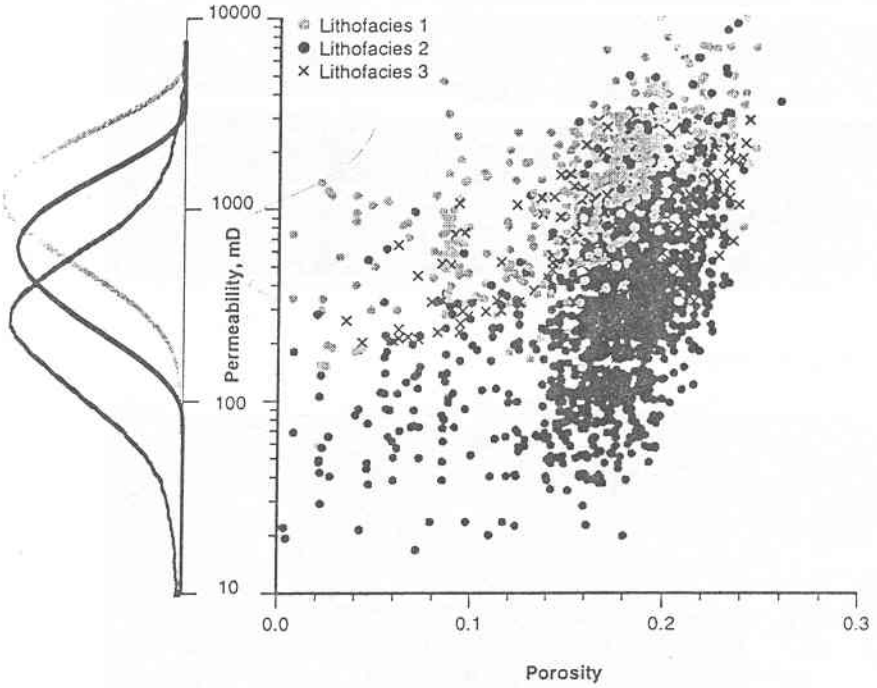


Figure 1.4: Porosity/permeability within the three sandstone facies. The permeability histograms are shown to the left.

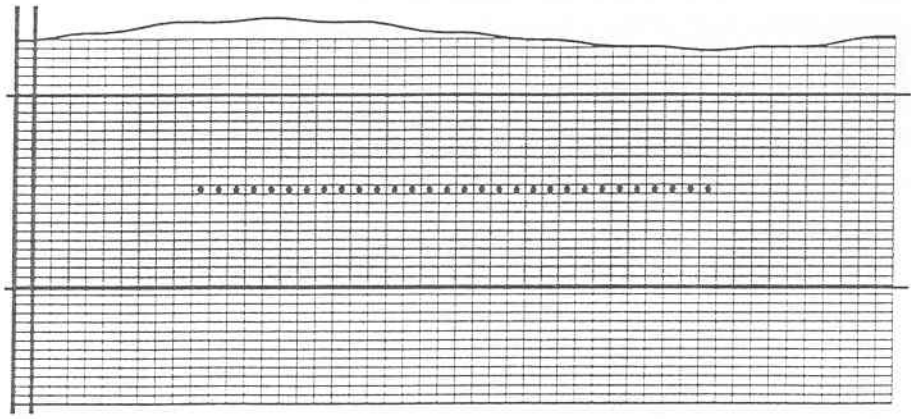


Figure 1.5: Gridding for flow simulation. The 50 (horizontal) by 40 (vertical) grid is aligned with the fluid contacts.

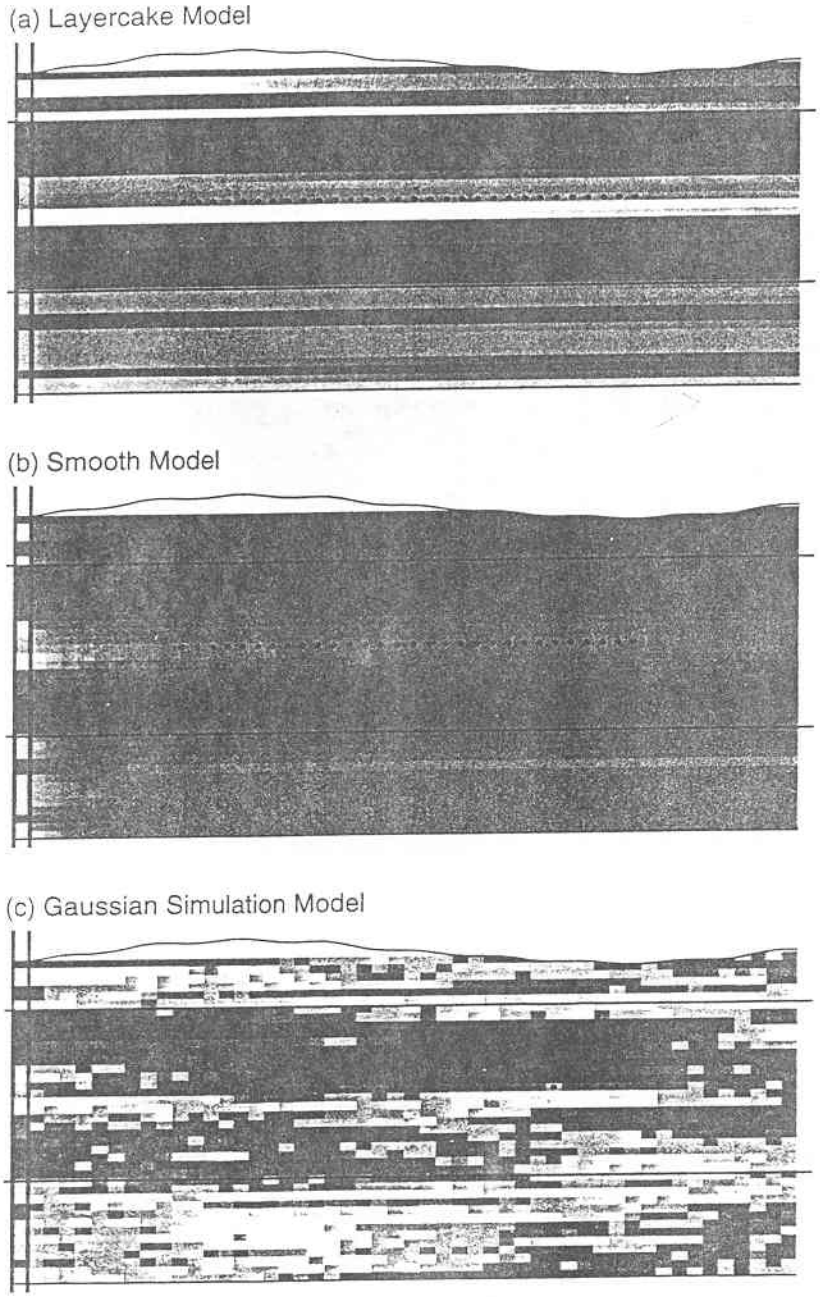


Figure 1.6: Three simple assignments of rock properties (a) a "layercake" or horizontal projection model, (b) a smooth inverse distance model, and (c) a simple Gaussian simulation. The gray scale ranges from 0% porosity (white) to 25% porosity (black). The lack of resolution on the inverse distance model visually reinforces the smoothing effect.

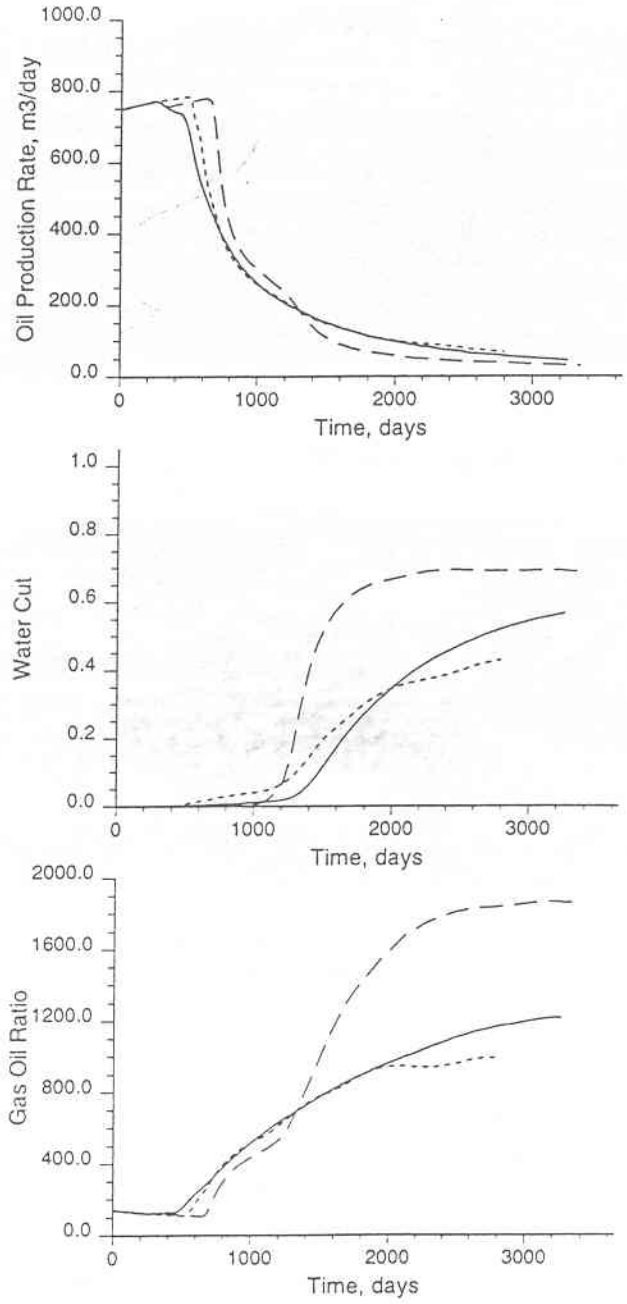
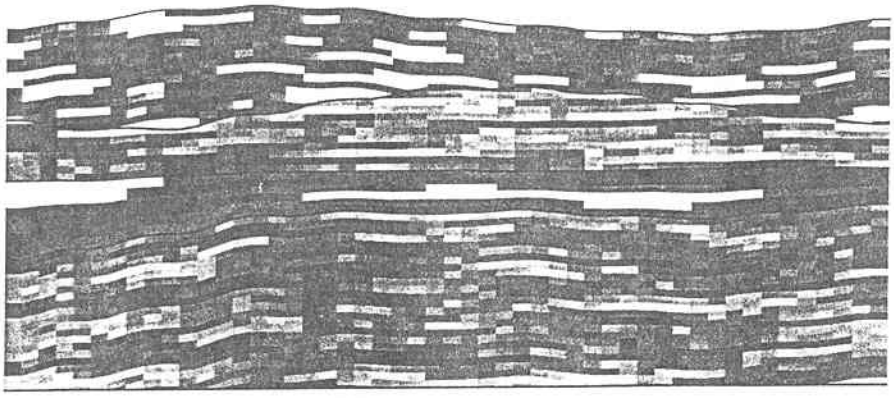


Figure 1.7: Flow results with three simple assignments of rock properties: solid line- layercake model; long dashes - smooth model; short dashes - simple geostatistics model.

(a) Geostatistical Model



(b) Geostatistical Model - Flow Grid

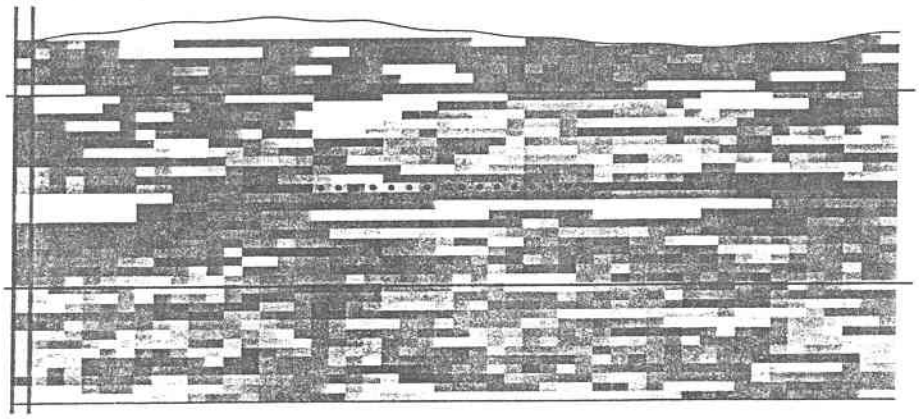


Figure 1.8: The first geostatistical realization shown on (a) the geological grid and (b) the flow simulation grid. As in Figure 1.6, the gray scale ranges from 0% porosity (white) to 25% porosity (black).

layers vary smoothly from the vertical well location, where they are fixed. Figure 1.8(b) shows the geostatistical realization at the 2000 regular grid blocks used for flow simulation. A nearest neighbor scheme was used to regrid the geological model for flow simulation. This regridding is necessary because virtually all finite difference flow simulators do not handle layers that pinch out (note how some cells are truncated near the top of Figure 1.8(a)).

Figure 1.9 shows 20 geostatistical realizations generated with the same geological gridding and porosity/permeability modeling. The vertical well, at the left, is honored in all realizations (note the two thin shales toward the top and the thicker shale near the center). Moreover, all realizations have the

same spatial character, for example, the proportion of high and low porosity is about the same and the horizontal extent of the shales is about the same. The details of the realizations, however, are uncertain as illustrated by the variability of the thick shale near the center of the vertical well; the horizontal extent varies from a few grid blocks to halfway across the cross section.

The flow results from 20 geostatistical realizations are shown in Figure 1.10. Note the significant variability in the results due to local differences in the porosity/permeability models. This uncertainty is further illustrated in Figure 1.11 where the histogram of the 20 cumulative oil productions after 1000 days and the times to water breakthrough are shown together with the earlier results. The geostatistical realizations lead to less cumulative production and earlier water breakthrough. Of course, this is not a general result. In other cases, the predictions of geostatistical modeling may be more favorable. The breakthrough time and recovery depend on the particular nature of heterogeneity.

This example has illustrated many of the steps that will be developed more fully in subsequent chapters. Some important subjects were not, however, presented including (1) scaling from core to geologic modeling cell, (2) the use of seismic data, (3) object-based facies modeling, (4) the integration of production data, and (5) decision making in the face of uncertainty.

1.6 Work Flow

Sequential Approach to Reservoir Modeling

A sequential approach is often followed for reservoir modeling. The large-scale features are modeled first followed by smaller, more uncertain, features. Figure 1.12 shows a high-level flow chart for geostatistical modeling:

1. The first step *Establish Stratigraphic Layering/Coordinates* involves defining the geometry and stratigraphy of the reservoir interval being modeled. This step also involves the development of a conceptual model for the major architecture and continuity of facies, porosity, and permeability within each layer.

2. The facies rock types are modeled by cell-based or object-based techniques within each stratigraphic layer.

3. The porosity is modeled on a by-facies basis before permeability because there are more porosity data (reliable well log data and, often, an abundance of porosity-related large-scale seismic data).

4. The 3-D models of permeability are constrained to the porosity, facies, and layering previously established. At times, there may be a conflict between the facies and porosity already established and well test or production-derived permeability data. In this case, such permeability

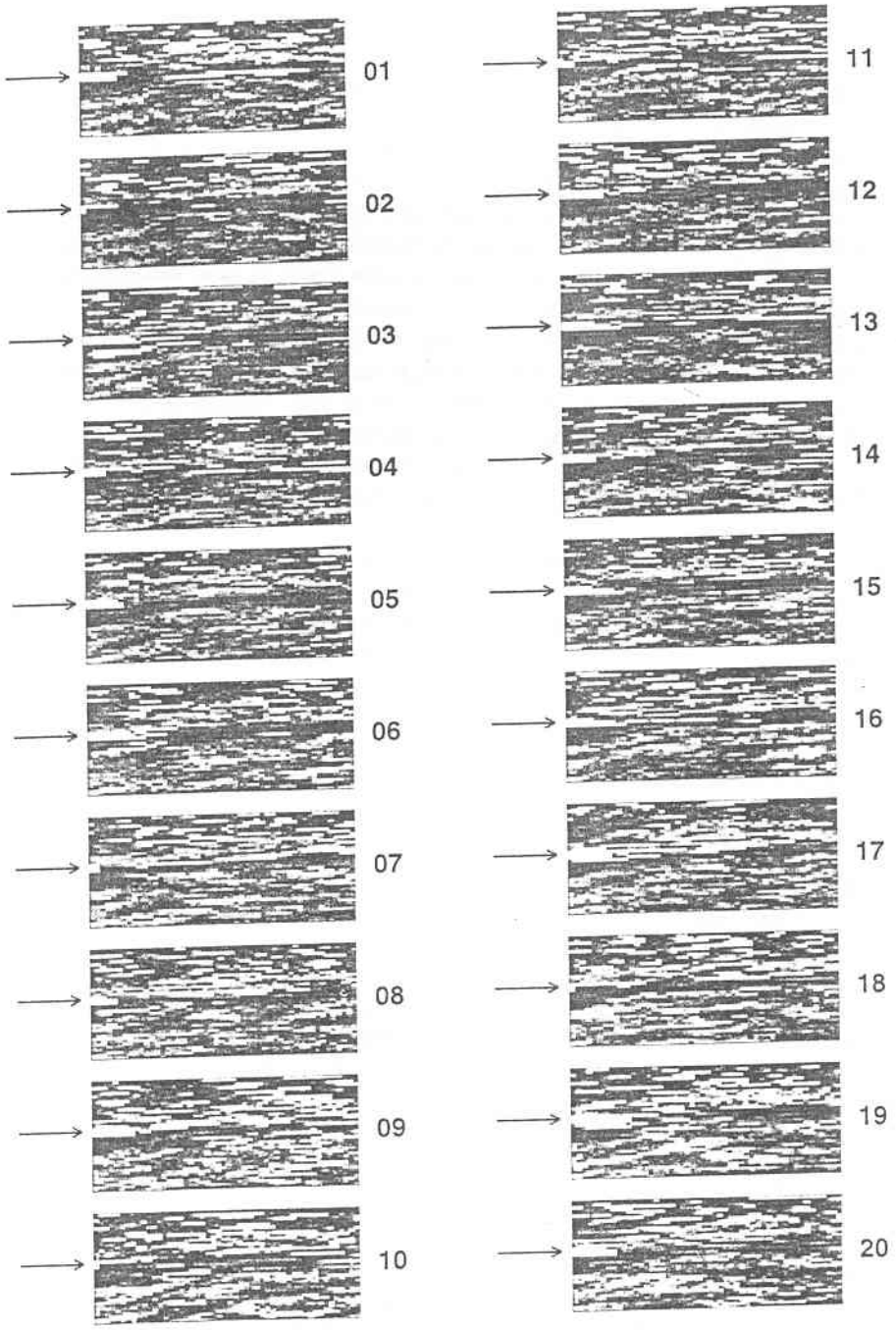


Figure 1.9: Twenty geostatistical realizations constructed using the geological grid and then regridded to the flow simulation grid. The gray arrows point out the location of the thick shale in the vertical well.

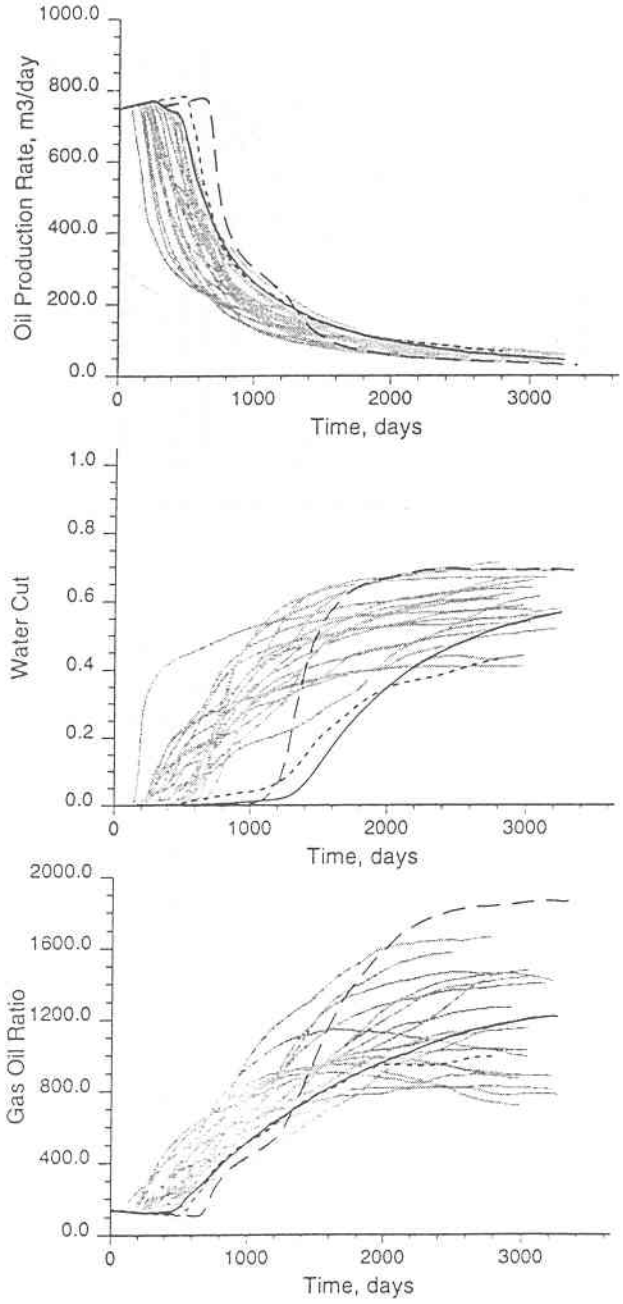
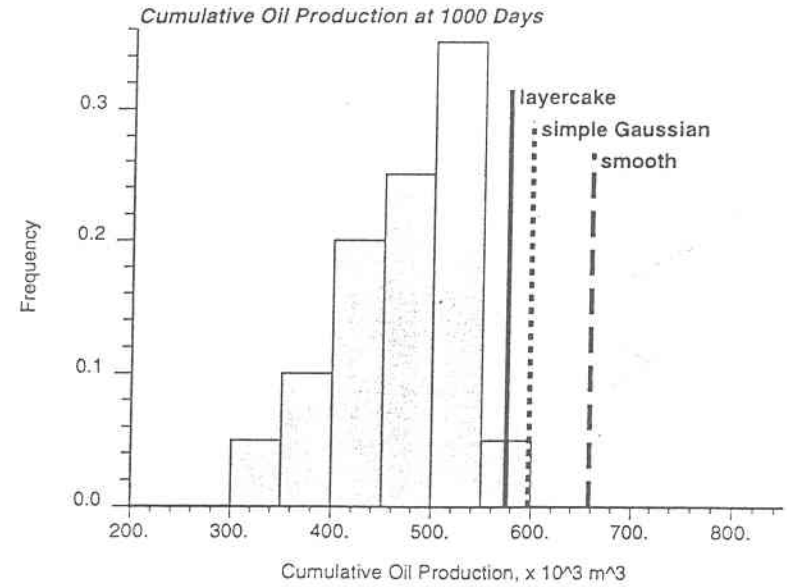


Figure 1.10: Flow results from 20 geostatistical realizations (solid gray lines) with simple model results of Figure 1.7 superimposed.



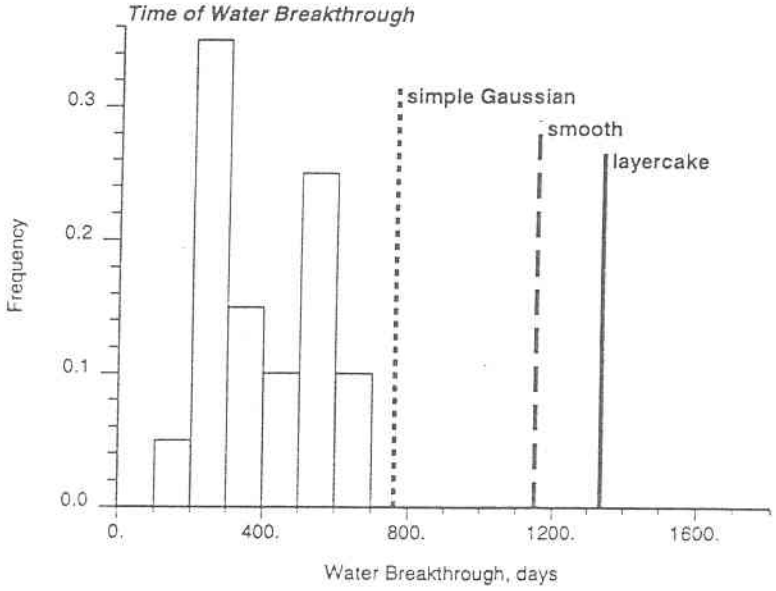


Figure 1.11: The cumulative oil production after 1000 days and the time to water breakthrough. Note the axis on the two plots. There is a significant difference between the simple models (represented by the vertical lines) and the results of geostatistical modeling (the histograms).

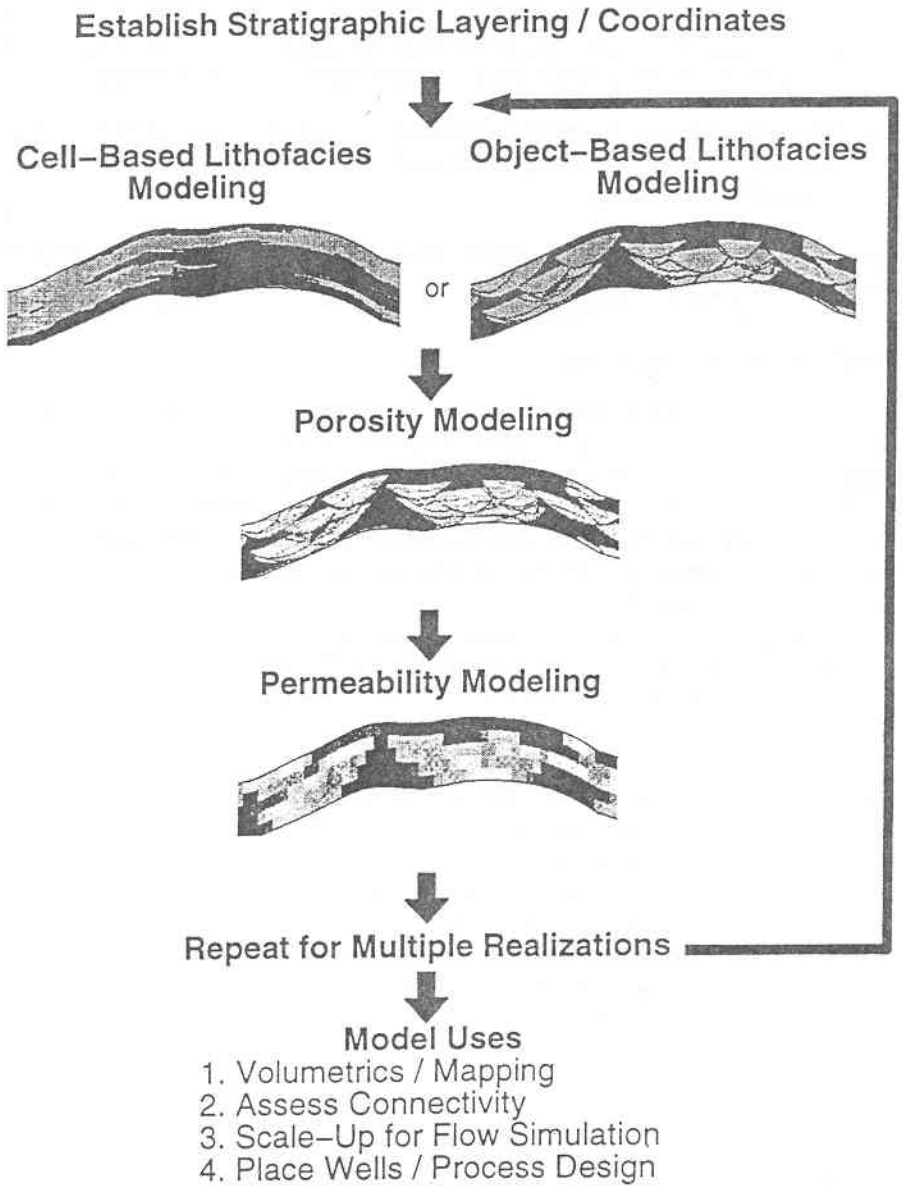


Figure 1.12: Geostatistical modeling flow chart: the structure and stratigraphy of each reservoir layer must be established, the facies modeled within each layer, and then porosity and permeability modeled within each facies.

data should be expressed as constraints on the porosity and permeability models and used in steps 2 and 3 above.

5. Multiple equally likely realizations are created by repeating the entire process. Each realization is "equally likely to be drawn"; however, some realizations are more similar to others, hence their class has higher probability; therefore, the term "equiprobable" is thus avoided.

6. These models are input to a simulator or visualized and used to aid in decision making, see Chapter 10 and other stochastic modeling references [118, 232, 277].

More detailed flow charts for each of these main tasks will be developed at the end of each chapter.

Work Flow Diagrams

Each chapter will close with work flow diagrams. The intent of the work flow diagrams is to help newcomers understand some of the linkages and required inputs to different modeling steps. The goal is neither to document all possible work flows nor to document each minor decision that must be made; the work flow diagrams are conceptual in nature. The diagrams are meant to be a summary of the procedures and material presented in the chapter. Readers should refer to the text in all cases for details about the data, operations, decisions, and interpretation guidelines.

The standard flow chart symbols shown in Figure 1.13 will be used. A brief description of how the symbols will be used is given on the right of Figure 1.13.

Figure 1.14 shows an illustration of some of the steps in reservoir characterization. The topic is diverse and involves many different disciplines. Geostatistical techniques are involved in a number of operations; however, the main focus is numerical algorithms to aid in surface mapping and modeling of rock properties (upper right of Figure 1.14). Any single book on petrophysics, seismic, geology, well test analysis, flow simulation, or geostatistics could not hope to cover "reservoir characterization." The aim here is to describe the geostatistical tools that enter into the critical problem of mapping surface and rock properties on the basis of sparse and imprecise data.

Geostatistical tools are most widely used to construct models of the geometry and rock properties for decision making and, perhaps, input to flow simulation. Figure 1.15 shows a work flow for this purpose. This figure is a re-expression of Figure 1.12 using the work flow symbols that will be used in later chapters.

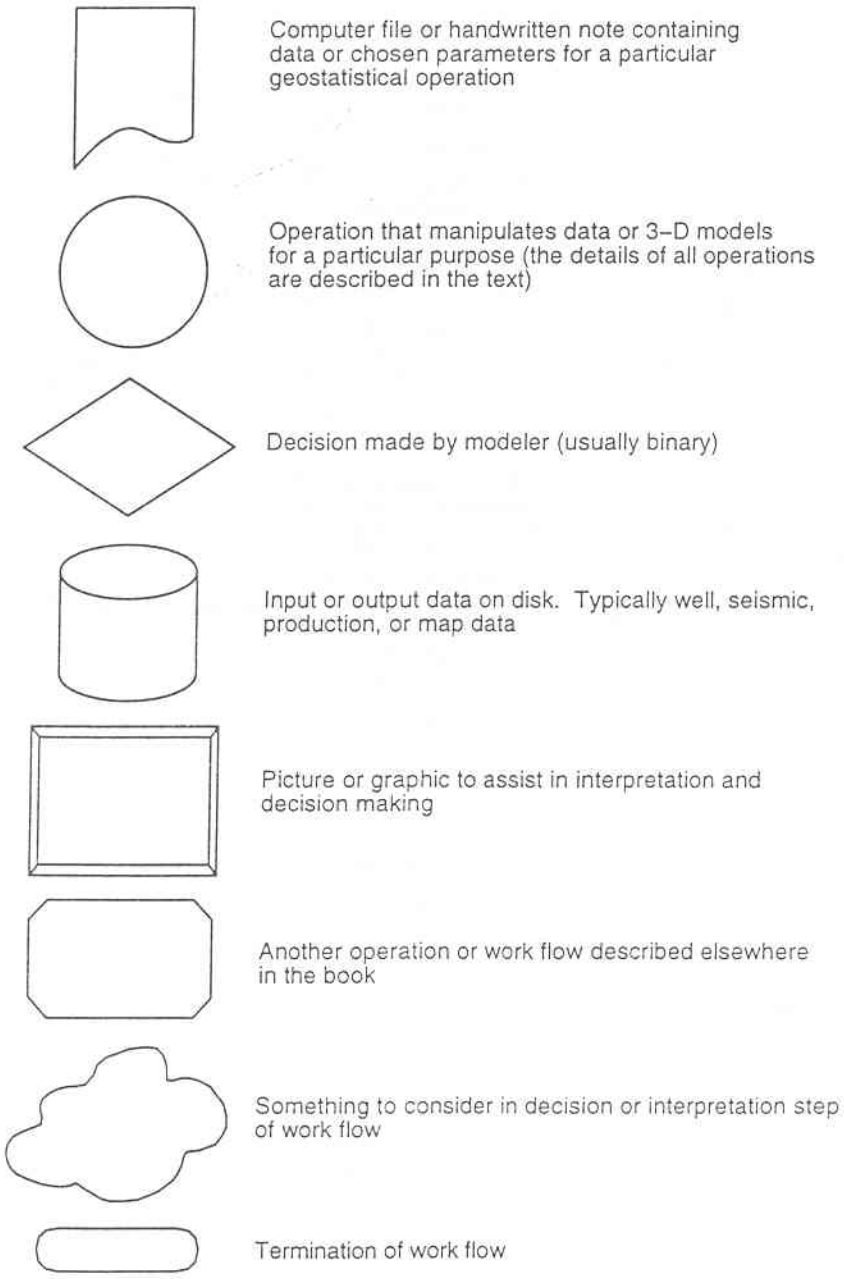


Figure 1.13: Work flow symbols used in subsequent chapters to clarify the operations and considerations involved in geostatistical reservoir modeling.

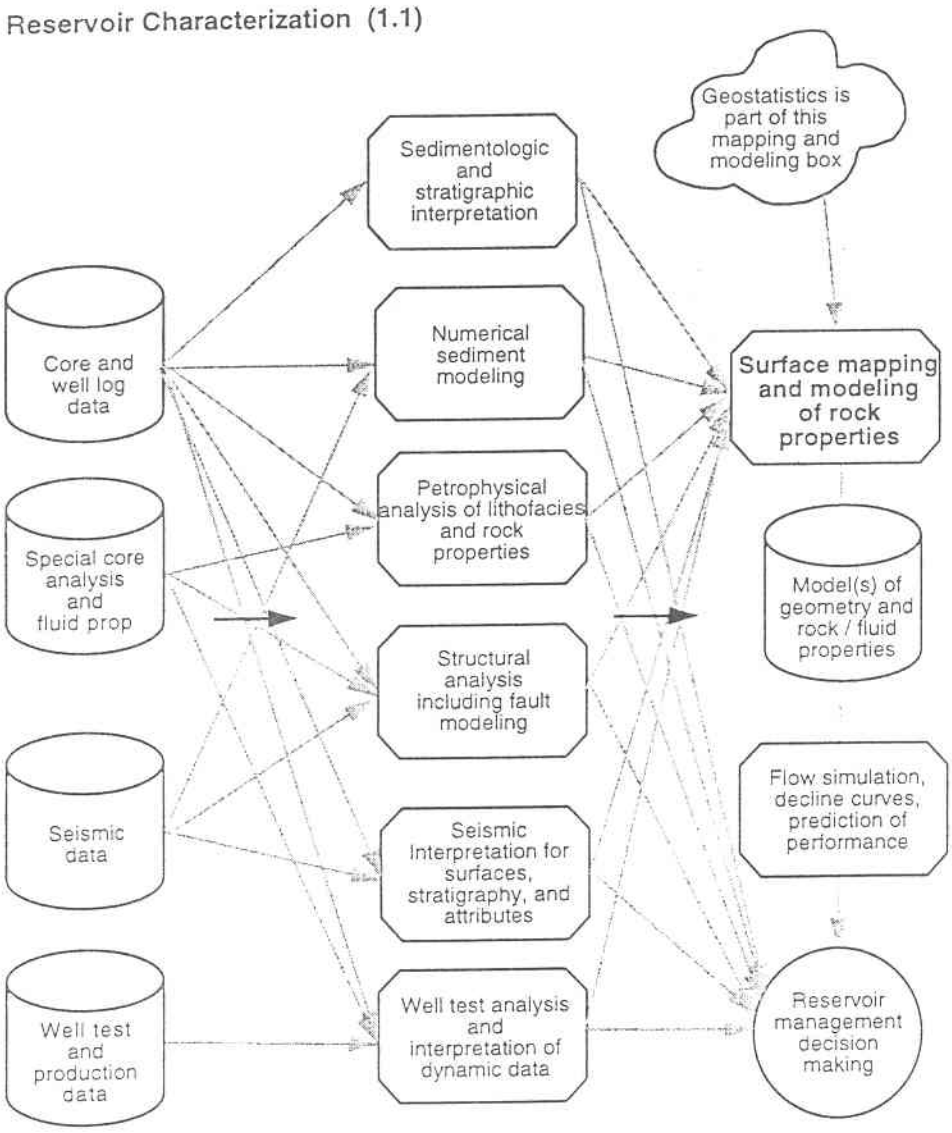


Figure 1.14: An illustration of *some* of the major operations that enter into reservoir characterization. In this large picture "Geostatistical Reservoir Modeling" constitutes part of the work flow to build numerical models of reservoir geometry and internal rock property heterogeneity.

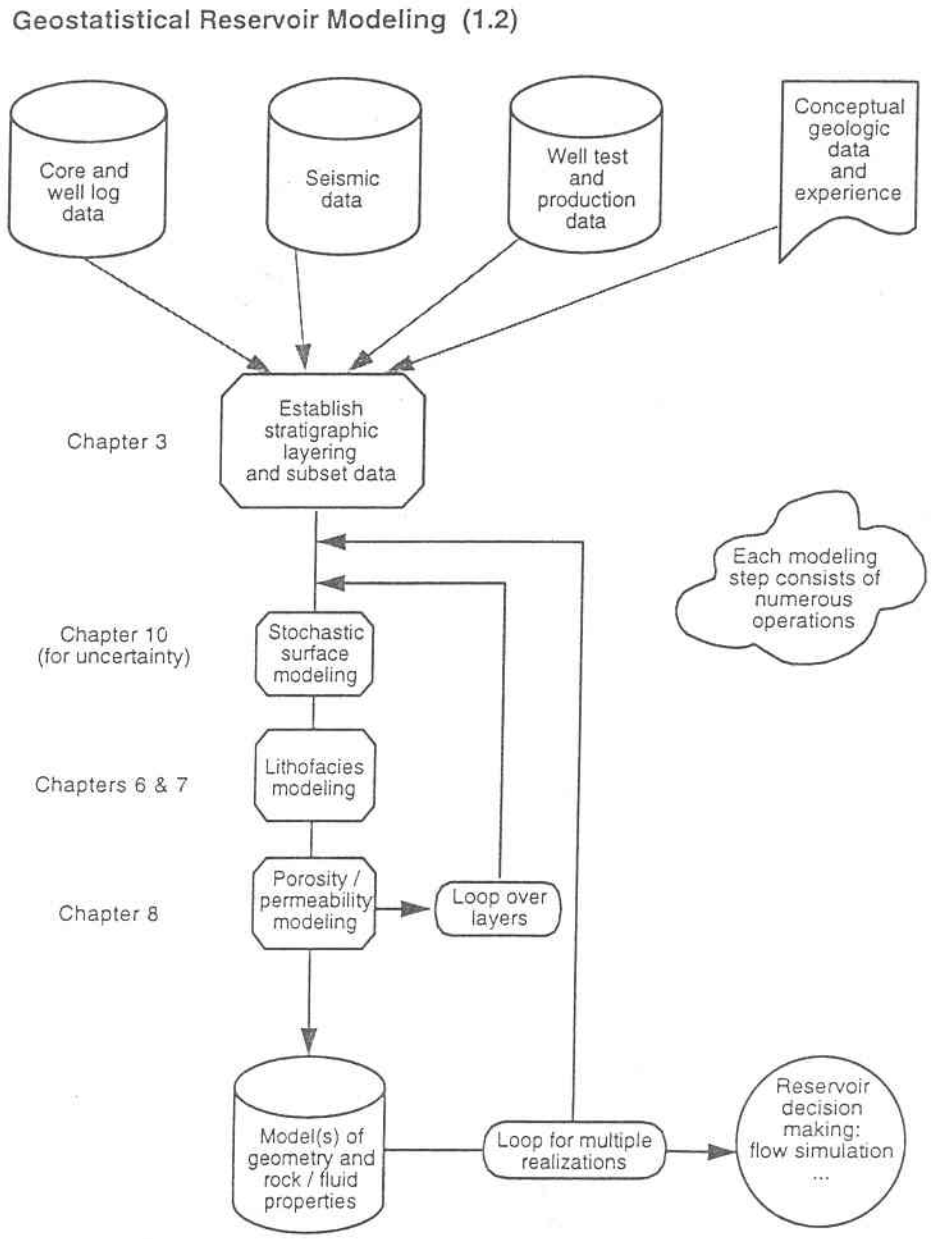


Figure 1.15: A high-level view of geostatistical reservoir modeling. Chapters 3, 6, 7, and 8 present critical operations in detail. Chapters 2 and 5 present preliminary concepts. Chapters 9, 10, and 11 present special topics related to data integration and uncertainty modeling.

Chapter 2

Preliminary Statistical Concepts

The field of statistics is concerned with quantitative methods for collecting, organizing, summarizing, and analyzing data, as well as for drawing conclusions and making reasonable decisions on the basis of such analysis. The field of statistics is extensive; the goal here is a pragmatic presentation of the concepts essential to geostatistical reservoir modeling. Geostatistics is distinct from statistics in three main respects: (1) focus on the geological origin of the data, (2) explicit modeling and treatment of spatial correlation between data, and (3) treatment of data at different volume scales and levels of precision.

Section 2.1 addresses the problem of defining geological and statistical populations for modeling. The concept of *stationarity* is discussed with the implications on subsequent modeling.

The basic notation and definition of probability distributions are given in Sections 2.2 and 2.3. Often, data must be transformed from one distribution to another. Section 2.4 describes a widely applicable technique for transformation that preserves the ordering and spatial relationships of the data. This section also introduces Q-Q plots and the concept of "despiking."

The important techniques of declustering and histogram smoothing are presented in Sections 2.5 and 2.6. These techniques allow going beyond the sample available and inferring a distribution and summary statistics more representative of the entire volume of interest.

Section 2.7 introduces Monte Carlo simulation methods for stochastic simulation of one variable. A variant of Monte Carlo methods, known as the bootstrap, is a useful tool in assessing uncertainty in input statistics under a strong assumption of independence between the constituent data. Such Monte Carlo techniques form the foundation of later 3-D geostatistical simulation methods.

Finally, Section 2.8 presents work flow diagrams for some operations related to straightforward statistical concepts.

2.1 Geological Populations and Stationarity

A critical decision in any statistical study is how to pool the data into populations for further analysis. This decision depends on the goals of the study, the specific data available, and the geological setting of that data; there is no universal answer. Often, a reservoir model is built on a by-facies basis one layer at a time. In this case, the data for each facies in each layer should be kept separate.

Populations must be defined because all sample statistics, including a histogram and a mean value, refer to a population and not to any sample in particular. Grouping everything together into one population may mask important trends. Excessive subdivision of the data may result in unreliable population statistics due to not enough data. With limited well data, it may be appropriate to combine the data from all of the layers, that is, separate the data on the basis of facies alone. At the time of modeling, the layers will be kept separate to ensure that the geological layering and layer-specific facies proportions are reproduced.

Facies

The main consideration in choosing the number of facies is to balance the level of geological input (tend to consider more) with the statistical and engineering significance of each facies classification (tend to consider less). Ideally, each chosen facies should be geologically significant and yet have enough data to allow reliable inference of the required statistics for reservoir modeling. In practice, it is difficult to support more than four different facies from the data.

The important petrophysical properties used in subsequent modeling are porosity, permeability, and multiphase flow properties such as relative permeability and capillary pressure. The facies must have (1) clearly different petrophysical properties, or (2) spatial features that make them easy to model. There is no benefit to separate the data according to facies that do not lead to distinguishable flow properties.

Typically, there are only a few wells with cores; most will only have a suite of well logs. The facies are most easily identified by studying core. Once identified from core, it is essential that the facies can also be separated (with reasonable accuracy) from well logs at the wells without core. This is done by some type of calibration procedure (discriminant analysis, neural networks, etc.). If the facies cannot be identified from well logs alone then "facies associations," or a broader definition of facies, would have to be considered prior to geostatistical modeling.

Stationarity

A more formal definition of stationarity will be presented later in Section 2.3; however, for all practical purposes, the decision of which data should be

pooled together for subsequent analysis *is* the decision of stationarity. There is always some temptation to consider each sample as unique and, therefore, not to lump it with other samples. This temptation must be avoided to allow enough data for geostatistical inference.

The decision of stationarity may be revisited once data analysis and geostatistical modeling has started. For example, we may notice a bimodal (two peaks) histogram of porosity within one of our chosen facies classifications. This does not mean the data are "non-stationary." It does mean that we should go back to the data and consider separating the data into two classes provided these display distinct "statistical" and "geological" properties.

Sometimes trends, such as a systematic fining upward trend or degradation in reservoir quality in the seaward direction, are sometimes referred to as non-stationarities. The presence of a trend or gradual areal/vertical change in facies proportions, porosity, or permeability need *not* be handled by further subdivision of the data. After all, there is no natural subdivisions in the case of a gradual, continuous, change in properties. This situation can be handled by modifications to the modeling technique, for example, simple kriging with locally varying means or considering that the average or mean property follows a specified trend, see later chapters.

2.2 Notation and Definitions

A *variable* is a measure, such as ϕ, K, i, which can assume any of a prescribed set of values. A variable that can assume a continuum of values between any two given values is called a *continuous variable*; otherwise it is a *discrete* or *categorical* variable. Porosity and permeability are continuous variables often denoted generically by the letter z. Facies classifications are categorical and commonly denoted with the indicator variable i_k where i_k is 1 if the category k is present and 0 if not.

The basic approach of predictive statistics is to characterize any unsampled (unknown) value z as a random variable (RV) Z. The probability distribution of Z models the uncertainty about the unknown true value. The random variable is traditionally denoted by a capital letter, say Z, while its outcome values are denoted with the corresponding lowercase letter, say z. The RV model Z, and more specifically its probability distribution, is usually location dependent; hence the notation $Z(\mathbf{u})$, with $\mathbf{u}$ being the location coordinates vector. The RV $Z(\mathbf{u})$ seen as a function of $\mathbf{u}$ is called a random function (RF). The RV $Z(\mathbf{u})$ is also information dependent in the sense that its probability distribution changes as more data about the unsampled value $z(\mathbf{u})$ become available.

Histograms and Probability Distributions

The histogram, a bar chart of the frequency of sample values observed, is a familiar statistical display. Figure 2.1 shows a typical histogram created with

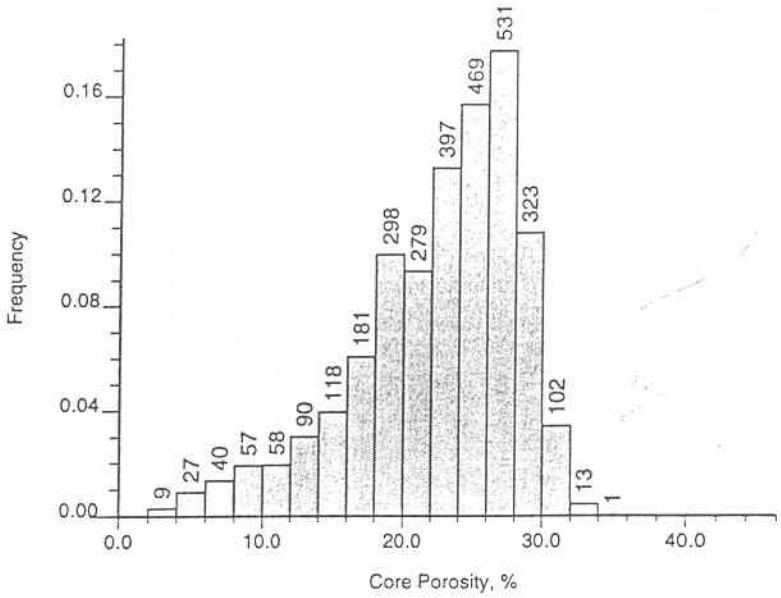


Figure 2.1: Histogram of 2993 core porosity measurements. The number of data in each bin is shown above each histogram bar and the relative frequency is shown on the vertical axis.

the GSLIB program `histplt`. This histogram has a fixed bin width. The frequency in each bin is the number of sample values falling into that bin divided by the total number of samples (2993 in this case). The original data values may be assigned unequal weights to correct for preferential clustering; see next section. A too-small bin width would create a noisy histogram with many short scale variations. A too-large bin width creates a too smooth histogram that may mask important features.

Histograms are suitable for gaining a visual appreciation for the distribution of data values; however, cumulative histograms are more extensively used in calculations. The cumulative distribution function (cdf) of the random variable Z is denoted:

$$F(z) = \text{Prob}\{Z \leq z\} \quad \in [0,1] \tag{2.1}$$

Where the cdf $F(z)$ is a non-decreasing function of z. A plot of $F(z)$ versus z, see Figure 2.2, does not require the data to be classified into bins. Note that $F(z)$ is valued between 0 and 1 and z is valued in the range of data values. We calculate $F(z)$ as the experimental proportion of data values no greater than a threshold z.

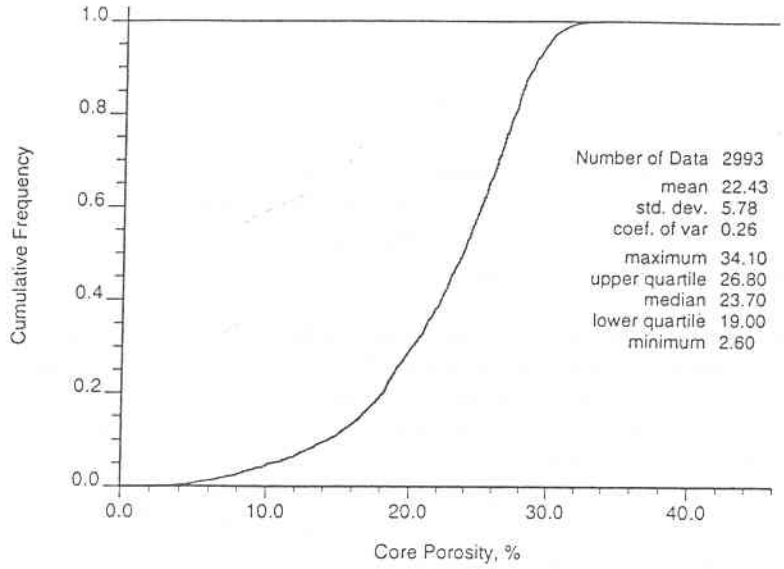


Figure 2.2: Cumulative distribution of 2993 core porosity measurements.

Probability Distributions

The *probability density function* (pdf) $f(z)$ is linked to the cdf $F(z)$ provided that $F(z)$ is derivable:

$$f(z) = F'(z) = \lim_{dz \longrightarrow 0} \frac{F(z+dz) - F(z)}{dz} \quad \text{if it exists} \tag{2.2}$$

where the pdf $f(z)$ is greater than or equal to zero ($f(z) \geq 0, \forall z$), the integral of $f(z)$ is the cumulative distribution function or cdf ($\int_{-\infty}^{x} f(u)du = F(x) = \text{Prob}\{X \leq x\}$), and the integral of $f(z)$ from $-\infty$ to ∞ is one ($\int_{-\infty}^{\infty} f(z)dz = 1$). As mentioned above, the cdf $F(z)$ is more extensively used in calculations although people would rather visualize a histogram or probability density function.

The probability density function or pdf $f(z)$ is related to the histogram; often, a parametric probability density function is overlaid on an experimental histogram. There is, however, an important conceptual difference between the discrete histogram and the continuous pdf curve. The frequency of samples in a histogram class may be interpreted as the probability of a sample value to fall in that class; however, $f(z)$ should not be interpreted as the "probability of z." It is the integral of $f(z)$ between a lower bound z_{lower} and upper bound z_{upper}, $\int_{z_{lower}}^{z_{upper}} f(z)dz$, where $z_{upper} > z_{lower}$, which can be interpreted as a discrete probability for z to be in the class $(z_{lower}, z_{upper}]$.

Parametric Distributions

A *parametric* distribution model is an analytical expression for the pdf or cdf, for example, for the normal or Gaussian density function:

$$f(z) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}(z-m)^2/\sigma^2}$$

with mean m and standard deviation σ that control the center and spread of the bell-shaped normal distribution. Parametric models sometimes relate to an underlying theory, for example, the normal distribution is the limit distribution for a number of theorems collectively known as the central limit theorem.

There is no general theory for earth science related variables that would predict the parametric form for probability distributions. Nevertheless, certain distribution shapes are commonly observed. A "lognormal" distribution is one where the logarithm of the variable follows a normal distribution. Often, however, any skewed distribution of a positive variable such as permeability is wrongly called *lognormal*. There are statistical tests to judge whether a set of data values follow a particular parametric distribution. These tests are of little value in reservoir characterization because they require that the data values all be independent one from another, in practice, taken far away from each other. Probability plots are useful for checking closeness to a normal or lognormal distribution and for identifying outlier data.

Probability plots are cumulative probability plots with the vertical probability axis scaled such that a normal distribution would plot as a straight line. Figure 2.3 shows a probability plot for the 2993 data we have been considering. Note the exaggerated probability scale for the high and low values. Such probability plots were originally used to see how well data fit a normal or lognormal probability distribution. The use of the histogram transformation procedures presented below alleviates the need for the data to follow any particular parametric distribution model; the main application of probability plots is to see all of the data values on a single plot. There are useful interpretations that can be made, for example, a sharp change of slope may indicate a different geologic population. Since we are often interested in the high and low values, the exaggeration of extreme values makes the normal probability scaling more useful than arithmetic probability scaling. The X-variable axis may be plotted with logarithmic scale; data falling on a straight line would then indicate a lognormal distribution. Program `probplt` in GSLIB creates a probability plot of the type shown in Figure 2.3.

Parametric distributions have three significant advantages: (1) they are amenable to mathematical calculations, (2) the pdf $f(z)$ and cdf $F(z)$ are analytically known for all z values, and (3) they are defined with a few parameters, which are chosen to be easy to infer.[1] The primary disadvantage

[1] A general rule is that a minimum of 10 data are needed for a particular "statistic," for example, 10 data would be needed to compute a mean with any reasonable level of

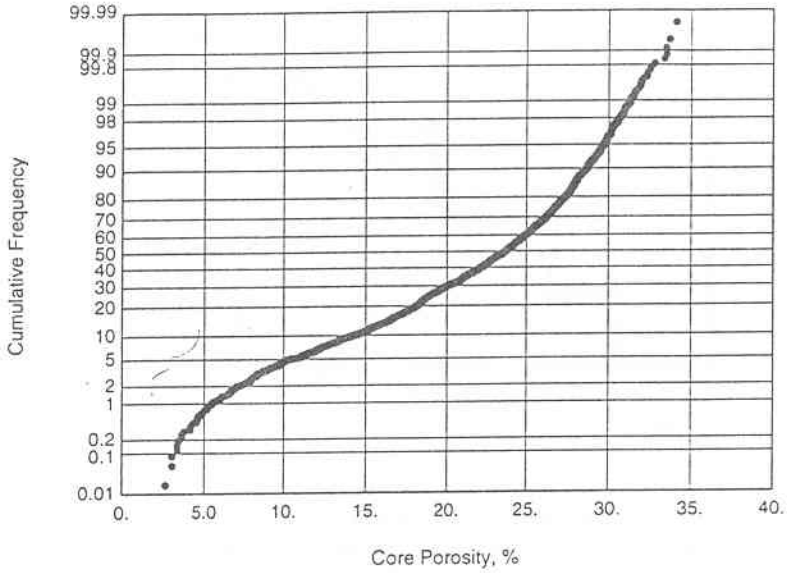


Figure 2.3: Probability plot of 2993 core porosity measurements.

of parametric distributions is that, in general, real data do not conveniently fit a parametric model. *Non-parametric* distributions (described below) are more flexible in capturing the behavior of real data. *Data transformation* (described below in Section 2.4) permits data following any distribution to be transformed to any other distribution, which permits us to capitalize on most of the benefits of parametric distributions.

Non-Parametric Distributions

In presence of sufficient data (say, more than 200), the data are often not well represented by a parametric distribution model. In this case, the cdf probability distribution may be inferred directly from the data, that is, *non-parametrically*. The cumulative distribution function $F(z)$ is inferred directly as the proportion of data less than or equal to the threshold value z:

$$F(z) = \text{Prob}\{Z \leq z\} \approx \frac{\text{Number of data less than or equal to } z}{\text{Number of data}} \tag{2.3}$$

Thus, a proportion is associated to a probability. The proportion in (2.3) assumes the data are equally representative, which may be a poor assumption in presence of data that are preferentially clustered in specific areas. For that reason, we will develop *declustering* procedures in Section 2.5 (below) to

precision. Two hundred data would be needed to compute a 90% probability interval (10 data below the 5% lower limit and 10 above the 95% upper limit). This is, of course, overly simplistic, but useful nonetheless.

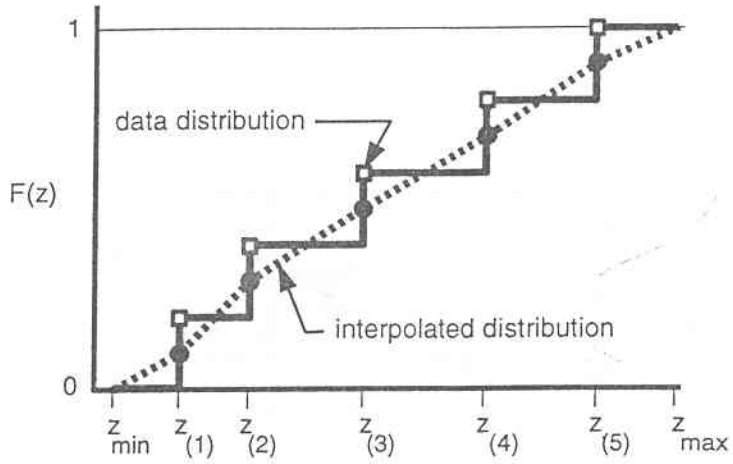


Figure 2.4: A non-parametric distribution inferred from sample data: the solid line is from the ordered data $z_{(1)}, \ldots, z_{(n)}$ and the definition of a probability (2.3). The dashed line is a linear interpolation of the cdf to provide a continuous relation between $F(z)$ and z.

calculate a relative weight $\{w_i, i = 1, \ldots, n\}$ to be assigned to each of the n data $\{z_i, i = 1, \ldots, n\}$. Lesser weight is assigned to redundant data. The sum of the weights is one: $\sum_{i=1}^{n} w_i = 1.0$. The cumulative distribution function $F(z)$ is inferred directly using the weights:

$$F(z) \approx \text{sum of weights to data with } z_i \leq z \tag{2.4}$$

A non-parametric cumulative distribution function $F(z)$ inferred with (2.4) is a series of step functions (see solid line on Figure 2.4). The subscript $z_{(i)}$ indicates the order in the data set; 1 being the smallest and n being the largest. Five equal-weighted data are illustrated in Figure 2.4). Some form of interpolation may be used to provide a more continuous distribution $F(z)$ that extends to arbitrary minimum z_{min} and maximum z_{max} values. Linear interpolation is often used. More complex interpolation models (see Chapter 5 in GSLIB [62]) could be considered for highly skewed data distributions with limited data.

Quantiles and Probability Intervals

A *quantile* is a z-value that corresponds to a fixed cumulative frequency. For example, the 0.5 quantile, also called the median and denoted $q(0.5)$, is the z-value that separates the data into two equal halves. $q(0.25)$ is the lower quartile. $q(0.75)$ is the upper quartile. The quantile function $q(p)$ is the inverse of the cumulative distribution function:

$$q(p) = F^{-1}(p) \text{ such that } F(q(p)) = p \ \in \ [0, 1]$$

The quantile function $q(p)$ contains the same information as the cdf $F(z)$; it is just another way of looking at a probability distribution.

A symmetric *p-probability interval* is defined by the following lower and upper quantiles

$$q\left(\frac{1-p}{2}\right) \quad \text{and} \quad q\left(\frac{1+p}{2}\right)$$

For example, the 90% symmetric probability interval is bounded by $q(0.05)$ and $q(0.95)$. Any probability interval can be determined once the cdf $F(z)$ or the quantile function $q(p)$ have been established, parametrically or non-parametrically.

Distributions of Discrete Variables

The probability distribution of a discrete or categorical variable is defined by the probability or proportion of each category, that is, $p_k, k = 1, \ldots, K$, where there are K categories. The probabilities must be positive $p_k \geq 0, \forall k = 1, \ldots, K$ and sum to 1.0: $\sum_{k=1}^{K} p_k = 1.0$. A table of the p_k values summarize the data distribution adequately. At times, however, we can consider a histogram and cumulative histogram format, see Figure 2.5.

The cumulative histogram in Figure 2.5 is a series of step functions for an arbitrary ordering of the discrete categories. Such a cumulative histogram is not useful for descriptive purposes but is needed for many calculations such as Monte Carlo simulation and data transformation, see below. In general, the ordering is not important. We will discuss cases when the ordering affects the results.

Continuous Variable Statistics

Continuous data are often summarized by a *central* value such as the average or *mean*:

$$\hat{m} = \frac{\sum_{j=1}^{N} w(\mathbf{u}_j) z(\mathbf{u}_j)}{\sum_{j=1}^{N} w(\mathbf{u}_j)} \tag{2.5}$$

where there are N data values $z(\mathbf{u}_j), j = 1, \ldots, N$ possibly weighted differently by $w(\mathbf{u}_j), j = 1, \ldots, N$. The notation $\hat{m}$ refers to the average or sample mean of some limited number of observations. The notation m is used for the true mean of a statistical model.

The spread of the data is often measured by the average squared difference from the mean:

$$\hat{Var}\{Z\} = \hat{\sigma}^2 = \sum_{j=1}^{N} w(\mathbf{u}_j)\left(z(\mathbf{u}_j) - \hat{m}\right)^2 \tag{2.6}$$

The notation $\hat{\sigma}^2$ refers to an observed variance and the notation σ^2 is used for the true variance of a statistical model. The square root of the variance is the standard deviation (σ).

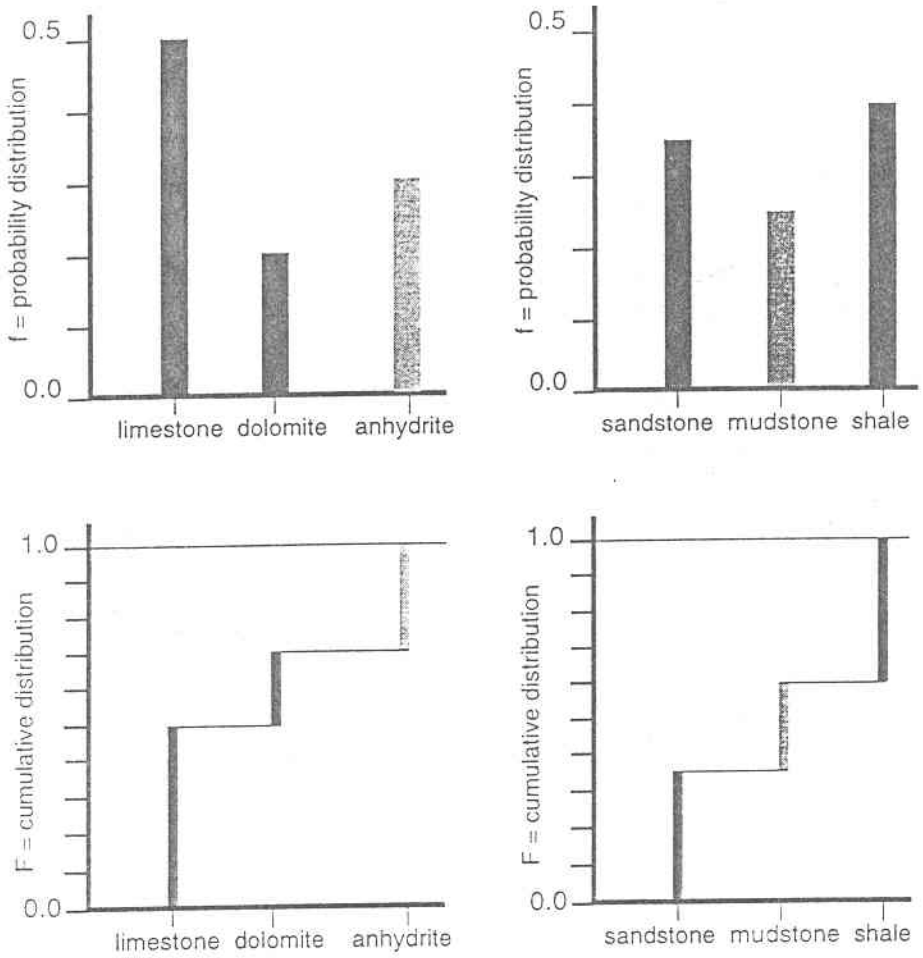


Figure 2.5: Two illustrations of the probability distribution and cumulative distribution of a categorical variable.

The coefficient of variation (CV) is the ratio of the standard deviation and the mean, that is, $CV = \sigma/m$. The CV is only defined for variables that are strictly positive. The CV is a useful measure of spread since it is unit free. The magnitude of the CV, however, does not tell us how difficult it will be to model a variable; the spatial variability must be considered. Nevertheless, it is generally understood that a CV of more than 2.0 indicates that a variable is highly *heterogeneous*. The CV could be high due to mixing of data from different facies with "homogenous" values in the facies. This may cause us to reconsider our decision of stationarity, since we should probably not mix homogeneous facies together.

There are other less commonly used measures of central tendency such as the median (0.5 quantile), mode (most common observation), and geometric mean $\left(\left[\prod_{j=1}^{N} z_j\right]^{\frac{1}{N}}\right)$. Less commonly used measures of spread include the range (difference between the largest and smallest observation), mean absolute deviation $\sum_{j=1}^{N} w(\mathbf{u}_j)|z(\mathbf{u}_j) - \hat{m}|$, and interquartile range (0.75 quantile minus 0.25 quantile).

Other high-order summary statistics intended to measure features such as shape, skewness, or modality are often rarely if ever used in geostatistics, they are also very sensitive to outlier and/or limited data.

Categorical Variable Statistics

Consider K mutually exclusive categories $s_k, k = 1, \ldots, K$. This list is also exhaustive; that is, any location $\mathbf{u}$ belongs to one and only one of these K categories. Let $i(\mathbf{u}; s_k)$ be the indicator variable corresponding to category s_k, set to 1 if location $\mathbf{u} \in s_k$, zero otherwise, that is:

$$i(\mathbf{u}_j; s_k) = \begin{cases} 1, & \text{if location } \mathbf{u}_j \text{ in category } s_k \\ 0, & \text{otherwise} \end{cases} \tag{2.7}$$

Mutual exclusion and exhaustivity entail the following relations:

$$\begin{aligned} i(\mathbf{u}; s_k) \cdot i(\mathbf{u}; s_{k'}) &= 0, \ \forall\, k \neq k' \\ \sum_{k=1}^{K} i(\mathbf{u}; s_k) &= 1 \end{aligned} \tag{2.8}$$

The mean indicator for each category $s_k, k = 1, \ldots, K$ is interpreted as the proportion of data in that category, for example estimated with declustering weights such as:

$$\hat{p}_k = \sum_{j=1}^{N} w_j i(\mathbf{u}_j; s_k) \tag{2.9}$$

The variance of the indicator for each category $s_k, k = 1, \ldots, K$ is a simple function of the mean indicator:

$$\begin{aligned} Var\left\{i(\mathbf{u}; s_k)\right\} &= \sum_{j=1}^{N} w_j \left[i(\mathbf{u}_j; s_k) - p_k\right]^2 \\ &= p_k\left(1.0 - p_k\right) \end{aligned} \tag{2.10}$$

2.3 Bivariate Distributions

Figure 2.6 shows a cross plot of permeability versus porosity. The vertical lines divide the data into six classes, each with an equal number of porosity

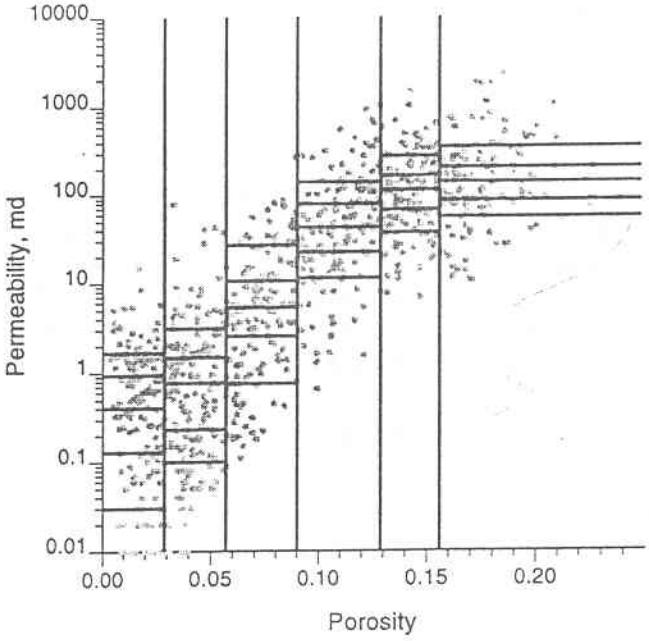


Figure 2.6: Cross plot of permeability versus porosity. The vertical lines divide the data into six classes with an equal number of data points. The horizontal lines within each porosity class divide the data into six subclasses with an equal number of data. The result is 36 classes with equal number of data.

values. The horizontal lines within each porosity class divide the data into six subclasses with an equal number of permeability values. The result is 36 classes with equal number of data.

The correlation coefficient, ρ, measures the degree of correlation between paired values $z(\mathbf{u}_j), y(\mathbf{u}_j), j = 1, \ldots, N$. It is calculated from the variance of each variable σ_z^2 and σ_y^2 (see equation 2.6) and the covariance between the two variables. The sample covariance is defined as:

$$\hat{c}_{z,y} = \sum_{j=1}^{N} w(\mathbf{u}_j)\,(z(\mathbf{u}_j) - \hat{m}_z) \cdot (y(\mathbf{u}_j) - \hat{m}_y) \tag{2.11}$$

where $\hat{m}_z$ and $\hat{m}_y$ are the sample means of the z and y data, see Equation 2.5. The sample linear correlation coefficient is then defined as:

$$\rho = \frac{c_{z,y}}{\sqrt{s_z^2 \cdot s_y^2}} \tag{2.12}$$

ρ is valued between -1 and 1, where -1 entails perfect negative correlation, 0 entails no correlation, and 1 entails perfect positive correlation.

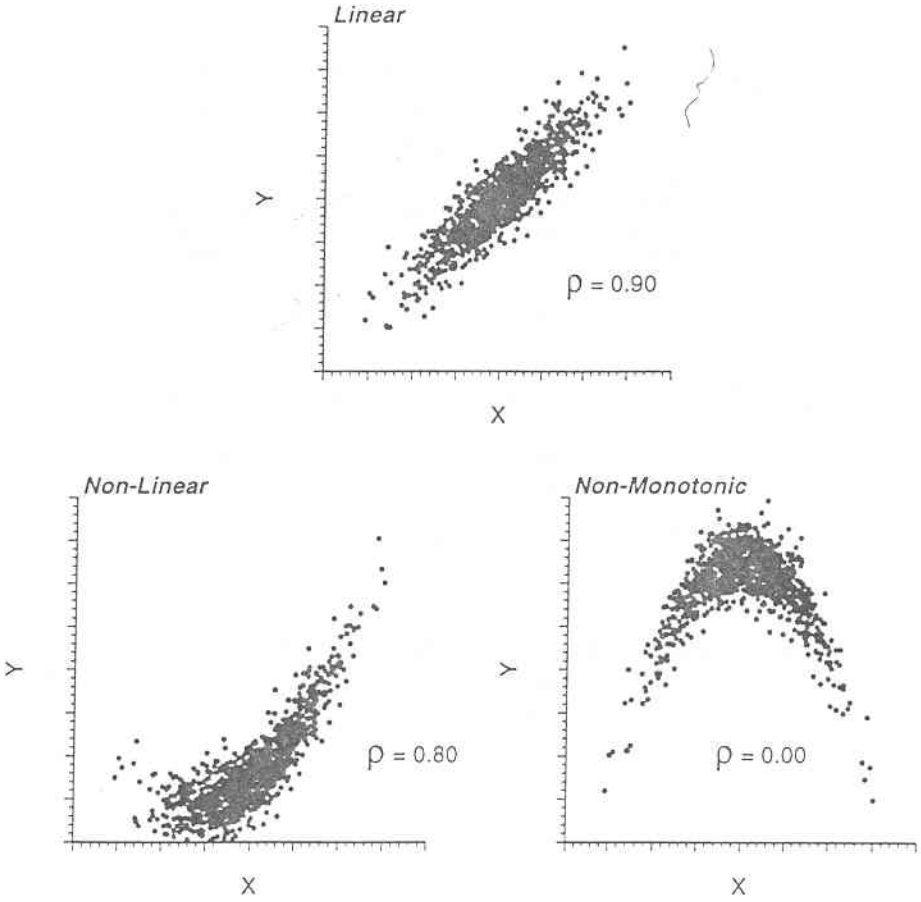


Figure 2.7: Linear, non-linear, and non-monotonic relationships with essentially the same "scatter." The correlation coefficient is significantly reduced for non-linear and non-monotonic relationships.

Note that the covariance $c_{z,y}$ and correlation coefficient ρ measure "linear" correlation. The scatter on non-linear or non-monotonic relationships is not adequately reflected by the correlation coefficients in Figure 2.7.

The GSLIB program `scatplt` creates a cross plot of paired values with either arithmetic or logarithmic scaling of both axes. The means, variances, and correlation coefficient are reported. A "rank correlation" is also presented. This rank correlation is calculated the same as above (equation 2.12) except that the cdf values $F_Z(z(\mathbf{u}_j))$ and $F_Y(y(\mathbf{u}_j))$ are used in place of the raw data values. This alleviates the impact of large outlier z- or y-data values; indeed, the histogram of both sets of values $F_Z(z(\mathbf{u}_j))$ and $F_Y(y(\mathbf{u}_j))$ are uniform between 0 and 1, hence, they do not show any specific tail shape.

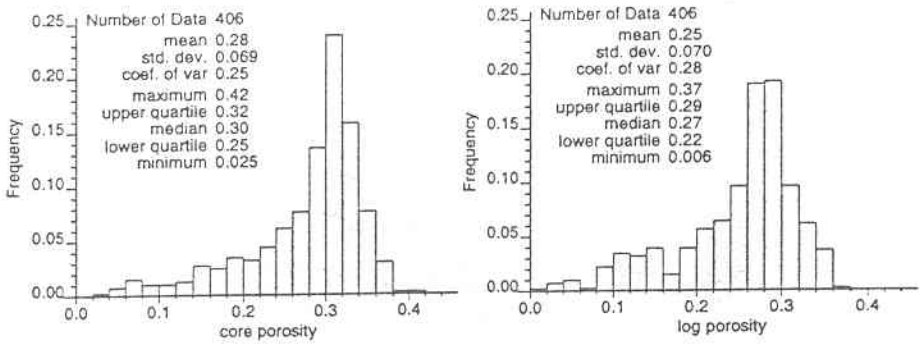


Figure 2.8: Histograms of paired core and log porosity data.

2.4 Q-Q Plots and Data Transformation

Until now the discussion of preliminary statistical concepts has focused on one variable. We turn our attention to comparing two different data distributions and data transformation.

Q-Q and P-P Plots

A *Q-Q plot* is a refined graphical tool for comparing two distributions. Of course, we could simply look at the two histograms side-by-side and compare the summary statistics. A Q-Q plot has advantages. A Q-Q plot is a cross plot of matching p-quantile values $q_1(p)$ versus $q_2(p)$ from the two different distributions.

For example, Figure 2.8 shows a histogram of core porosity and a histogram of well log derived porosity. The core porosity has a different histogram than well log derived porosity. Figure 2.9 shows the corresponding two cumulative distributions and the Q-Q plot. As highlighted in the figure, the 0.6 quantile of the core porosity is 0.309 and the 0.6 quantile of the log porosity is 0.279. The pair (0.309, 0.279) is one point on the Q-Q plot. The other points are associated to the other 406 quantiles determined from the data distributions.

When all points on a Q-Q plot fall on the 45°line, the two distributions are exactly the same. Departure from the 45°line indicates a difference in the two histograms, in particular,

1. A systematic departure above or below the 45°line indicates that the *center* or mean of the distributions is different. A shift above the 45°line implies that the Y-distribution is higher valued than the X-distribution. A shift below the 45°line entails that the X-values are higher.

2. A slope different than 45°indicates that the *spread* or variance of the two distributions is different. A slope greater than 1 (or 45°) indicates

that the Y-variance is higher than X. A slope less than 1 indicates that the X-variance is greater.

3. Curvature on a Q-Q plot indicates that the two distributions have different shape.

In practice, differences in all three features (center, spread, and shape) are observed. Looking at Figure 2.9 we note (1) the difference in average porosity from core to log (the vertical shift), and (2) the different shape of the histograms for porosity values less than 0.15.

Data Transformation

It is often necessary to transform data from one histogram to another histogram, for example, we may want to transform (1) well log derived porosity to the histogram of core porosity to correct for the approximative nature of log interpretation and the deemed more reliable histogram of core data, (2) a simulated realization of porosity values to the approximate sample data distribution, or (3) the available data to the congenial normal/Gaussian distribution for subsequent geostatistical analysis. In data transformation, the ordering of the data should be preserved, that is, high values before transformation should remain high after transformation. The transformation method described below preserves such rank ordering.

The traditional method of standardizing data is to calculate the standard residual, $Y = (Z - m)/\sigma$; where m and σ^2 are the mean and variance of the original variable Z. The mean and variance of the new Y data are 0 and 1; however, the shape of the distribution remains unchanged. A straightforward transformation procedure that also corrects the entire shape of the histogram is the rank-preserving quantile transform.

The idea of quantile transformation is to match the p-quantile of the data distribution to the p-quantile of the target or reference distribution. Consider the data variable z with cdf $F_Z(z)$. This will be transformed to a y-value with cdf $F_Y(y)$ as follows:

$$y = F_Y^{-1}\left(F_Z(z)\right) \tag{2.13}$$

A graphical representation of this procedure is easier to understand.

The example shown in Figure 2.9 can be used as an illustration. We want to transform the well log derived porosity values to have the same histogram as the core porosity values to (1) put the two data types on the same basis for further geostatistical analysis, (2) correct the histogram for the scale (support volume) difference between the two data types, thus removing any bias introduced by the well log interpretation software. As stated above, the transformation procedure amounts to *match quantiles*. That is, a well log derived porosity of 0.279 is transformed to 0.309 (the corresponding 0.6 quantile on the core porosity distribution). Any z value can be transformed provided some within-class interpolation scheme is considered to fully specify the original distribution $F_Z(z)$.

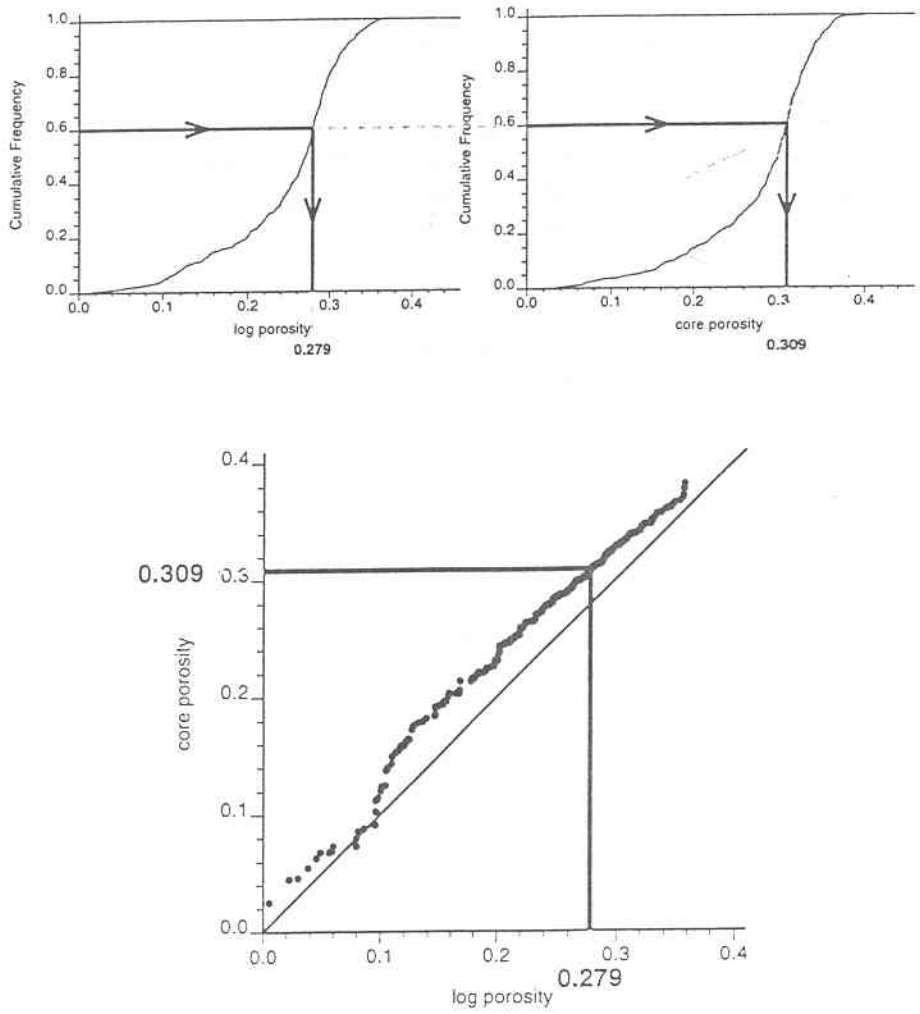


Figure 2.9: Cumulative histograms and Q-Q plot of paired core and log porosity data. The arrows illustrate the steps to get the 0.60 cumulative frequency point on the Q-Q plot: (1) the 0.6 quantile of the core porosity distribution is determined to be 0.309, (2) the 0.6 quantile of the log porosity distribution is 0.279, and (3) the values 0.309 and 0.279 are cross plotted on the Q-Q plot.

For some applications, there are far more z data to transform than there are "hard" y data for defining the target cdf $F_Y(y)$. The $Q-Q$ plot is built at the resolution of the smallest data set and, then, interpolation options are used in the lower tail, between quantile pairs, and in the upper tail. Often, linear interpolation is used because of its simplicity. More elaborate alternatives are discussed in GSLIB [62]. The `trans` program in GSLIB implements the quantile transformation procedure described above.

The normal scores transformation has the standard normal distribution as a target distribution. A normal distribution with mean m and standard deviation σ has the following probability density:

$$f(z) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{z-m}{\sigma}\right)^2} \tag{2.14}$$

Figure 2.10 illustrates the quantile transformation procedure to the standard normal distribution. The methodology is the same as the quantile transformation procedure described above; the target distribution is now the standard normal distribution. Note that although the cumulative distribution for the standard normal distribution has no closed form analytical solution, there are excellent numerical approximations [160]. The `nscore` program in GSLIB implements normal scores transformation.

Despiking

The quantile transformation or normal scores transformation procedures described above are reversible provided that there are no "spikes" in either the starting histogram or the target histogram. For example, there would be no unique transformation to any target distribution if there are 10% original zero values. There is a need to break the ties prior to transformation.

The simplest approach is to break the ties randomly. Often, the ordering of ties is left to the sorting algorithm or the order the values are encountered in the data file. More sophisticated random despiking would consist of adding a small random number to each tie and then ordering from smallest to largest. In either case, random despiking could have an effect on subsequent modeling. In particular, the variogram calculated in the transformed data space (often, after normal score transformation) could show a too-high nugget effect[2] and unrealistic short-scale variations. Notwithstanding the limitations of random despiking, it is acceptable when there is a small proportion of constant values.

It should be noted that there is no need for despiking with indicator-based methods, that is, where the range of variability is naturally divided by a series of thresholds. Each set of constant values or "spike" could constitute a separate geological population. For example, a spike of zero porosity values could correspond to completely cemented rock or shales. The cemented population might be considered a distinct geological population and modeled

[2] Details of variogram terminology and inference are presented in Chapter 4.

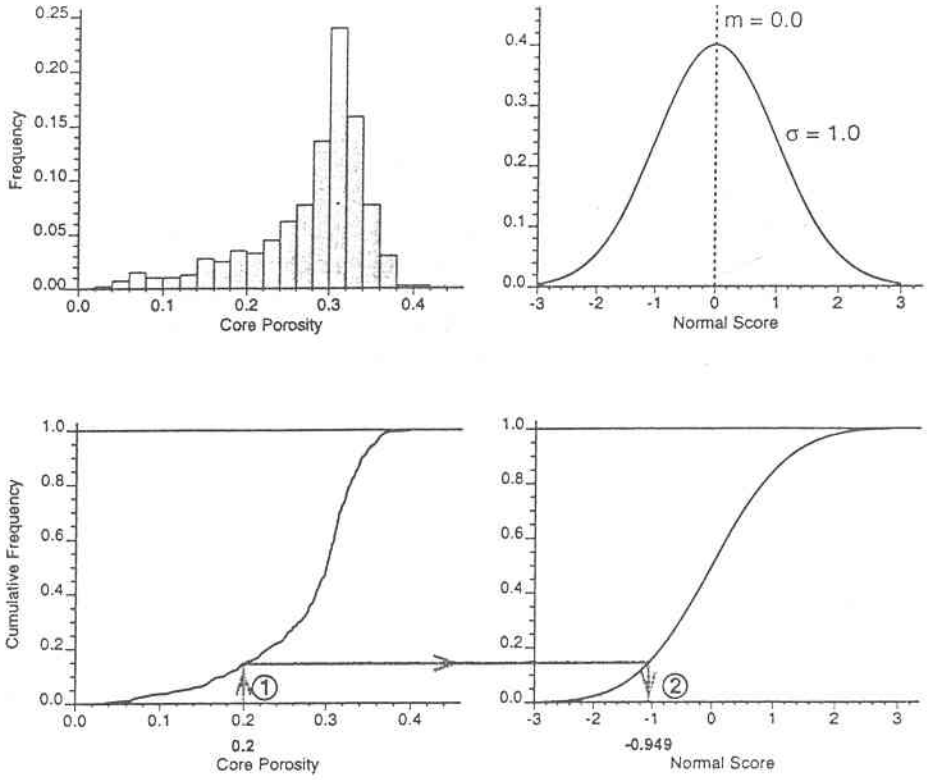


Figure 2.10: Procedure for transforming core porosity values to normal score values. The histograms are shown at the top of the figure. The cumulative distributions, shown at the bottom, are used for transformation. To transform any core porosity (say 0.2): (1) read the cumulative frequency corresponding to the porosity, and (2) go to the same cumulative frequency on the normal distribution and read the normal score value (-0.949). Any porosity value can be transformed to a normal scores value in this way, and vice versa.

as a separate facies; the properties within certain geological populations can be held constant.

Another approach must be considered when random despiking is considered unacceptable and there is no geological basis to isolate the spike as a separate population. A straightforward and effective approach for despiking was proposed by Verly [257]. The idea is to compute local averages within local neighborhoods centered at each tied data value. The data are then ordered or despiked according to the local averages; high local averages rank higher.

The `despike` program will perform the local average despiking of 1-, 2-, or 3-D data. The user must choose the size of the window. Clearly, if the window is too large or too small then no despiking will be done. There is

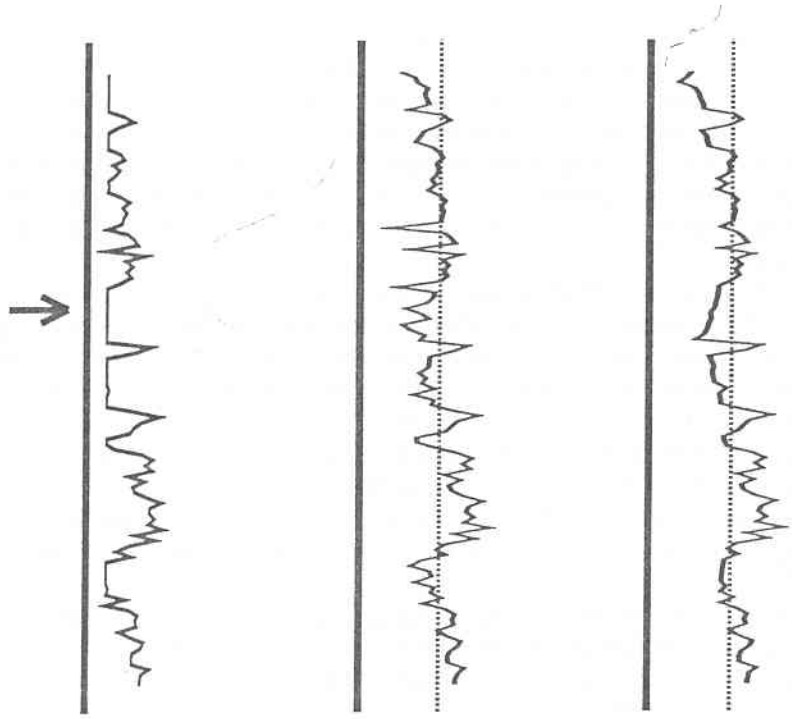


Figure 2.11: 1-D example of despiking: the original porosity values are on the far left, normal score transform of the porosity values with random despiking, and, on the right, normal score transform with local average despiking. Note that there is less randomness in the regions of constant 0.0 porosity, for example, at the location of the arrow.

little benefit to study the window size in great detail; the despiking results are relatively robust with respect to the window size.

Figure 2.11 shows an example of despiking. The normal score transformed values with the local average despiking show less small-scale fluctuations (in the constant valued zones) than the results with random despiking.

Consideration of the "despiking" problem marks a shift in emphasis between *sample* statistics and *population* statistics. Although we only have limited data, we are interested in the statistical parameters that represent the entire population or reservoir volume. A concern much more consequential than "spikes" is clustered/non-representative data.

2.5 Declustering

Data are rarely collected with the goal of statistical representivity. Wells are often drilled in areas with a greater probability of good reservoir quality. Core measurements are taken preferentially from good quality reservoir rock. These "data collection" practices should not be changed; they lead to the best economics and the greatest number of data in portions of the reservoir that contribute the greatest flow. There is a need, however, to adjust the histograms and summary statistics to be representative of the entire volume of interest.

The volume of interest is most often a particular facies of a particular stratigraphic sequence. The fact that no core porosity and permeability data have been taken from shale is accounted for in the construction of a prior sand/shale model. A concern is that, even in a particular facies, there may be clustering of data in the low- or high- valued areas.

Most contouring or mapping algorithms automatically correct this preferential clustering. Closely spaced data inform fewer grid nodes and, hence, receive lesser weight. Widely spaced data inform more grid nodes and, hence, receive greater weight. A map constructed by ordinary kriging is effectively declustered.

Even though modern stochastic simulation algorithms are built on the mapping algorithm of kriging, they do *not* correct for the impact of clustered data on the target histogram; these algorithms require a distribution model (histogram) that is representative of the entire volume being modeled. Simulation in an area with sparse data relies on the global distribution which must be representative of all areas being simulated.

Polygonal and Cell Declustering

Declustering techniques assign each datum a weight, $w_i, i = 1, \ldots, n$ based on its closeness to surrounding data. Then the histogram and summary statistics are calculated with these declustering weights. The mean and standard deviation are calculated as:

$$\overline{z} = \sum_{i=1}^{n} w_i z_i \quad \text{and} \quad s = \sqrt{\sum_{i=1}^{n} w_i (z_i - \overline{z})^2}$$

where the weights $w_i, i = 1, \ldots, n$ are between 0 and 1 and add up to 1.0. The most intuitive approach to declustering is to base the weights on the volume of influence of each sample.

As a first example, consider the 122 well locations shown on Figure 2.12. The underlying gray scale values represent the "true" porosity values, which were simulated for this example. The mean and standard deviation of the true distribution are 18.72% and 7.37%. Although hand contouring and a knowledge of the regional geology would reveal the central trend of high

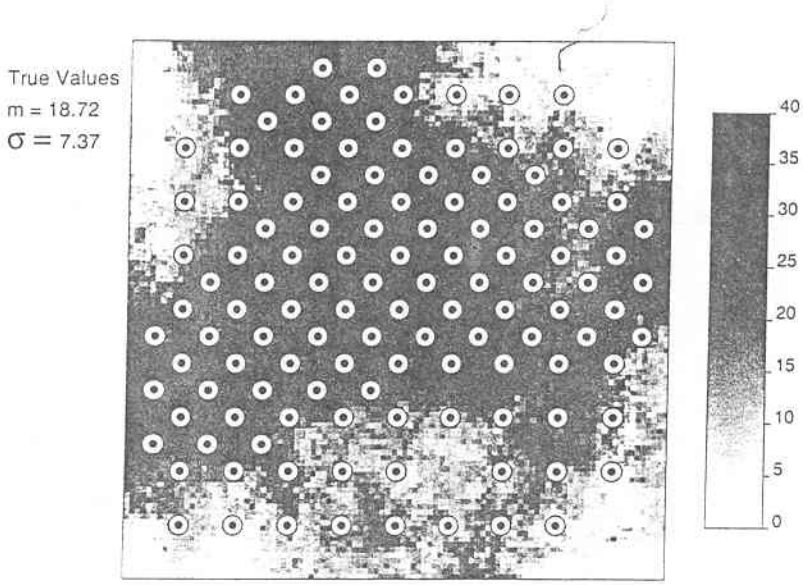


Figure 2.12: Location map of 122 wells. The gray-scale code shows the underlying "true" distribution of porosity. In practice, we may know onlythe central trend of high porosity.

porosity, the details of the true porosity distribution would be inaccessible in practice. Figure 2.13 shows an equal-weighted histogram of the 122 well data. The sample equal-weighted mean (22.26%) significantly overestimates the true mean and the standard deviation (5.58%) underestimates the true standard deviation.

The polygonal areas of influence are shown in Figure 2.14. There are many algorithms to establish the polygons around each datum [214] and calculate the area of influence $A_i, i = 1, \ldots, n$. These algorithms become CPU-intensive and somewhat unstable in 3-D. Declustering weights are taken proportional to the areas:

$$w_i^{(p)} = \frac{A_i}{\sum_{j=1}^{n} A_j}$$

where $w_i^{(p)}, i = 1, \ldots, n$ are the polygonal declustering weights. The histogram and summary statistics using these weights are shown on Figure 2.15. The proportion of data between 25 to 30% has been corrected downward and the weight given to data in the 0 to 20% porosity range has been increased. The summary statistics are now much closer to the true values (see Table 2.1).

The technique of *cell declustering* is another commonly used declustering technique [55, 144]. Cell declustering works as follows:

1. Divide the volume of interest into a grid of cells $l = 1, \ldots, L$.

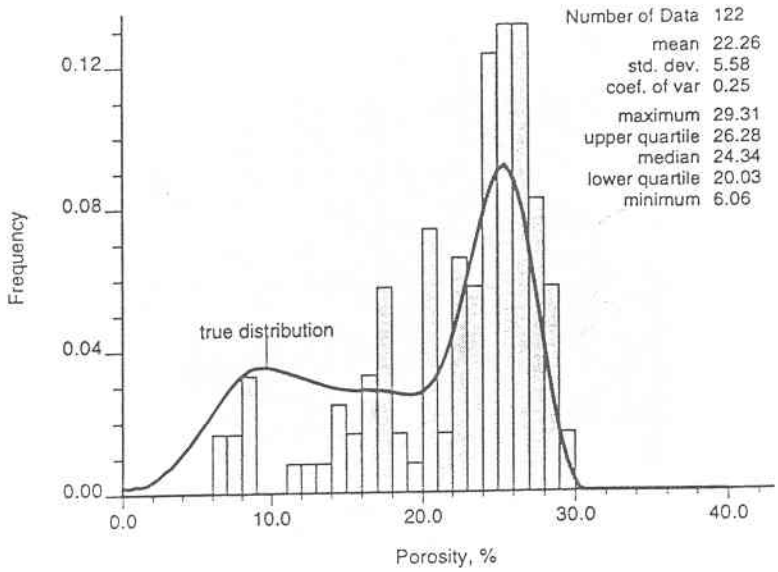


Figure 2.13: Equal-weighted histogram of 122 well data with the true reference histogram shown as the black line. Note the greater proportion of data between 25 to 30% and the sparsity of data in the 0 to 20% porosity range.

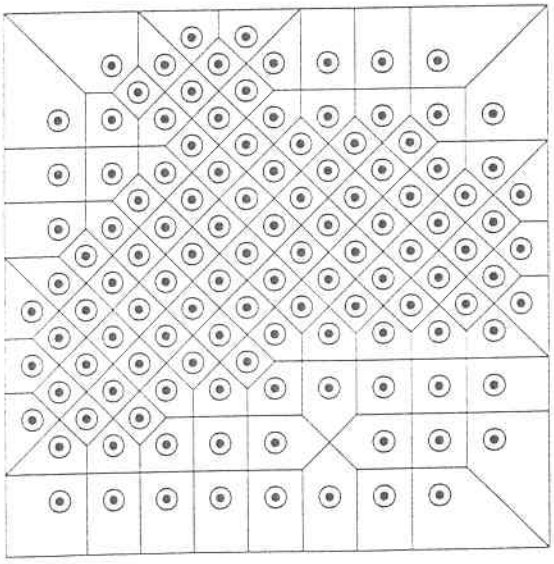


Figure 2.14: Location map of 122 wells with polygonal areas of influence.

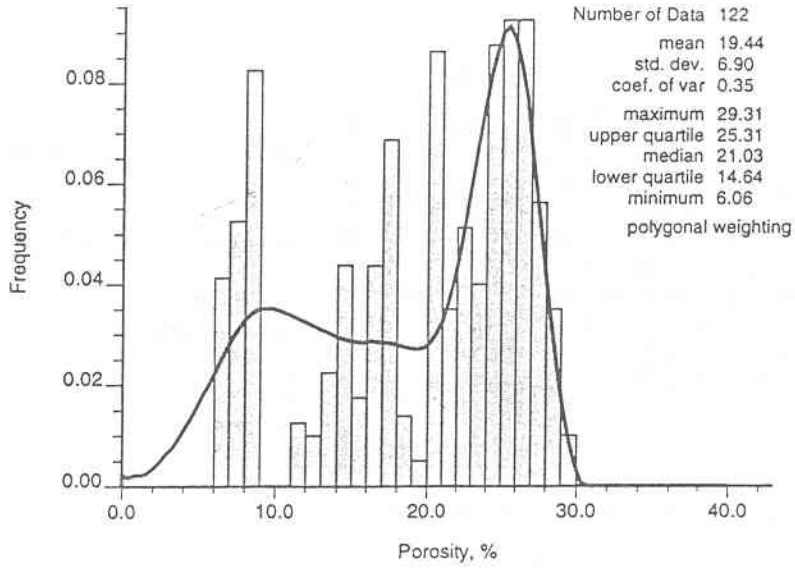


Figure 2.15: Histogram of 122 well data considering the polygonal declustering weights with the true reference histogram shown as the black line. The proportion of data between 25 to 30% has been corrected and the weight given to data in the 0 to 20% porosity range has been increased.

2. Count the occupied cells L_o, $L_o \leq L$, and the number of data in each occupied cell $n_{l_o}, l_o = 1, \ldots, L_o$ where $\sum_{l_o=1}^{L_o} n_{l_o} = n =$ the number of data.

3. Wweight each data according to the number of data falling in the same cell, for example, for datum i falling in cell l, $l \in [0, L_o]$ the cell declustering weight is:

$$w_i^{(c)} = \frac{1}{n_l \cdot L_o}$$

The weights $w_i^{(c)}, i = 1, \ldots, n$ are between $(0, 1]$ and sum to one, $\sum_{i=1}^{n} w_i^{(c)} = 1.0$. These weights are inversely proportional to the number of data in each cell: data in cells with one datum receive a weight of $1/L_o$, data in cells with two data receive a weight of $1/(2 \cdot L_o)$, data in cells with three data receive a weight of $1/(3 \cdot L_o)$, etc.

Figure 2.16 illustrates the cell declustering procedure. The area of interest is divided into a grid of $L = 36$ cells, there are $L_o = 33$ occupied cells, the number of data in each occupied cell is established by arbitrarily moving data on the grid boundaries to the right and down, the weights for data falling in the column of cells at the far right are shown in Figure 2.16.

The weights depend on the cell size and the origin of the grid network. It is important to note that the cell size for declustering is *not* the cell size for geologic or flow modeling; it simply defines an intermediate grid that allows

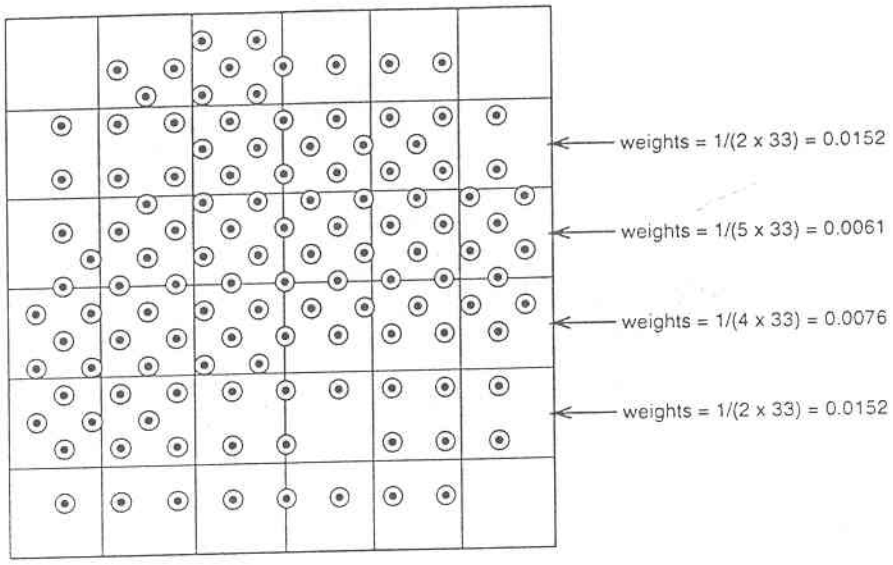


Figure 2.16: Illustration of cell declustering. The weights for the 13 data in the last column are shown. Weights for all other data are established in the same way.

assigning declustering weight. When the cell size is very small, each datum is in its own cell ($L_o = n$) and receives an equal weight. When the cell size is very large, all data fall into one cell ($L_o = 1$) and are equally weighted. Choosing the *optimal* grid origin, cell shape, and size requires some sensitivity studies; consider the following guidelines:

- Perform an areal 2-D declustering when the wells are vertical (or nearly vertical). Consider 3-D declustering when there are horizontal or highly deviated wells that may preferentially sample certain stratigraphic intervals.
- The shape of the cells depends on the geometric configuration of the data; at this point we are not seeking to quantify the spatial continuity. Thus, adjust the shape of the cells to conform to major directions of preferential sampling. For example, if the wells are more closely spaced in the X direction than in the Y direction, the cell size in the X direction should be reduced.
- Choose the cell size so that there is approximately one datum per cell in the sparsely sampled areas. Check the sensitivity of the results to small changes in the cell size; large changes in the result most likely indicate that the declustering weight is changing for one or two anomalously high or low wells.

Table 2.1: Mean and variance for example illustrated in Figure 2.12.

	Mean		Standard Deviation	
Reference	18.72		7.37	
Equal weighted	22.26	+18.9%	5.58	-24.3%
Polygonal declustering	19.44	+3.8%	6.90	-6.4%
Cell declustering	20.02	+6.9%	6.63	-10.0%

- In mature fields where there are many data and it is known that the high- or low-valued areas have been over-sampled, then the cell size can be selected such that the weights give the minimum (or maximum) declustered mean of the data. Plot the declustered mean versus the cell size for a range of cell sizes and choose the size with the lowest declustered mean (data clustered in high-valued areas) or the size with the highest declustered mean (data clustered in low-valued areas).

- The origin of the cell declustering grid and the number of cells L must be chosen such that all data are included within the grid network. Fixing the cell size and changing the origin often leads to different declustering weights. To avoid this artifact, a number of different origin locations N_{origin} should be considered for the same cell size [55, 62]. The declustering weights are then averaged for each origin offset.

Returning to the 122 well example, an optimal cell size of 280 grid units was determined by trying 95 different cell sizes from 75 through 1025 grid units in increments of 10 with 20 different origins per cell size (GSLIB program `declus` does this automatically and very fast). The plot of declustered mean versus cell size is shown on Figure 2.17. A cell size of 280 units with declustered mean 20.02% was chosen. The histogram of the 122 well data considering cell declustering weights is shown in Figure 2.18. The proportion of data between 25 to 30% has largely been corrected and the weight given to data in the 0 to 20% porosity range has been increased. The summary statistics are shown in Table 2.1.

Polygonal declustering led to statistics closer to the reference histogram in this 122 well example. This is not a general conclusion. It has been observed that polygonal declustering works well when the limits (boundaries) of the volume of interest are well defined and the polygons do not vary in size by more than a factor of, say, 10 (largest area/smallest area). Due to the difficulty of determining Voronoi volumes in influence in 3-D, cell declustering is also preferred in complex 3-D situations. In spite of the results shown in Table 2.1, polygonal declustering is known not to perform well beyond the mean. The important point is that both declustering procedures lead to a histogram closer to the truth.

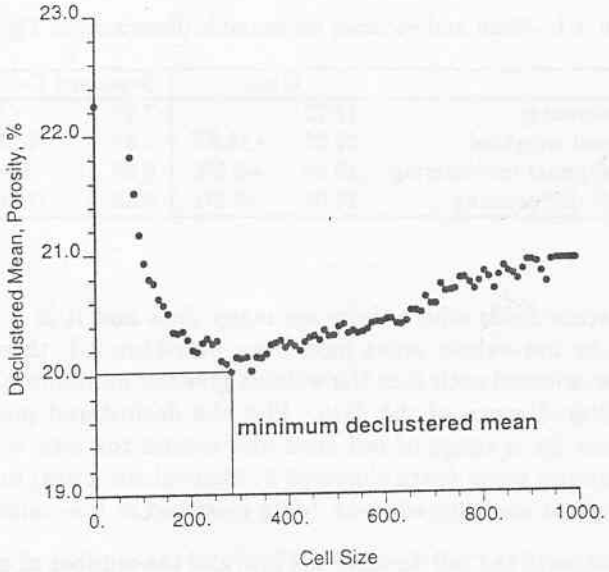


Figure 2.17: The declustered mean versus the cell size. Note the equal-weighted mean of 22.26%.

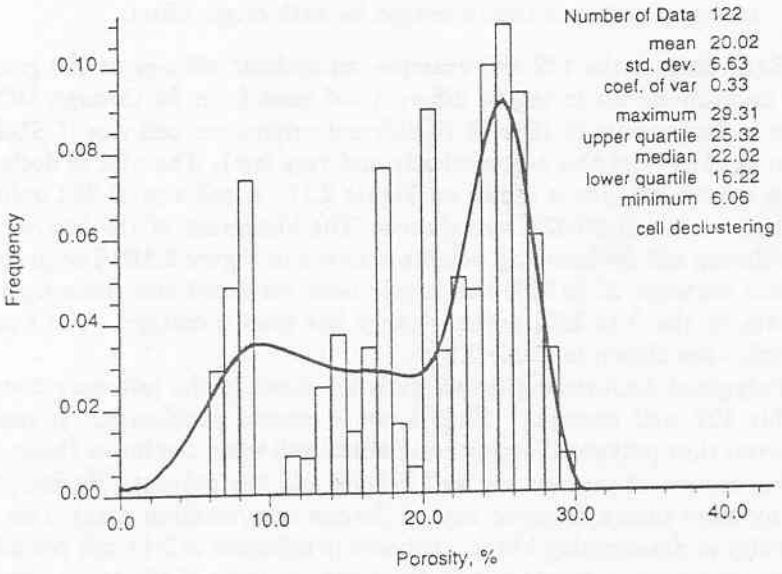


Figure 2.18: Histogram of 122 well data considering cell declustering weights. The true reference histogram is shown as the black line. The proportion of data between 25 to 30% has largely been corrected and the weight given to data in the 0 to 20% porosity range has been increased.

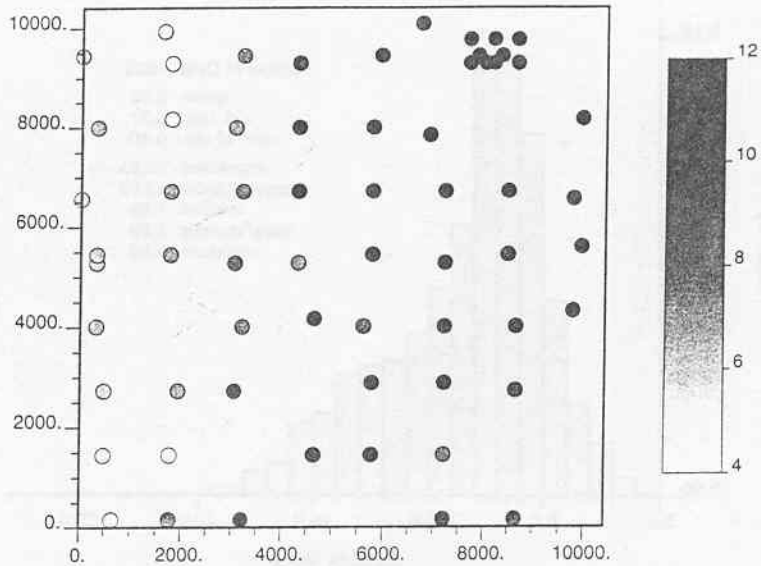


Figure 2.19: Location map of 63 wells. The well locations are shown with a gray-scale circle that represents the vertical average of porosity in that well. Note the higher porosity and the clustering of wells in the northeast corner.

Another Example

The 63 wells shown in Figure 2.19 are from a West Texas carbonate reservoir. The wells have been shown with a gray-scale coding that represents the vertical average of porosity. The cluster of wells in the northeast corner appears to be in a high-porosity area. The reservoir layer of interest is an average of 52 feet thick. An equal-weighted histogram of well log derived porosity sampled at a 1-ft spacing is shown in Figure 2.20. Note the average porosity of 8.33% and the standard deviation of 3.37%.

Applying cell declustering with an areal cell size of 1500 ft and a vertical cell size of 10 ft yields the histogram shown on Figure 2.21. The declustered mean is 7.99% and the standard deviation is 3.23%. The average porosity has decreased by 0.34 porosity units or 4.1% and the standard deviation has decreased by 0.14 porosity units or 4.2%. A Q-Q plot of the equal-weighted porosity distribution and the cell declustered porosity distribution is shown in Figure 2.22. The declustered distribution gives less weight to the porosity values between 8 to 16%.

For some applications a decrease of 4% in the average may be inconsequential. There are times, however, when 4% represents an important volume of hydrocarbon.

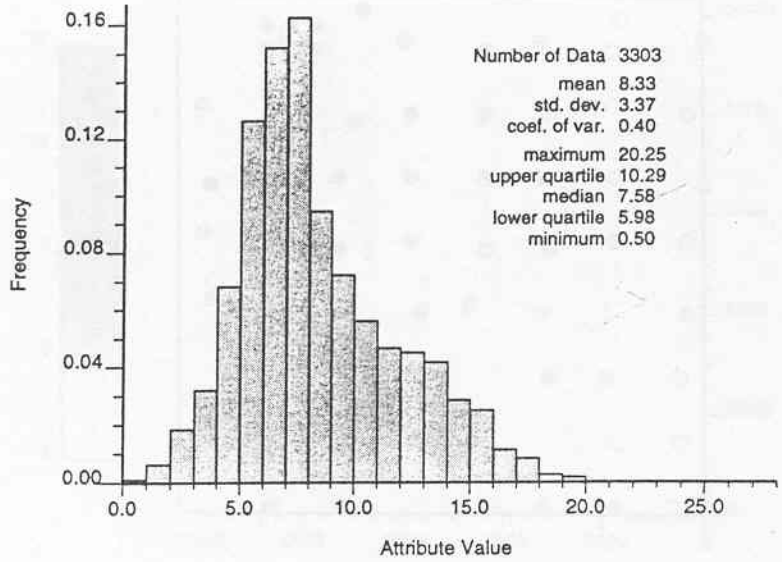


Figure 2.20: Equal-weighted histogram of porosity corresponding to the 63 wells shown in Figure 2.19.

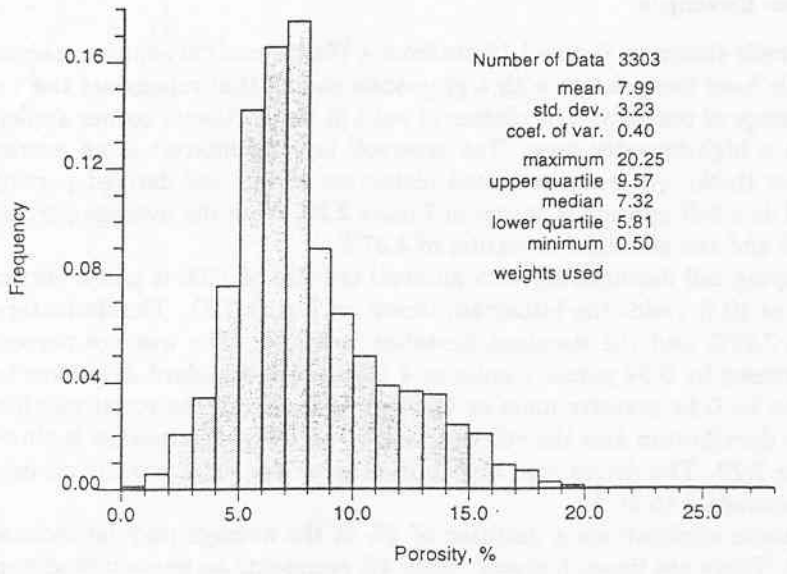


Figure 2.21: Weighted histogram of porosity where the cell declustering weights were derived from with an areal cell size of 1500 ft and vertical cell size of 10 ft.

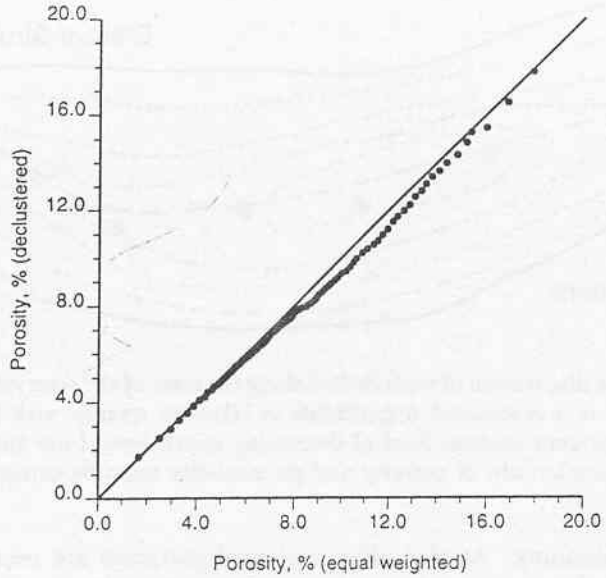


Figure 2.22: Q-Q plot of porosity values with and without declustering. Declustering has decreased the probability to be in the 8 to 16% porosity range.

Declustering with Multiple Variables

Declustering weights are largely determined on the basis of the geometric configuration of the data; therefore, only one set of declustering weights is calculated in presence of multiple variables that have been equally sampled. Different declustering weights will need to be calculated when there is unequal sampling. For example, there are often fewer permeability data than porosity data, which would require two sets of declustering weights.

Declustering weights are primarily used to determine a representative histogram for each variable; however, we also require the correlation between multiple variables. The same set of declustering weights can weight each pair contributing to the correlation coefficient (the `scatplt` program in GSLIB permits such weighting).

Correcting Nonrepresentative Data

Declustering only works when there are enough data to assign greater and lesser weights. The conventional polygonal and cell declustering methods are inadequate to determine a representative distribution unless there is adequate data coverage in both *good* and *poor* areas of the reservoir. There may be sufficient seismic or geological data to know a-priori where the high-pay *good* areas are; the first wells would be located in the best areas. Then, later in reservoir development, a geostatistical model may be constructed for further

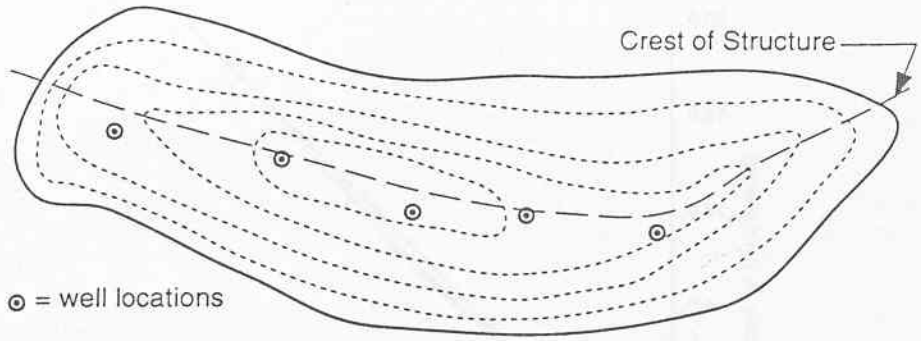


Figure 2.23: An illustration of wells drilled along the crest of the reservoir structure. At times, there is a systematic degradation in reservoir quality with depth. The dashed lines represent contour lines of decreasing depth away from the center. In this case, the distributions of porosity and permeability must be corrected.

development planning. At that time, unbiased statistics are required. The central idea in this section is to use the secondary seismic or geological data to determine representative statistics and probability distributions, in other words, correct for the non-representative data.

The first requirement is a spatial distribution of the secondary variable. The secondary variable could be a seismic attribute, a hand-contoured reservoir quality (RQ) variable, or simply structural depth if reservoir quality is linked to depth. An early application of the method presented below was to a Danish chalk reservoir, where the porosity exhibits a clear decrease with depth.

Figure 2.23 shows a schematic areal view of a reservoir where the wells are located near the crest of the structure. This is done to maximize production from a limited number of wells early in development. The dashed lines on this figure represent depth contours, which decrease away from the center. A-priori knowledge of systematic trends is common; such knowledge is often used for the placement of early production wells.

In addition to an exhaustive map of the secondary variable, a "calibration" relationship between the secondary variable (say, depth) and the reservoir property under consideration (say, porosity) is required. A schematic calibration relationship is shown in Figure 2.24. It would be preferable to have more actual data to support the trend shown in Figure 2.24. In this case, with five wells at the crest, the data would have to come from some external source of information, for example, experience with similar reservoirs in similar depositional settings.

The information illustrated in Figures 2.23 and 2.24 could be merged to provide a representative distribution of porosity. The two required pieces of information are:

1. a secondary variable, for example, depth (Y) at all locations, and

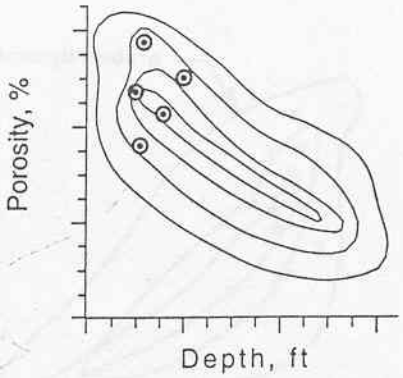


Figure 2.24: A calibration cross plot between the depth (see Figure 2.23) and the porosity. The five well data are at relatively shallow depths whereas most of the reservoir is deeper. The contour lines are equal-probability lines.

2. a bivariate relationship between the secondary variable Y and the Z variable of interest.

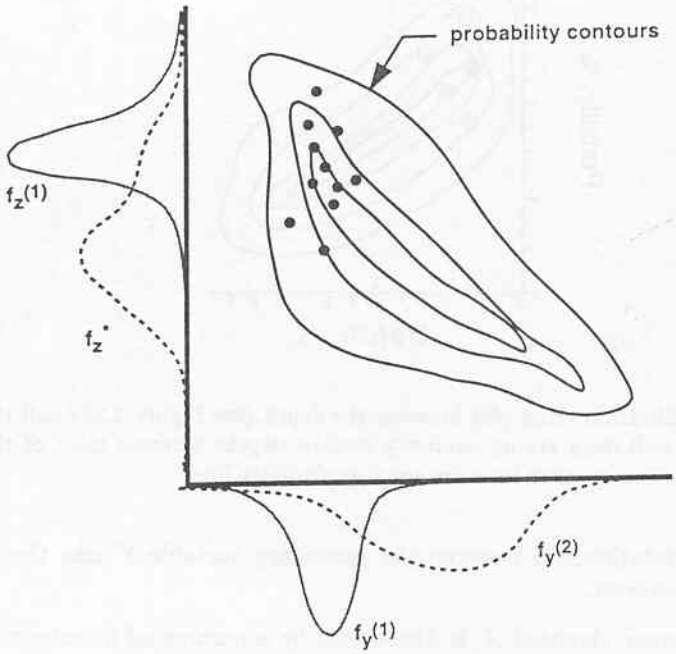
The region of interest, denoted A, is discretized by a number of locations $\mathbf{u}$, $\mathbf{u} \in A$. The notation $\mathbf{u}$ represents the 3-D coordinates of a geological modeling cell. The geological modeling cells are all about the same volume. The secondary variable is available at all locations, $y(\mathbf{u}), \mathbf{u} \in A$. For each secondary variable $y(\mathbf{u})$-value there is a corresponding probability distribution of z, denoted $\hat{f}_{Z|Y}(z|Y = y(\mathbf{u}))$. These conditional probability distributions come from the calibration relationship.

In the schematic example of Figures 2.23 and 2.24, the conditional distributions would show high z-values for shallow (low) depth y values and low z-values for greater depth. The distribution of all y-values (all $\mathbf{u} \in A$) is representative of the entire area of interest; therefore, a representative distribution of the primary Z-value may be constructed by accumulating all of the conditional distributions:

$$f_Z^*(z) = \sum_{\text{all } \mathbf{u} \in A} \frac{1}{C} f_{Z|Y}\left(z|Y = y(\mathbf{u})\right) \tag{2.15}$$

where C is a normalizing constant to ensure $f_Z^*(z)$ sums to one. Figure 2.15 gives an illustration of this procedure: the data distributions (probability distributions with solid lines) are biased due to the location of the wells, the secondary data is known everywhere and a representative distribution may be established (dashed line below y axis), and equation (2.15) may be used to calculate the representative distribution shown in a dashed line to the side of the vertical z axis.

This simple summation procedure is encoded in the `getrep` program available as a supplement to GSLIB [62]. We see that the final representative or



$f_z(1)$, $f_y(1)$ – distributions of primary (z) and secondary (y) variables at well locations

$f_y(2)$ – distribution of secondary (y) variable representative of entire area

f_z^* – derived distribution of z variable representative of entire area

Figure 2.25: Illustration of the calibration procedure of deriving representative probability distribution. Distribution $f_Z^*(z)$ is calculated with equation (2.15).

declustered distribution $f_Z^*(z)$ depends entirely on the secondary variable and the calibration. The actual z-data only contribute to establishing the calibration relationship. Of course, these z-data are also used as local data in subsequent geostatistical modeling.

The secondary data and calibration could come from seismic data. Figure 2.26 illustrates the calculation of a representative histogram using an exhaustive grid of seismic data (top left), the corresponding histogram of the seismic attribute (top right), the bivariate relationship between the seismic attribute and porosity (middle) to arrive at a distribution of porosity (bottom). The "vertical lines" on the central calibration cross plot arise from the vertical wells; there are many porosity values down each well that correspond to a single seismic attribute value.

Skipping declustering can lead to biased reservoir models. In addition to correcting bias through declustering, it is sometimes necessary to *smooth* the histograms and cross plots.

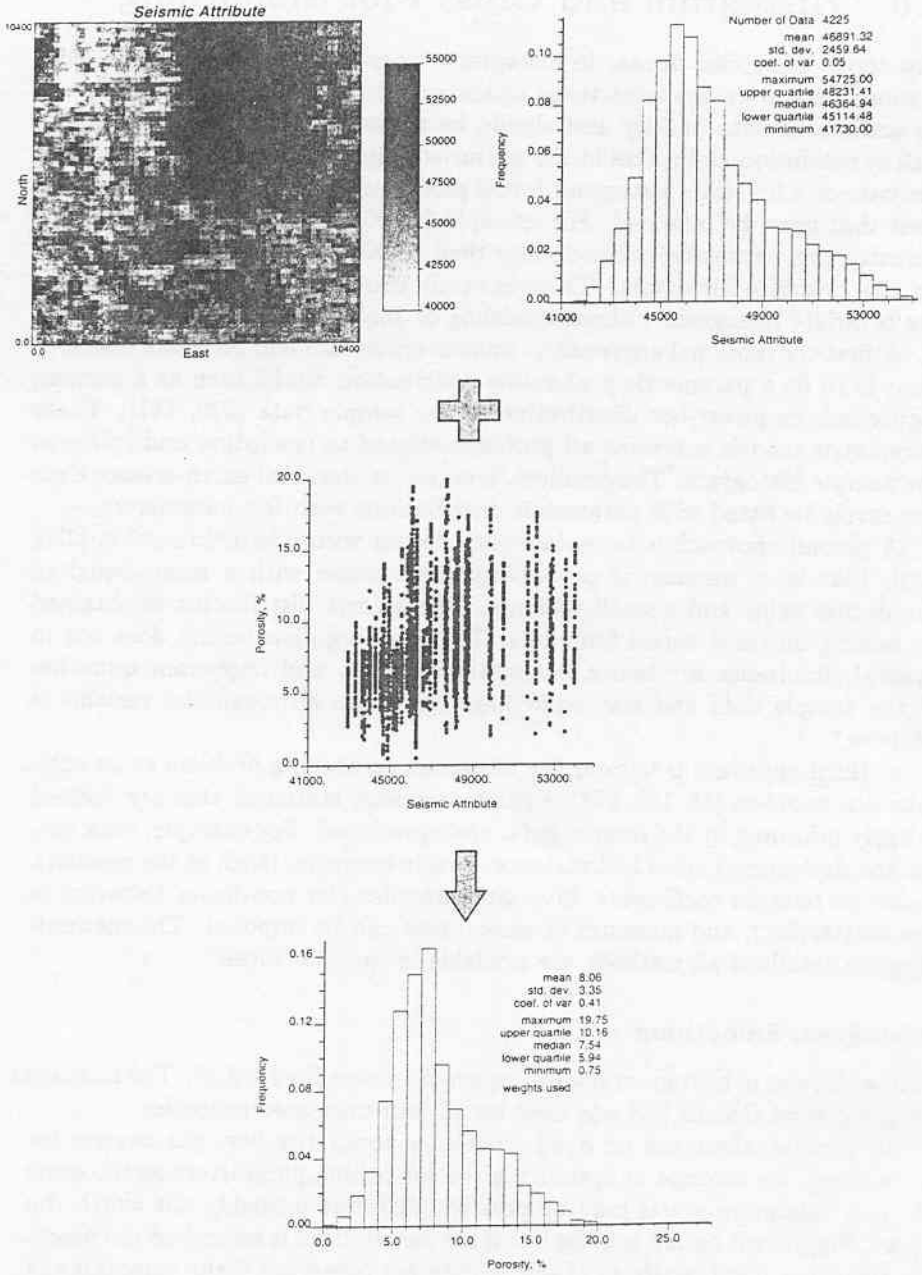


Figure 2.26: Establishing a representative histogram by using an exhaustive secondary variable and a calibration cross plot.

2.6 Histogram and Cross Plot Smoothing

Saw-tooth-like spikes appear in histograms when there are few sample data. If more data were available these spikes would not likely appear; they are an artifact of data paucity and should be smoothed out. The problems of lack of resolution and spikes in the univariate histogram become far worse in the case of a bivariate histogram (cross plot) due to the number of bivariate class that must be informed. For example, if 100 classes are required for the porosity and permeability histograms then 10,000 classes would be required for the bivariate histogram. There is rarely enough data to reliably inform the bivariate histogram without modeling or *smoothing*.

A first conventional approach to smooth univariate and bivariate distributions is to fit a parametric probability distribution model such as a normal, lognormal, or power-law distribution to the sample data [220, 141]. These parametric models overcome all problems related to resolution and spikes in the sample histogram. The problem, however, is that real earth-science data can rarely be fitted with parametric distributions with few parameters.

A second approach is to replace each datum with a kernel function [220, 223], that is, a parametric probability distribution with a mean equal to the datum value and a small variance. The *smooth* distribution is obtained by adding up these kernel functions. The resulting distribution does not in general simultaneously honor the mean, variance, and important quantiles of the sample data and may show negative values although the variable is positive.

A third approach is to pose the histogram smoothing problem as an optimization problem [58, 159, 273]. Critical summary statistics, that are deemed reliably informed by the sample data, are reproduced. For example, data limits, the declustered mean and variance, certain quantiles (such as the median), linear correlation coefficients, bivariate quantiles (for non-linear behavior in the scatterplot), and measures of smoothness can be imposed. The methodological details of all methods are available in the references.

Histogram Smoothing

Four examples of histogram smoothing are shown on Figure 2.27. The `histsmth` program from GSLIB [62] was used for all four unrelated examples.

In general, there are no fixed guidelines about the best parameters for smoothing. An attempt at optimizing the smoothing parameters on the basis of cross validation scores may be possible [220] but probably not worth the effort. Significant details may be lost if the distribution is smoothed too much. On the other hand, artifacts of scarce data are preserved if the smoothing is too "small."

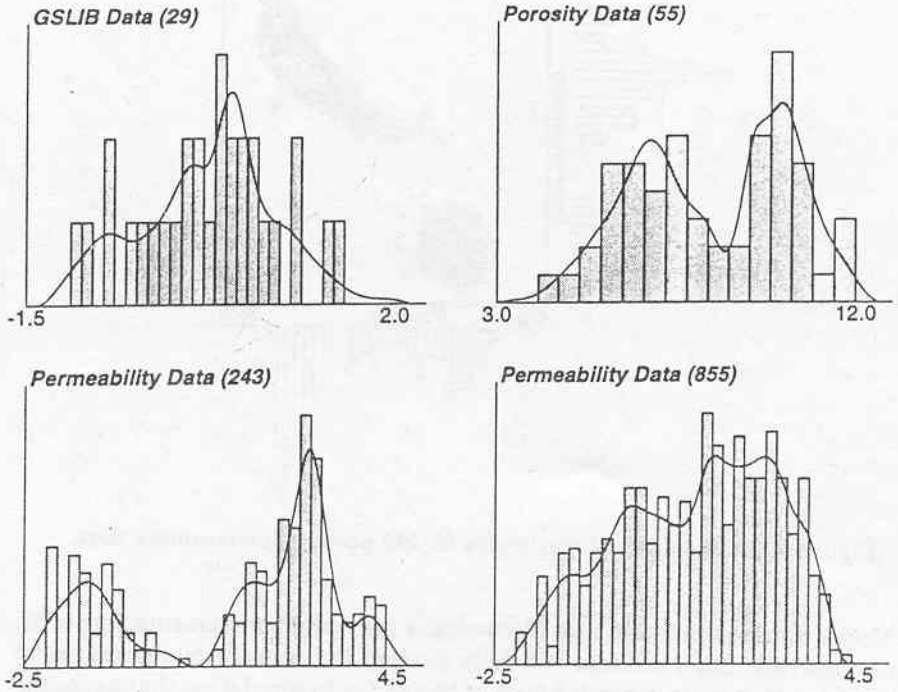


Figure 2.27: Four examples of histogram smoothing, clockwise from the upper left: 29 data taken from GSLIB [62], 55 porosity data, 243 permeability data, and 855 permeability data. The four data sets are unrelated.

Scatterplot Smoothing

Figure 2.28 shows a smoothed scatterplot for 243 porosity/permeability data pairs. The smoothed porosity and permeability distributions, shown below the porosity axis and to the left of the permeability axis, were calculated first. The bivariate distribution was then smoothed to honor these two marginal distributions. The gray scale is white for low probability and black for high probability. The final smoothed distribution shown has been constrained to the marginal distributions, some selected bivariate quantiles, and a measure of smoothness.

The `histsmth` and `scatsmth` programs from GSLIB [62] were used for this example. Once again, there is no objective estimate of the exact amount of smoothing for scatterplots. A visual verification often suffices.

2.7 Monte Carlo Methods and The Bootstrap

The application of Monte Carlo or stochastic methods is ubiquitous in modern science. In the present Section we limit ourselves to introducing the notion

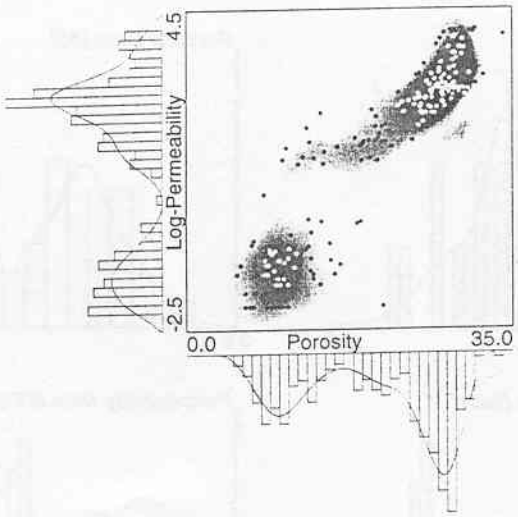


Figure 2.28: A smoothed scatterplot for 243 porosity/permeability data.

of Monte Carlo simulation and discussing a particular implementation of it, the bootstrap. Later sections will fully develop the concept of geostatistical simulation, that is, an extended form of Monte Carlo simulation that accounts for spatial constrants as imposed by a variogram model.

Monte Carlo simulation is the terminology used for the process of drawing realizations with specified probability. Drawing colored stones from an urn or picking pieces of paper (with different outcomes written on them) from a hat are classic illustrations of Monte Carlo simulation. The *Monte Carlo* prefix is dropped from "simulation" for brevity. Mechanical and electronic devices for generation of realizations were quickly replaced by numerical algorithms more than 50 years ago. These algorithms were devised to generate pseudo-random numbers which have the mathematical properties of random numbers, that is, uniformly distributed between zero and one with no correlation. The *pseudo* prefix is dropped from "random numbers" for brevity. There is a rich and varied history and literature related to random number generation. One of the latest random number generators is the ACORN generator [269] that is commonly used in GSLIB. Marsaglia's random number generator [182] is also used. The book of Ripley [210] presents some additional background and testing criteria for randomness. No major problems have been yet documented with the random number generators currently used in geostatistics.

The quantile method of simulating of realizations from arbitrary probability distributions is accomplished by (1) generating a random number, p, uniformly distributed between zero and one, and (2) calculating the inverse of the cumulative probability distribution (cdf) function:

$$y = F^{-1}(p)$$

The cdf $F(y)$ and its inverse $F^{-1}(p)$ are defined for both categorical and continuous variables. A large number of simulated realizations is typically required, that is, $y^{(l)} = F^{-1}(p^{(l)}), l = 1, \ldots, L$ where L is a large number and $p^{(l)}, l = 1, \ldots, L$is a set of random numbers generated with an algorithm and a particular "seed" number. Regardless of the number L, the set of random numbers can always be regenerated with the algorithm and seed number.

This simple yet powerful concept of Monte Carlo simulation is at the heart of most modern geostatistical techniques. Accounting for spatial correlation and integrating data of different types in simulation of reservoir models are explained in detail in later chapters. The simple bootstrap technique, described next, accounts for neither spatial correlation nor multiple data sources.

The Bootstrap

A popular application of the Monte Carlo simulation technique is the *bootstrap* method developed by Efron [81, 82]. The bootstrap is a statistical resampling technique that permits the quantification of uncertainty in statistics by resampling from the original data, in other words, "pulling yourself up by your bootstraps." There are situations where we need to know the uncertainty in input statistics because they affect uncertainty in reservoir prediction.

Consider the 17 permeability data shown at the top of Figure 2.29. There is significant uncertainty in the mean permeability since we only have 17 data. The bootstrap can be used to quantify uncertainty in the mean by the following procedure.

1. Draw (simulate) 17 values from the distribution of 17 permeability values. This can be seen as drawing with replacement, that is, some values may be chosen more than once and other values may never be chosen.

2. Calculate the average of the permeability values $\overline{K}$ and save it as one possible average.

3. Go back to step 1 many times to assess the uncertainty in the mean.

The histogram at the bottom of Figure 2.29 shows the uncertainty in the average permeability if the procedure is repeated 1000 times.

The same Monte Carlo procedure can be used to assess uncertainty in any calculated statistic, including the correlation coefficient, for example, (1) draw N pairs from the N paired observations, (2) calculate ρ, and (3) repeat many times. Figure 2.30 shows a calibration cross plot between a seismic attribute and porosity with a linear correlation coefficient of 0.54. Applying the bootstrap procedure to assess the uncertainty in the correlation coefficient leads to the distribution below.

It is also possible to use the bootstrap to assess uncertainty in more complex statistics such as recoverable hydrocarbon. The procedure in this case is as follows:

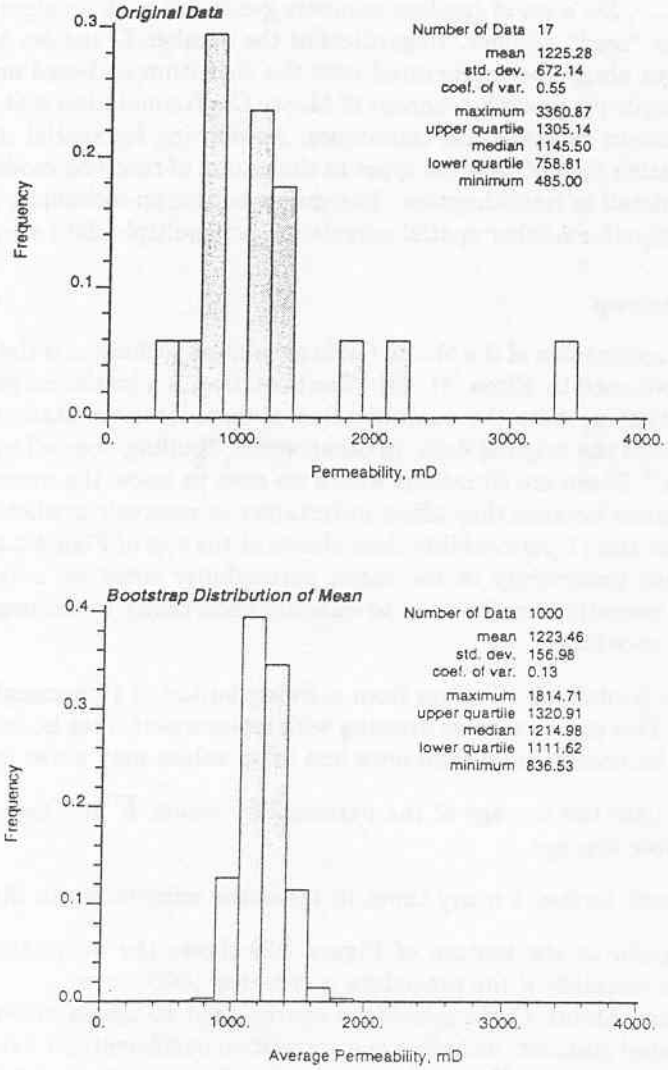


Figure 2.29: A histogram of 17 permeability data is shown at the top. The bootstrap was applied to arrive at a distribution of uncertainty in the mean permeability, shown at bottom. Note that the sample mean permeability 1225 remains unchanged, the bootstrap simply quantifies the uncertainty in the average permeability.

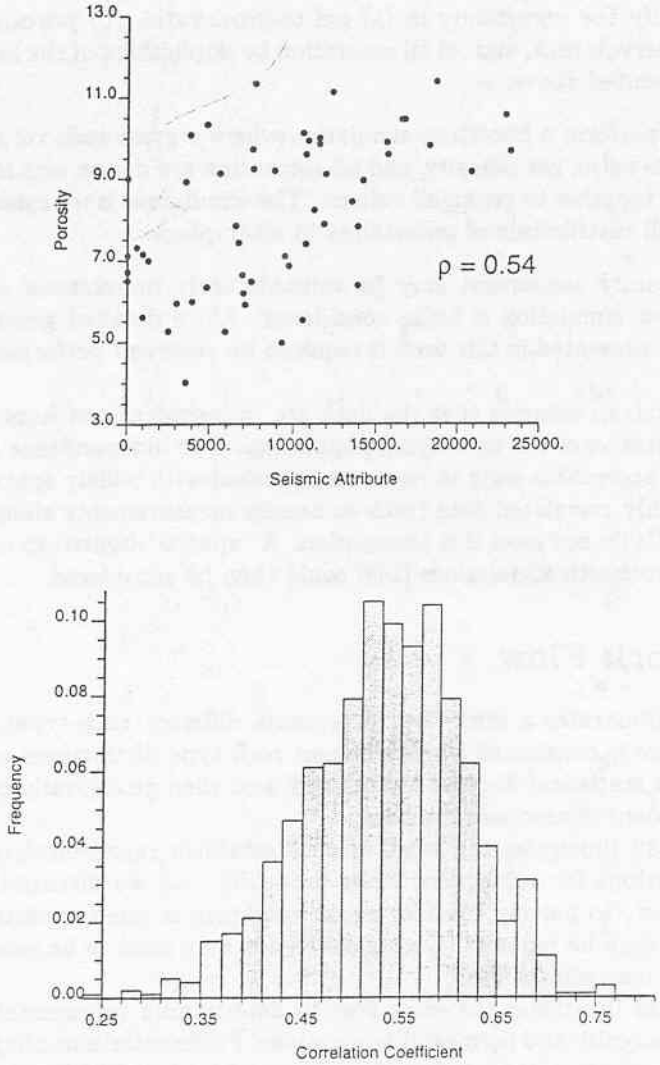


Figure 2.30: Calibration cross plot on the top with a linear correlation coefficient of 0.54. Applying the bootstrap procedure to assess the uncertainty in the correlation coefficient leads to the distribution of uncertainty shown below.

1. Quantify the uncertainty in gross rock volume by modeling the surfaces and the fluid contacts stochastically (techniques for this are developed in later chapters).

2. Quantify the uncertainty in (a) net-to-gross ratio, (b) porosity in the net reservoir rock, and (c) oil saturation by application of the bootstrap, as presented above.

3. Then, perform a bootstrap simulation where a gross rock volume, net-to-gross ratio, net porosity, and oil saturation are drawn and then multiplied together to get an oil volume. The simulation is repeated to lead to a full distribution of uncertainty in oil-in-place.

Such uncertainty assessment may be valuable early in reservoir appraisal where no flow simulation is being considered. More detailed geostatistical modeling, as presented in this text, is required for reservoir performance predictions.

The bootstrap assumes that the data are *independent* one from another and representative of the underlying population. The independence assumption may be acceptable early in reservoir appraisal with widely spaced wells; however, highly correlated data (such as nearby measurements along a well) and close wells do not meet this assumption. A "spatial" bootstrap or resampling from stochastic simulations [148] could then be considered.

2.8 Work Flow

Figure 2.31 illustrates a work flow to separate different rock types. An iterative scheme is considered where different rock type distinctions are used, statistics and statistical displays are created, and then geologically or statistically significant distinctions are kept.

Figure 2.32 illustrates the work flow to establish representative global facies proportions for subsequent facies modeling. As we discussed earlier in this chapter, to naively consider equal weighting is often unsatisfactory. Declustering may be required. Geological trends may need to be considered. Seismic data may also be used.

Figure 2.33 illustrates the work flow to determining representative histograms for porosity and permeability variables. Preferential sampling affects the histograms and statistics of continuous variables just like facies proportions. The required statistics for subsequent modeling are different, but the declustering steps are similar to those needed for facies.

Figure 2.34 illustrates the work flow for data transformation. There are a number of situations where data transformation is required: (1) correcting bias in well log derived measurements, (2) fixing the distribution of values from a geostatistical simulation, (3) transformation to the normal or Gaussian distribution prior to spatial statistics and geostatistical modeling, and (4)

conversion to a standard distribution so that purely spatial differences can be discerned. The transformation is straightforward, but checking must follow.

Figure 2.35 illustrates the work flow to establish a bivariate calibration relationship for subsequent cosimulation or modeling. The use of seismic data, production data, or a geological trend variable is common in geostatistical reservoir modeling. In such cases, the calibration between the variable of interest and the secondary variable is critical. The main steps in this work flow are to get the data pairs and smooth the corresponding cross plot.

Figure 2.36 illustrates the work flow to check histogram and statistics reproduction of a geostatistical model. We must compare statistical displays and interpret Q-Q plots.

Figure 2.37 illustrates the work flow to use the bootstrap to calculate the uncertainty in a global statistic. The global statistics one would usually consider include (1) the mean and variance, (2) the entire histogram, and (3) the correlation between two variables of interest. The bootstrap should not be applied to data with significant spatial correlation; hence, the check for spatial correlation.

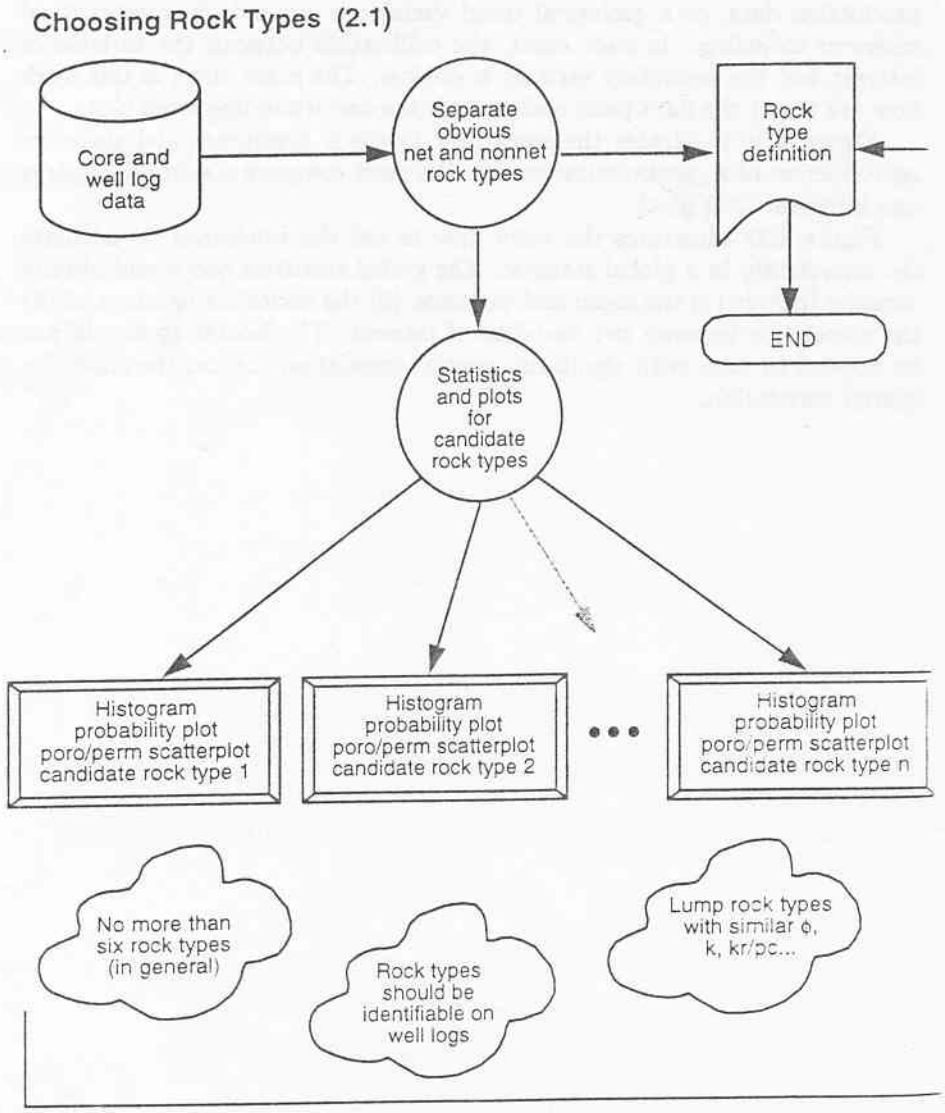


Figure 2.31: Work Flow 2.1: choosing rock types for geostatistical modeling.

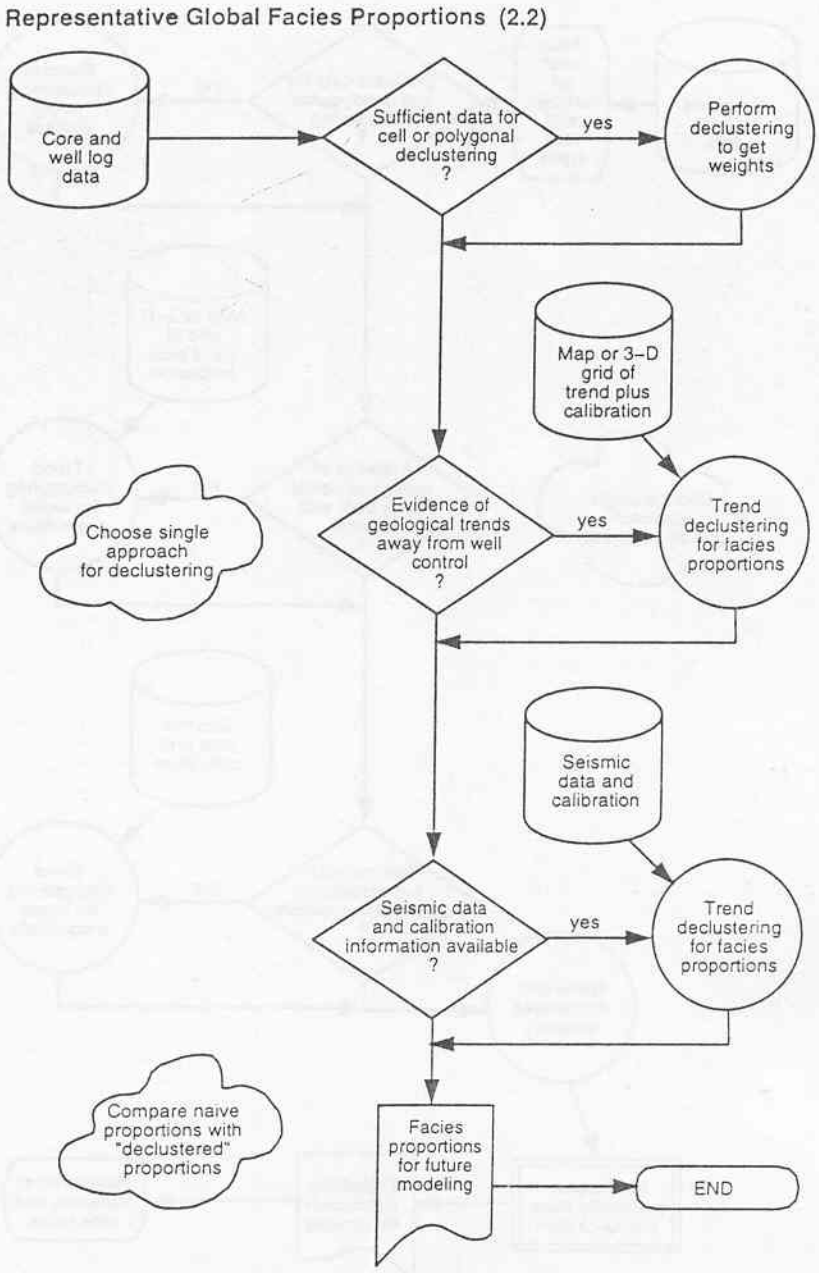


Figure 2.32: Work Flow 2.2: determining representative global facies proportions for subsequent facies modeling.

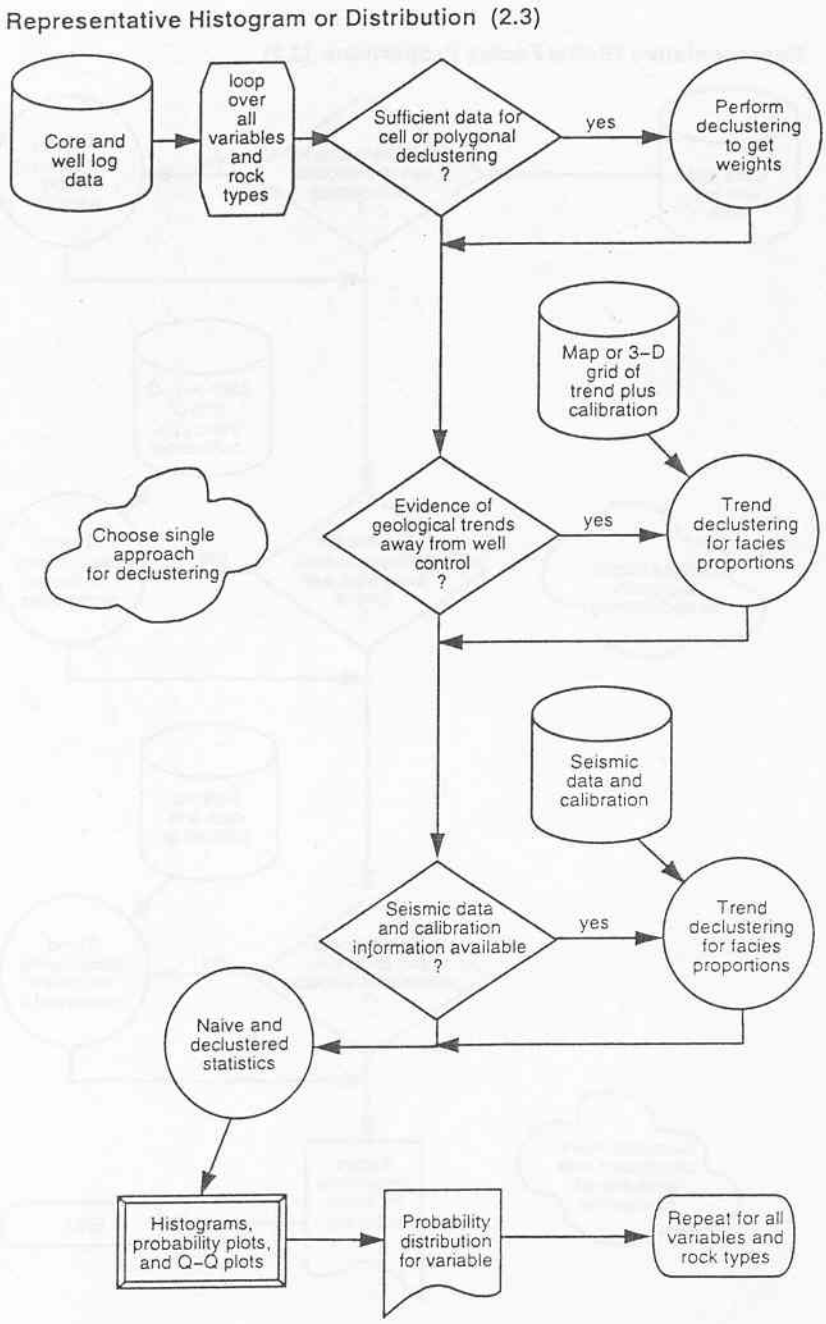


Figure 2.33: Work Flow 2.3: determining representative histograms for porosity and permeability variables for subsequent facies modeling.

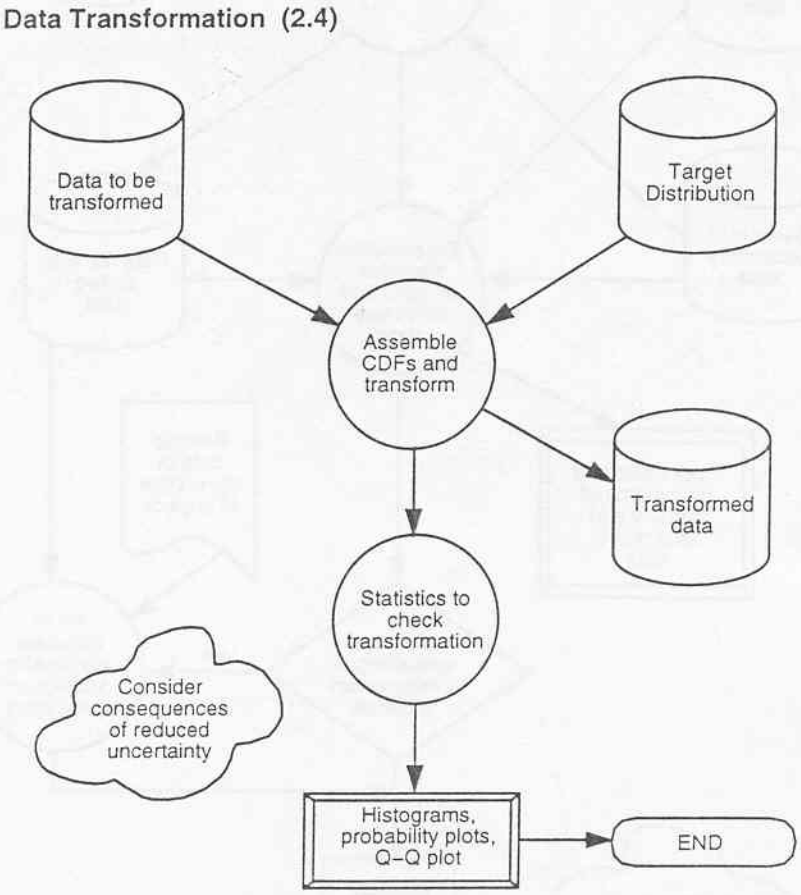


Figure 2.34: Work Flow 2.4: data transformation.

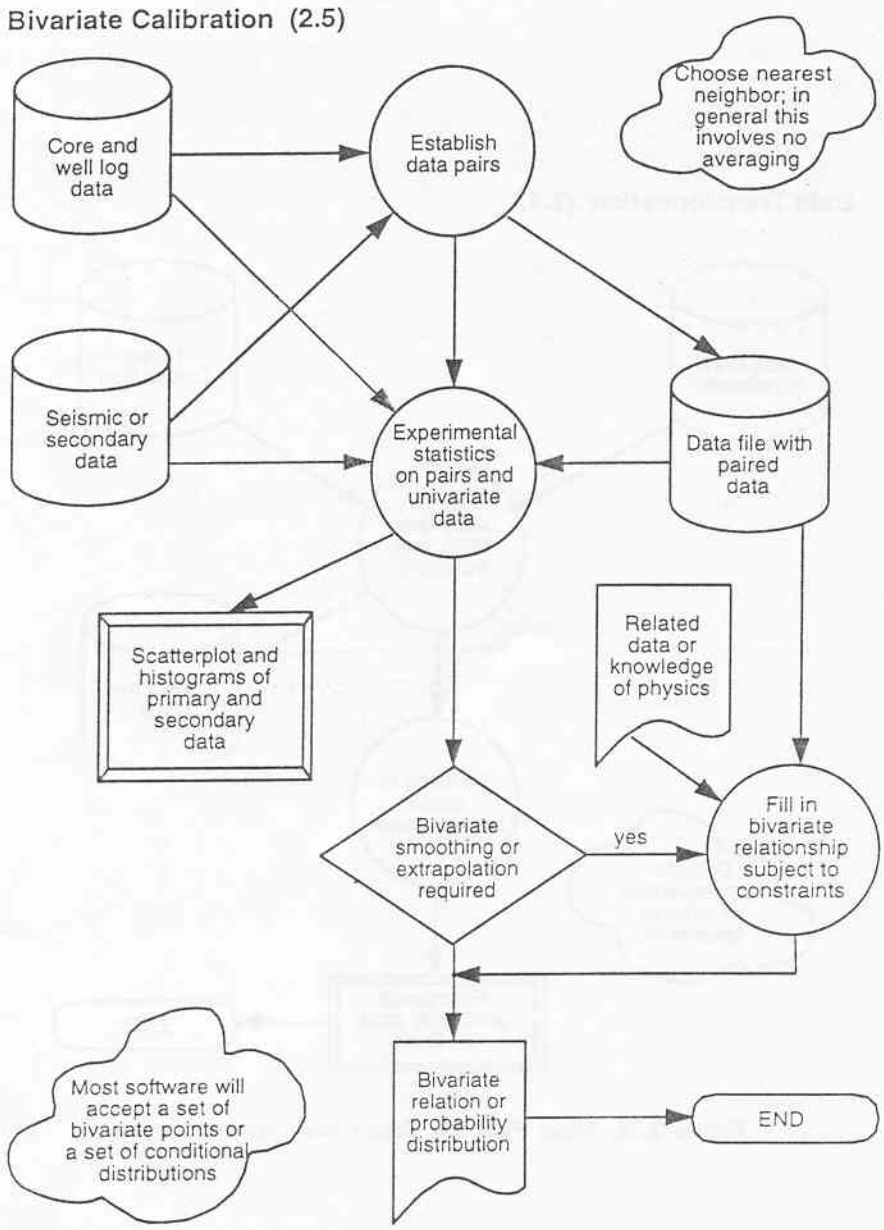


Figure 2.35: Work Flow 2.5: establish bivariate calibration relationship for subsequent cosimulation or modeling.

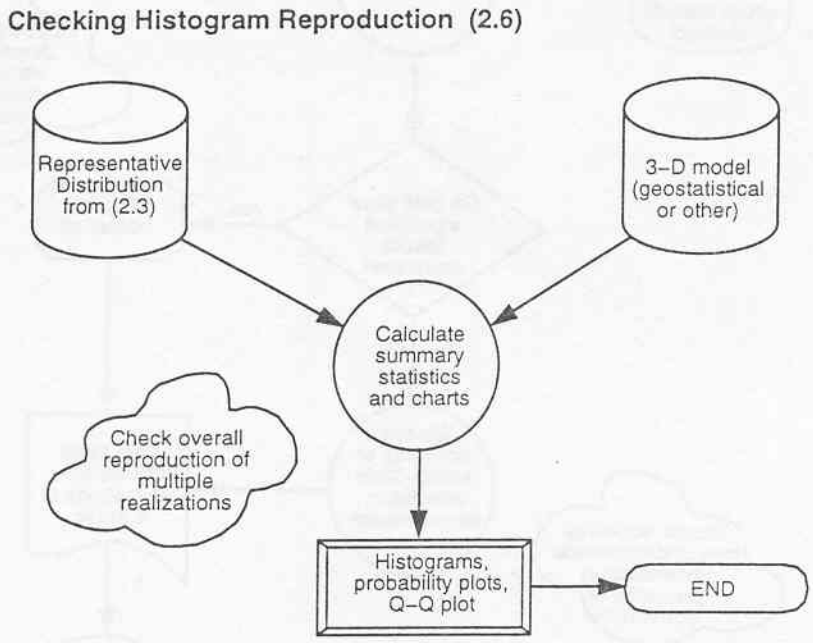


Figure 2.36: Work Flow 2.6: check histogram and statistics reproduction of model.

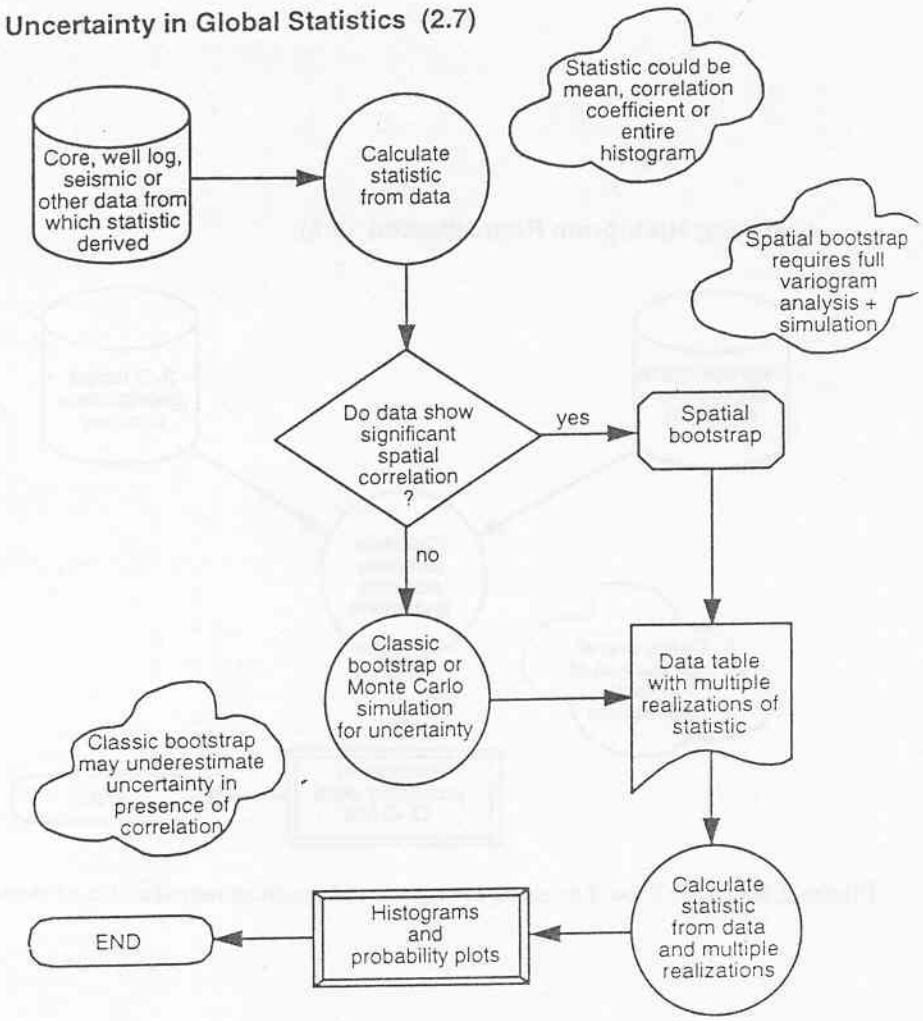


Figure 2.37: Work Flow 2.7: calculate uncertainty in global statistic.

Chapter 3

Gridding Reservoir Layers

A first step in reservoir modeling is to establish the major reservoir architecture, that is, the geometry of the hydrocarbon-bearing formation and important adjacent geological formations. The reservoir "volume" must be filled with a grid network that makes it straightforward to account for the variety of available data, in particular the geological information concerning layering and stratigraphic correlation.

Section 3.1 describes alternatives for gridding geological models. Simple Cartesian grids with rectangular parallelepiped blocks are almost always preferred. A number of more sophisticated gridding schemes are presented for special applications.

Section 3.2 describes how the reservoir is divided into significant reservoir layers. Each reservoir layer, defined on the basis of chrono-stratigraphy, has a specific stratigraphic correlation style and corresponding coordinate system.

Section 3.3 describes elementary techniques for handling faults. Faults compartmentalize the reservoir into fault blocks that must be unraveled prior to geostatistical modeling. Specifically, Section 3.3 focuses on deterministically modeling known faults observed on seismic data. Stochastic fault modeling is only mentioned briefly.

Section 3.4 presents techniques to assess uncertainty in the reservoir geometry. This is unnecessary when constructing a single reservoir model; however, the largest contribution to uncertainty in reservoir volumetrics and performance forecasting is often due to uncertainty in large-scale reservoir geometry.

This chapter is concerned with establishing stratigraphic, layer-specific, Cartesian grid systems, that is, the appropriate grid system within which lithofacies, porosity, and permeability will be modeled. Techniques for lithofacies, porosity, and permeability modeling are presented in Chapters 6, 7, and 8. Often, the stratigraphic grid system for geostatistical modeling is not the grid system chosen for flow simulation; for the latter, well placement, fluid flow processes, numerical accuracy, and CPU cost must also be considered. Scale averaging for flow simulation is discussed in Chapter 11.

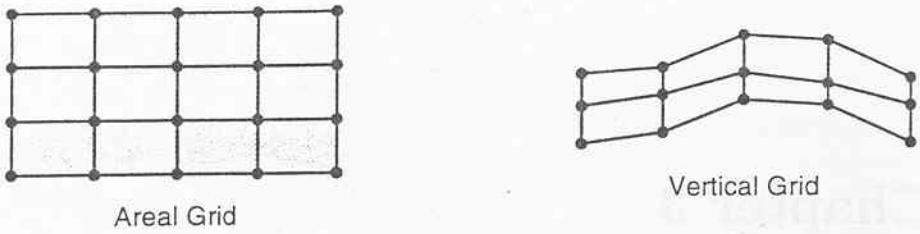


Figure 3.1: Basic topology used for geostatistical model grid: corner-point surface and isochore grids.

Figure 3.2: Cell-centered grids are not connected according to the Z-coordinate elevation (left-hand side), rather according to cell index number, see dashed lines on right-hand side.

3.1 Gridding for Geologic Modeling

The basic topology used in most reservoir modeling and throughout this book is *Cartesian*, defined by corner-point areal grids and isochore vertical grids (see Figure 3.1). An elevation is associated to each grid node or cell corner. Elevations associated to entire grid cells would constitute a "cell-centered" specification. The areal spacing of the grid nodes is typically regular, that is, with constant X and Y intervals.

The vertical coordinate is measured perpendicular to the horizontal plane, that is, *isochore* thickness, which is measured vertically even when the stratigraphic formation dips. At times, the vertical coordinate is measured perpendicular to stratigraphic dip, that is, *isopach* thickness is considered.

Standard practice is to have the cells aligned along the vertical Z axis. Figure 3.3 shows this approach (on the left) and an alternative, where the cells are aligned on lines perpendicular to the bounding surfaces. This latter alternative would allow more correct horizontal distances at an added computational cost. Figure 3.4 illustrates the distortion in calculated horizontal distances that results from having all cell boundaries line- up vertically. To alleviate the consequences of this decision, and to make the grid more efficient, the primary X, Y, and Z coordinates could be rotated to conform to the overall strike and dip of the reservoir.

A new set of orthogonal coordinates can be established if the dip is significant. Although a single-step transformation could be defined, a two-step

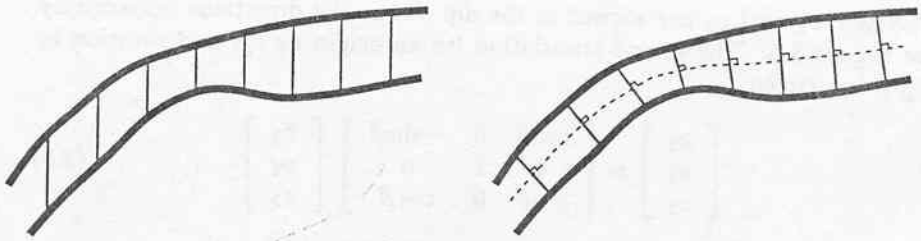


Figure 3.3: Two approaches to establish areal gridding (1) on the left, the cells/nodes are stacked vertically, and (2) on the right, the cells/nodes are aligned on lines perpendicular to the bounding surfaces. The first approach, with a universal vertical coordinate, is usually adopted because of computational simplicity.

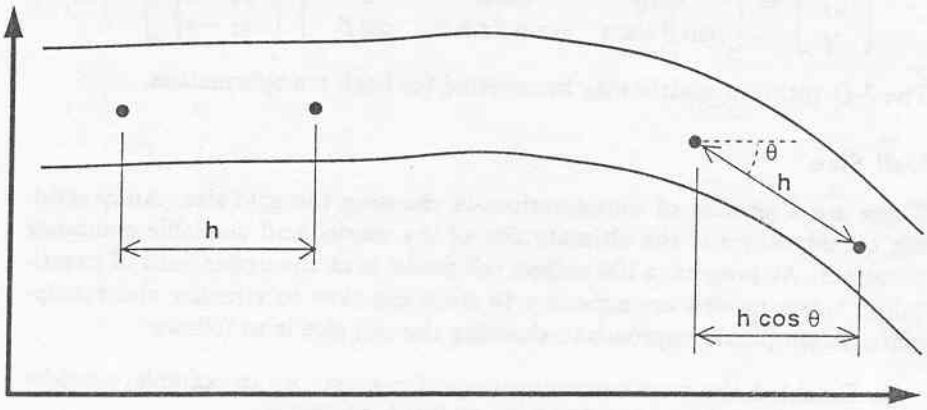


Figure 3.4: Both pairs of data points are at the same stratigraphic Z coordinate and separated by the same distance h; however, the pair of points to the right appear closer.

procedure is clearer: (1) rotate the horizontal X and Y coordinates to be aligned along the strike and dip direction of the stratigraphic unit (see Figure 3.5), and (2) rotate the new X direction to be aligned with the dip direction. The first translation (to an origin at x_1^0, y_1^0) and clockwise rotation of the original X (x_1) and Y (y_1) by α is written:

$$\begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix} = \begin{bmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 - x_1^0 \\ y_1 - y_1^0 \\ z_1 - z_1^0 \end{bmatrix} \tag{3.1}$$

$$\text{or} \quad [X_2] = [R_{Z_1}]\,[X_1]$$

The new x_2 and y_2 are aligned in the dip and strike directions respectively, see Figure 3.5. The second translation (to an origin at z_1^0) and rotation by dip β is written:

$$\begin{bmatrix} x_3 \\ y_3 \\ z_3 \end{bmatrix} = \begin{bmatrix} \cos\beta & 0 & -\sin\beta \\ 0 & 1 & 0 \\ \sin\beta & 0 & \cos\beta \end{bmatrix} \begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix} \tag{3.2}$$

$$\text{or} \quad [X_3] = [R_{Y_2}]\,[X_2]$$

So, $[X_3] = [R_{Y_2}]\,[R_{Z_1}]\,[X_1]$, or multiplying the two rotation matrices:

$$\begin{bmatrix} x_3 \\ y_3 \\ z_3 \end{bmatrix} = \begin{bmatrix} \cos\alpha\cos\beta & -\cos\beta\sin\alpha & -\sin\beta \\ \sin\alpha & \cos\alpha & 0 \\ \sin\beta\cos\alpha & -\sin\beta\sin\alpha & \cos\beta \end{bmatrix} \begin{bmatrix} x_1 - x_1^0 \\ y_1 - y_1^0 \\ z_1 - z_1^0 \end{bmatrix} \tag{3.3}$$

The 3-D rotation matrix may be inverted for back transformation.

Cell Size

There are a number of considerations in choosing the grid size. An overriding consideration is the ultimate size of the model and available computer resources. At present, a 100 million cell model is at the upper limit of practicality; larger models are expensive to store and slow to visualize and manipulate. A simplistic approach to choosing the cell size is as follows:

1. Establish the gross reservoir volume of interest. As an example, consider a total envelope of 10,000 ft by 10,000 ft by 200 ft.

2. Choose the vertical "stratigraphic" cell size ds needed to represent the heterogeneities observed in the wells. Consider, say $ds = 2$ ft.

3. Pick a target model size, say 10 million cells.

4. Calculate the consequent areal grid size, da, in this case,

$$\frac{10,000}{da} \cdot \frac{10,000}{da} \cdot \frac{200}{ds} = 10,000,000$$

   yielding $da = 31.6$. Rounding da to 50 ft would yield a 200 by 200 by 100 cell model with a total of 4 million cells.

5. Iterate as necessary to achieve a balance between the total number of cells, the stratigraphic cell size, and the areal cell size. Of course, there is no need to use square areal cells. Pronounced areal anisotropy could be accommodated by rectangular cells.

Choosing the cell size according to a preconceived maximum model size is not completely satisfactory. It is worth considering more *reservoir-specific* considerations. Perhaps the cell size could be much coarser and yet meet

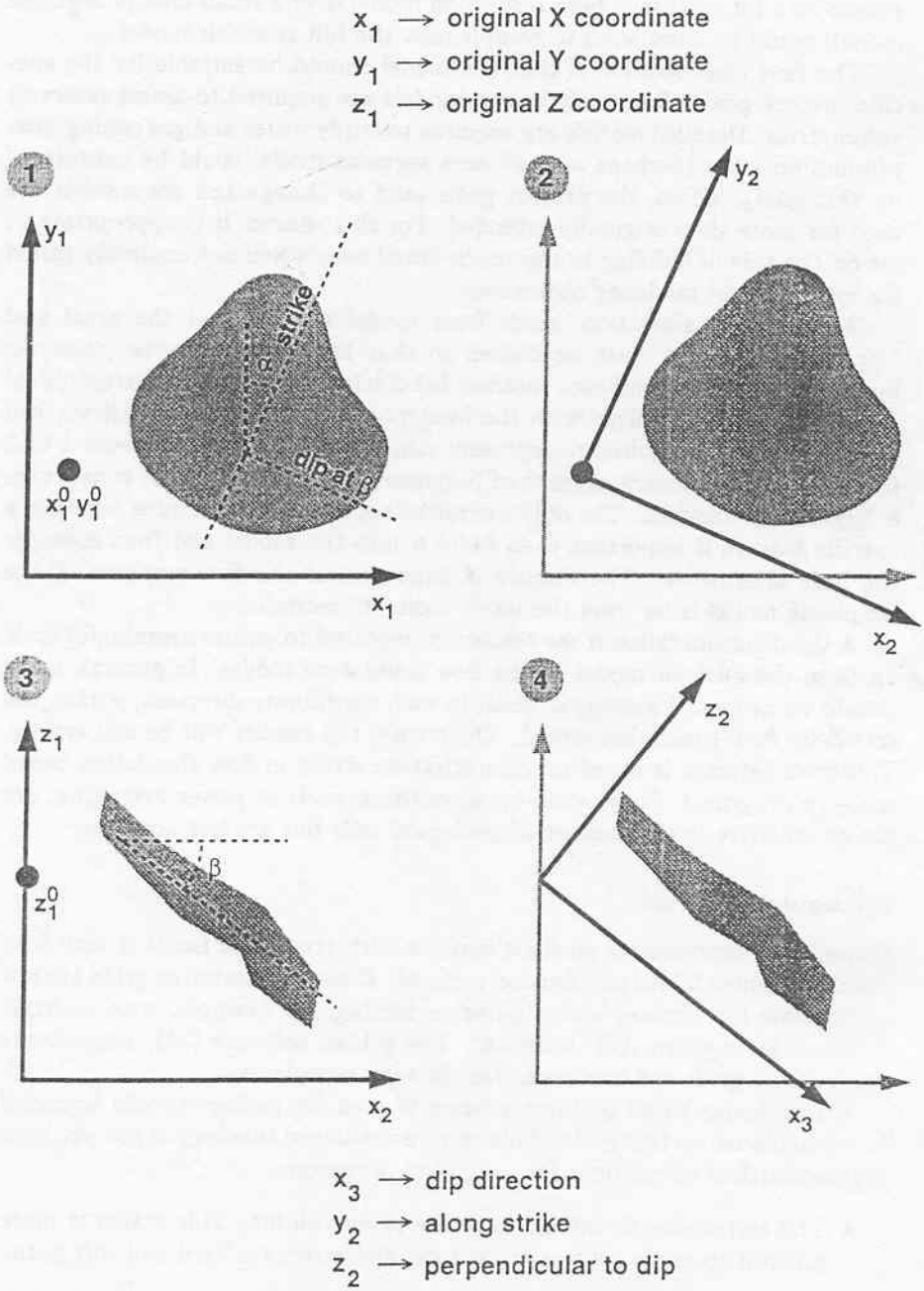


Figure 3.5: Rotated Z coordinate system for a dipping stratigraphic unit.

all objectives of the reservoir modeling study? Alternatively, if the cell size *should* be a lot smaller, a high resolution model over a small area (a segment model) could be considered to complement the full reservoir model.

The first consideration is that the model should be suitable for the specific project goals. Relatively coarse models are required to assess reservoir volumetrics. Detailed models are required to study water and gas coning near production wells (perhaps a small-area segment model could be considered for this goal). Often, the project goals tend to change and the models are used for more than originally intended. For this reason, it is appropriate to err on the side of building in too much detail even when not explicitly called for by the initial modeling objectives.

A second consideration, aside from model size, is that the areal and "stratigraphic" grid must be chosen so that important features (reservoir boundaries, faults, significant internal lithofacies changes, and petrophysical variations) can be resolved with the final model. For example, a 1-ft vertical cell size would be required to represent relatively thin shales (between 1 to 3 ft). There is a necessary element of judgment required as to what constitutes a "significant" feature. The only definitive approach to determine whether a specific feature is important is to build it into the model and then consider the best alternative. The feature is important if the flow response of the simplistic model is far from the more "correct" model.

A third consideration is the resolution required to ensure meaningful scale up from the geologic model to the flow simulation model. In general, there should be at least 3 geological cells, in each coordinate direction, within the scaled-up flow simulation model. Otherwise, the results will be too erratic. This consideration is based on discretization errors in flow simulation based scale-up programs. Some scale-up algorithms, such as power averaging, are not as sensitive to the number of geological cells but are less accurate.

Triangulated Grids

Numerically representing geologic surfaces with triangular facets is very flexible for complex folded and faulted surfaces. Classical Cartesian grids are not appropriate for surfaces with significant folding, for example, with multiple Z values for a given X-Y location. The gOcad software [241] implements triangulated grids and can easily handle such complexity.

A tetrahedra-based gridding scheme is used for geological cells bounded by triangulated surface grids. This tetrahedra-based topology is not yet used in geostatistical calculations for a number of reasons:

- The tetrahedra do not all have the same volume. This makes it more difficult to apply geostatistical tools and integrate hard and soft data.
- That topology is less commonly accepted by finite-difference scale-up and flow simulation programs.

Presently in the oil industry Cartesian grids are used almost exclusively because of their direct relation to finite-difference flow scale-up and simulation programs. The complex geometry defined by the tetrahedra grids in gOcad is presently transformed to a Cartesian (U, V, W) grid for geostatistical calculations. The resulting models are then back transformed. Techniques are being developed for direct population of tetrahedra-based and other non-Cartesian grids.

3.2 Stratigraphic Correlation/Coordinates

Reservoirs consist of a series of genetically related strata that may be correlated over the areal extent of the reservoir [260]. The surface grids that define this geologic correlation correspond to a chrono-stratigraphic or sequence stratigraphic framework, that is, the bounding surfaces between the layers correspond to a specific geologic time that separates two different periods of deposition or a period of erosion followed by deposition. Although outside the scope of this book, it is essential to have a sound geological framework prior to geostatistical modeling. Some references on the determination and importance of stratigraphic frameworks include [19, 263, 264].

Figure 3.6 shows a photograph of channel features at an outcrop of deep sea sediments (turbidite). The area photographed is about 69m long by 43m high. The lower photo shows geological surfaces that have been identified at the outcrop; four (partial) geological layers are defined. Figure 3.7 shows two possible Cartesian gridding schemes. The top scheme is appropriate for modeling the geological details. The simpler bottom scheme may be enough for larger-scale 3-D modeling. The procedure to establish the "stratigraphic" coordinates on the top of Figure 3.7 are presented below.

Reservoirs are made up of a number of reservoir layers. Each layer corresponds to a particular time period in the creation of the reservoir. The surfaces that separate these layers relate to a significant geological change. The definition of a "layer" is ambiguous. In the present context, a geological layer is a reservoir unit between 5 to 100 ft thick that can be correlated between multiple wells and mapped over a substantial areal extent. The layers are defined to provide a large-scale subdivision of the reservoir into geologically homogeneous units. Most reservoirs are divided into 5 to 10 layers. These layers are sometimes called "sequences" after sequence stratigraphy; however, the layers considered in geostatistical modeling need not correspond directly with the formal definition of geological sequences.

Each layer is defined by existing top and base surface grids: $z_{et}(x, y)$ and $z_{eb}(x, y)$ where x and y are areal coordinates, *et* refers to existing or present top and *eb* refers to existing (present) base. These surface grids are not flat because of differential compaction and subsequent structural deformation. Also, the layer may have been eroded by later depositional events or the sediments may have "filled" existing topography. The continuity of the facies and reservoir properties within a layer do not necessarily follow grids based

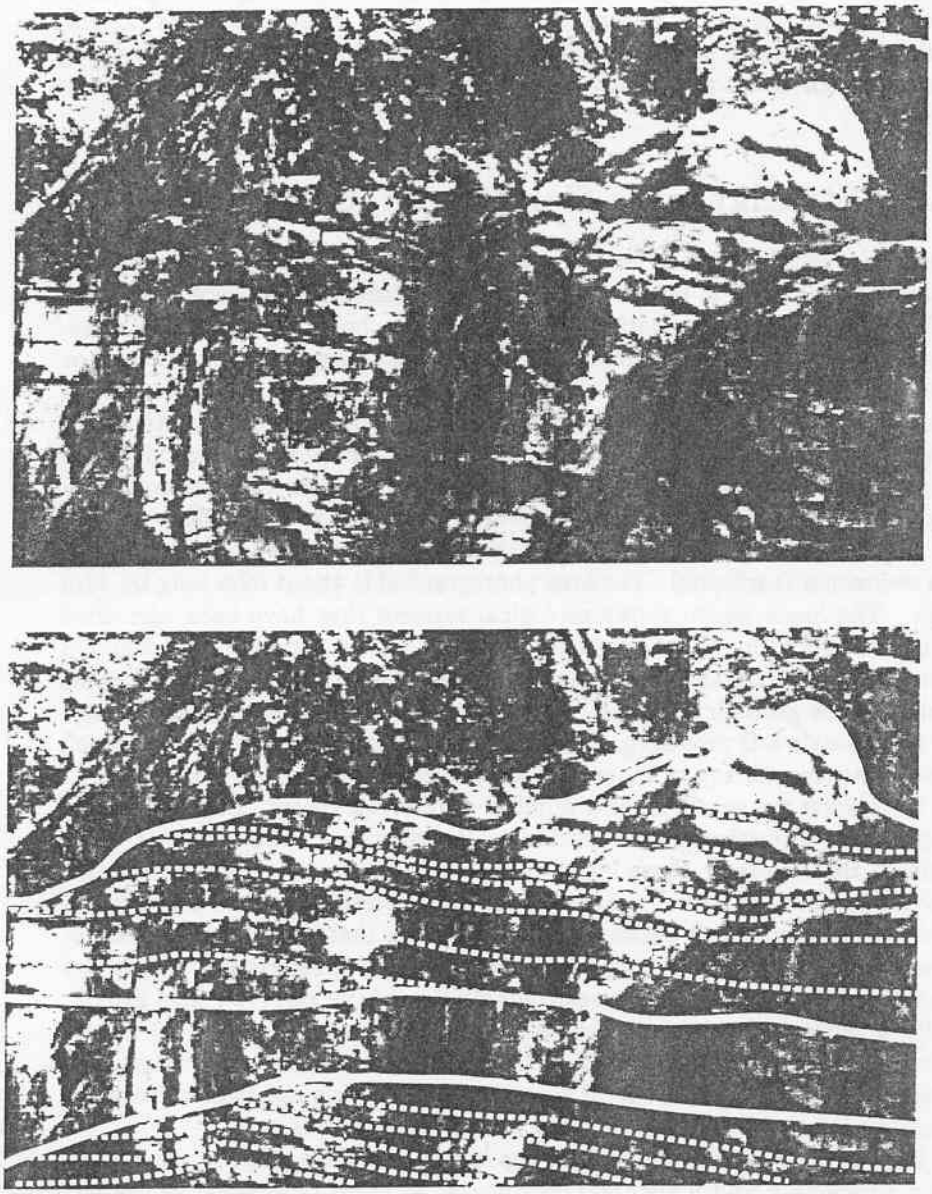


Figure 3.6: A photograph of a channel features at an outcrop of deep sea sediments (turbidite). The area photographed is about 69 m long by 43 m high. The lower photo shows geological surfaces that have been identified at the outcrop.

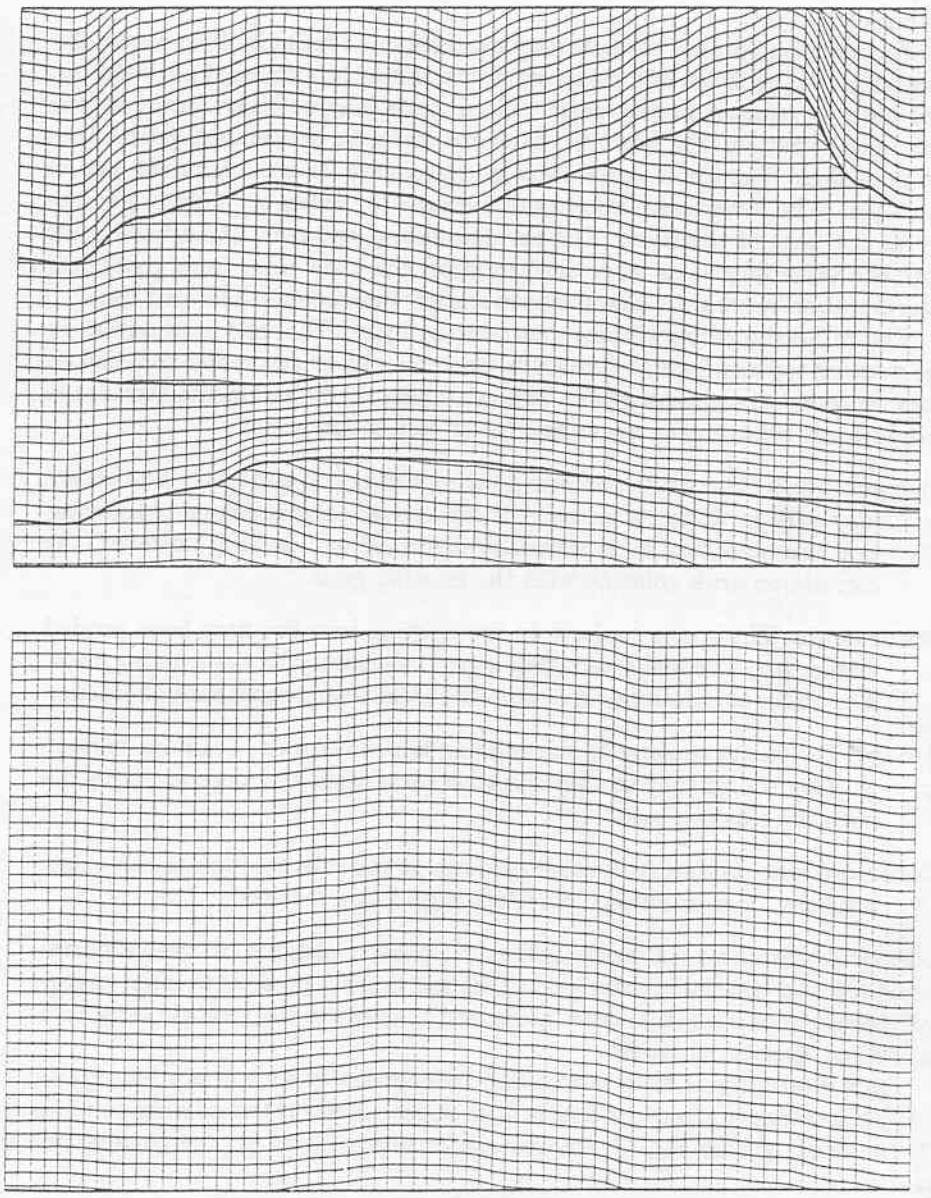


Figure 3.7: Two possible Cartesian gridding schemes for the geologic features in Figure 3.6. The top scheme is appropriate for modeling the geological details. The simpler bottom scheme may be enough for larger-scale 3-D modeling.

on the existing (present) bounding surfaces. Additional "correlation grids" may be required to define the stratigraphic continuity or "correlation style" within each layer.

A common approach to capturing stratigraphic continuity is to model the continuity of the facies and reservoir properties as proportional between top and base *correlation surface grids*. These grids may not correspond to any existing or past surface; however, they are useful for subsequent geostatistical modeling. The upper correlation grid accounts for erosion at the top of the layer. The lower correlation grid accounts for onlap or "filling" geometry at the base of the layer. The correlation top and correlation base grids have the same format as other surface grids, but they come from geological interpretation more then direct interpretation from seismic and well data.

Figure 3.8 shows four common correlation styles. The solid lines represent the existing top and base surface grids (same in all four cases) and the dashed lines represent correlation grids. The continuity of the sediments (strata) is proportional between the upper and lower correlation grids:

Proportional: The strata conform to the existing top and base. The strata may vary in thickness because of differential compaction or sedimentation rate and may be structurally deformed and faulted; however, the correlation grids coincide with the existing grids.

Truncation: The strata conform to the existing base but have been eroded at the top. The lower correlation grid coincides with the existing base. The upper correlation grid defines the areally varying amount of erosion.

Onlap: The strata conform to the existing top (no erosion) but have "filled" existing topography so that the base correlation grid does not coincide with the existing base.

Combination: The strata neither conform to the existing top nor the existing base. Two additional correlation grids are required.

A vertical coordinate will be defined as the relative distance between a correlation top and correlation base grid. This will make it possible to infer "natural" measures of horizontal correlation and to preserve the geologic structure in the final numerical model.

Each stratigraphic layer is modeled independently with a relative stratigraphic coordinate z_{rel} derived from four surface grids (1) the existing top z_{et}, (2) the existing base z_{eb}, (3) the correlation top z_{ct}, and (4) the correlation base z_{cb}:

$$z_{rel} = \frac{z - z_{cb}}{z_{ct} - z_{cb}} \cdot T \tag{3.4}$$

The coordinate z_{rel} is 0 at the stratigraphic base and T at the stratigraphic top. T is a thickness constant chosen to yield reasonable distance units for the z_{rel} coordinate; most commonly, T is the average thickness of the layer using the z_{et} and z_{eb} surface grids.

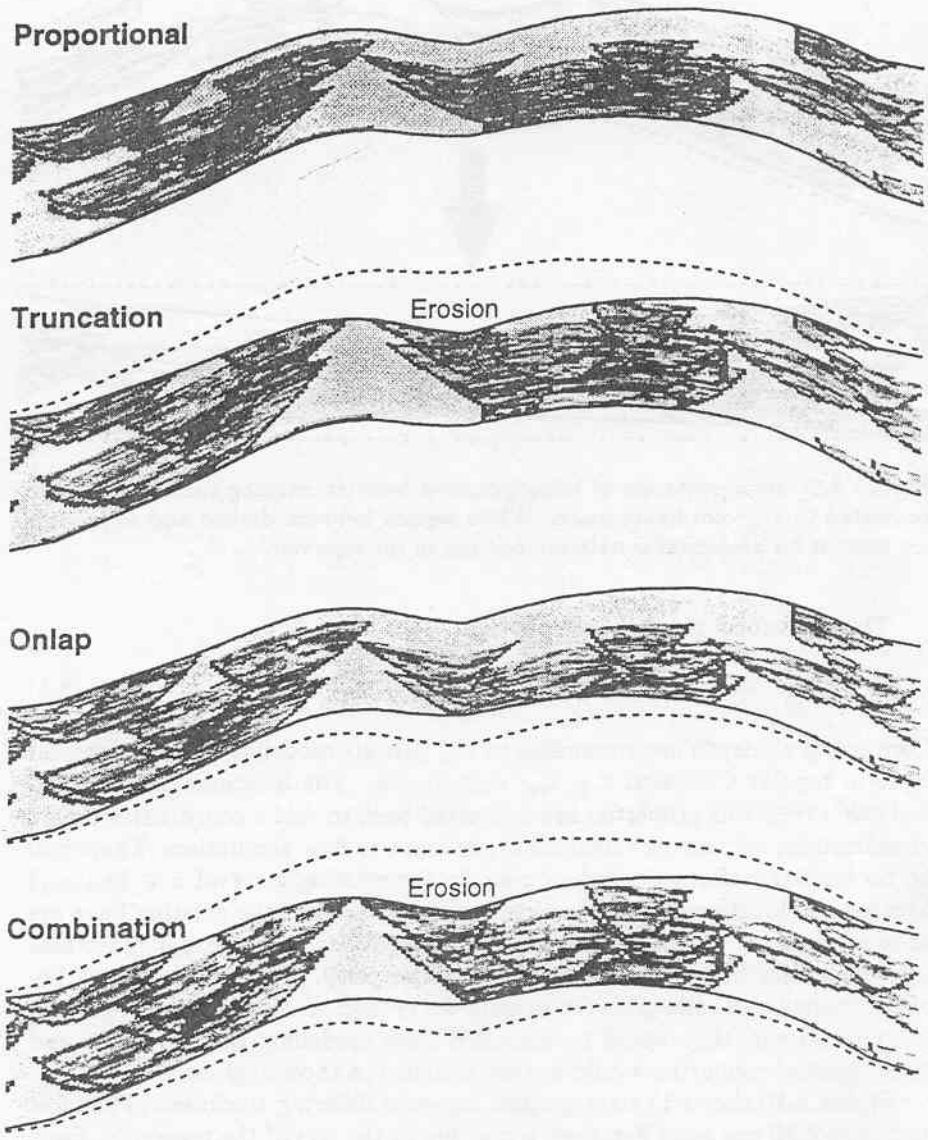


Figure 3.8: Example of different stratigraphic correlation styles. Internal channels illustrate proportional deformation, erosion, and onlap. Note that the solid lines, representing the existing reservoir layer, are the same in all four cases. The dashed lines in the cases of truncation, onlap, and combination represent geological correlation grids.

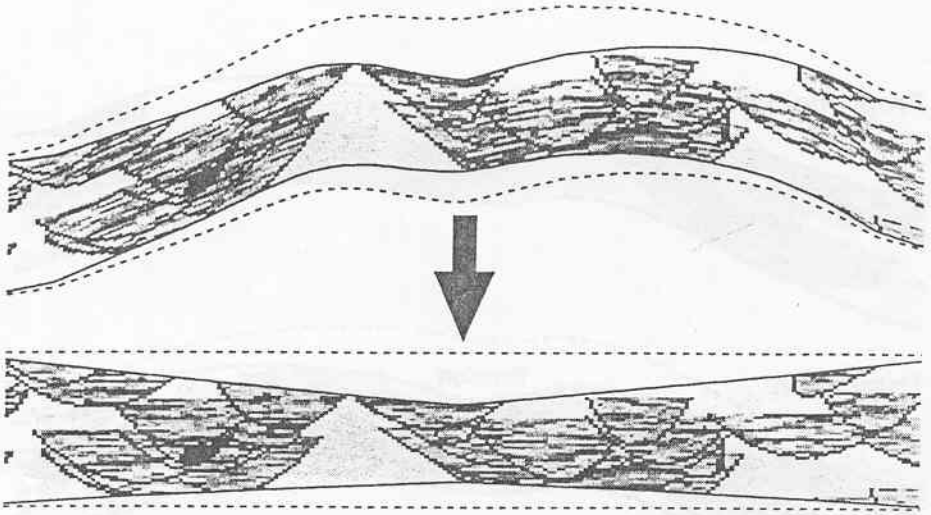


Figure 3.9: Representation of heterogeneities between existing surface grids and converted to z_{rel} coordinate space. White regions between dashed and solid grids are present for geological correlation but not in the reservoir.

This transform may be reversed by:

$$z = z_{cb} + \frac{z_{rel}}{T} \cdot (z_{ct} - z_{cb}) \tag{3.5}$$

Converting all depth measurements to z_{rel} permits modeling of each reservoir layer in regular Cartesian x, y, z_{rel} coordinates. The locations of the facies and other reservoir properties are converted back to real z coordinates before visualization, volumetric calculations, or input to flow simulation. There will be no back-transformed z-values outside the existing interval $z \notin (z_{eb}, z_{et})$ because the locations above the existing top and below the existing base are known ahead of time and excluded from modeling. Figure 3.9 illustrates heterogeneities between existing and restored grids. The white regions between dashed and solid grids are present for geological correlation but are not in the reservoir; they would be excluded from modeling, that is, facies and petrophysical properties would not be assigned in those regions.

Figure 3.10 shows 11 stratigraphic layers of differing thickness. The view in Figure 3.10 has been flattened according to the top of the reservoir. Each stratigraphic layer has a different stratigraphic correlation style.

Figure 3.11 shows four stratigraphic layers of differing thickness. The view in Figure 3.11 is of an actual outcrop.

Straightening Functions

As stated above, two motivations for coordinate rotation and stratigraphic coordinate transformation are to (1) allow the calculation of *natural* "hor-

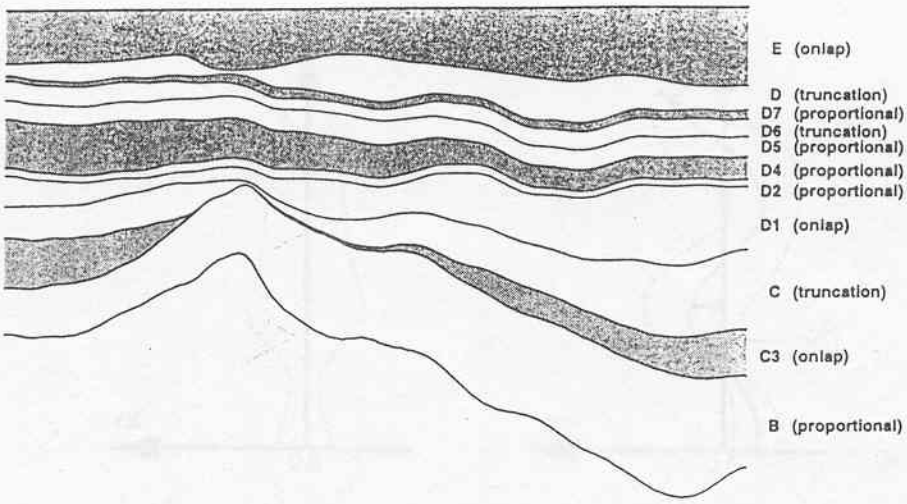


Figure 3.10: A cross section through a reservoir with 11 different stratigraphic layers of differing thickness and correlation style.

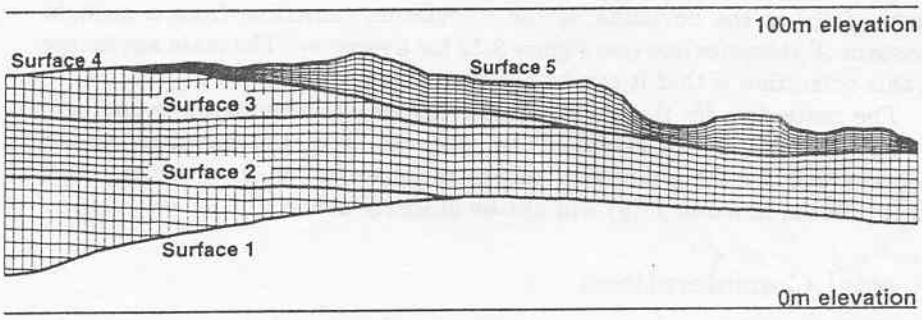


Figure 3.11: A cross section through an outcrop with four different stratigraphic layers of differing thickness and correlation style. The correlation styles cannot be seen on the figure since no within-layer properties are shown.

izontal" measures of correlation, and (2) to reflect the geological structure in the final numerical model. Other transformations could be considered in cases where there is large-scale undulation or curvature that is difficult to honor with geostatistical modeling techniques that assume a single direction of continuity. As an example, Figure 3.12 illustrates the *straightening* of a large channel or valley-type structure.

The Y coordinate along the primary direction of continuity is left unchanged. The X coordinate is corrected to be the distance from the centerline, that is, for any point x, y the new X coordinate is defined as:

$$x_1 = x - f^c(y) \tag{3.6}$$

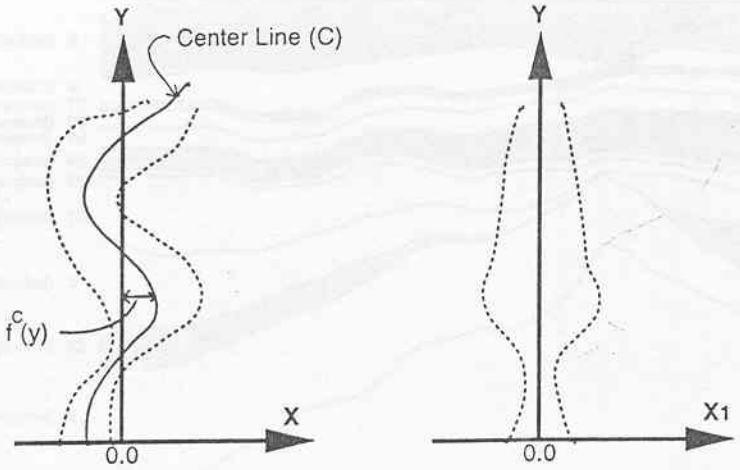


Figure 3.12: Transformation from the channel direction to a channel coordinate system following the sinuous channel centerline.

where $f^c(y)$ is the deviation of the undulating centerline from a straight constant-X reference line (see Figure 3.12 for a picture). The main advantage of this correction is that it can be easily reversed.

The center line for the large scale undulation *must* be known beforehand for all y values over the extent of the reservoir. Also, the curvature cannot be too great; the main direction cannot meander too much or reverse itself otherwise the function $f^c(y)$ will not be uniquely defined.

Special Considerations

The z_{rel} vertical coordinate requires top and base correlation grids over the areal extent of the reservoir layer. These are poorly defined when the stratigraphic dip is "steep." Figure 3.13 shows a schematic example with fluid contacts defining the layer of interest. It may be more straightforward to simply use the original Z vertical coordinate and consider a dip angle in subsequent geostatistical modeling steps.

Top and base correlation surface grids for Figure 3.13 would be separated by a large distance. The top correlation grid would start close to the existing top at the left hand side; however, it would be at a very high elevation at the right hand side. The base correlation grid would have the same "tilt" as the correlation top, but it would be at a very low elevation at the left hand side. The separation between the two grids could be very large. Constant elevation grids aligned with the OWC or GOC could be used for the top and base correlation grids. These grids would only be separated by the thickness of the reservoir. As indicated above, the direction of greatest continuity would not be "horizontal" in the x, y, z_{ref} coordinate space; strike and dip angles would need to be considered in subsequent modeling.

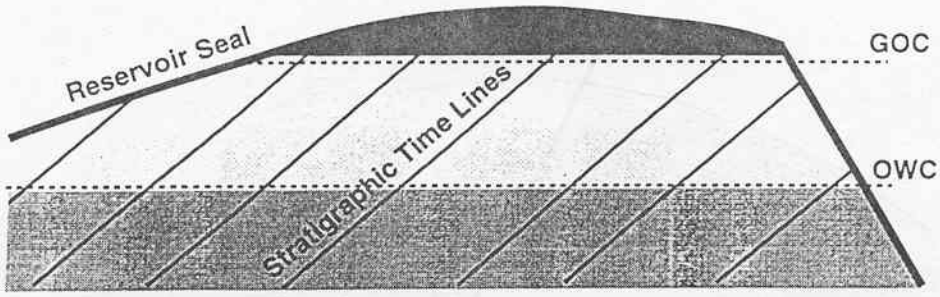


Figure 3.13: Illustration of *steeply* dipping (a steep dip may be only 1 to 5 degrees; shown here with significant vertical exaggeration) strata with gas oil (GOC) and water oil (WOC) contact.

3.3 Faults

The problems presented by faults are diverse, challenging, and somewhat outside the scope of this book. Straightforward normal faults, such as illustrated by the cross section in Figure 3.14, can be handled by correlation grids and stratigraphic coordinate transformation as described above.

Reverse faults cause problems because there are multiple surface values at some x, y locations, see Figure 3.15. One solution to this problem is to assume the faults are vertical through each geologic layer. Then, we are back to the case of simply requiring four surface grids for each layer. Such fault "verticalization" appears crude but is a reasonable approximation if the faults are relatively steep and the offset is not too great. As an example, Figure 3.16 shows a schematic illustration of a geologic layer 50 ft thick and 2000 ft across with a normal fault at 70°with an offset of 150 ft. The horizontal separation is $cos(70) \cdot 150 = 51.3$ ft, which is small relative to the total areal extent of the geological layer. Assuming a vertical fault in the center would cause a mismatch of less than 26 ft. The areal size of geological modeling cells often exceeds this size.

There are cases, in thick formations with significant offset, where fault verticalization does not adequately capture the geometry of the reservoir layers. There are a number of solutions (1) each fault block can be modeled separately, or (2) the horizontal coordinate can be expanded to remove overlaps. The first alternative is simpler and does not distort horizontal distances. The correlation of petrophysical properties across fault blocks can be enforced by sequential modeling where results from previously simulated fault blocks are considered. The procedure in more detail is as follows:

- Identify fault locations and model the fault surfaces.
- Flag all geologic modeling cells and well data by fault block.
- Construct Z_{rel} coordinate system in each fault block.

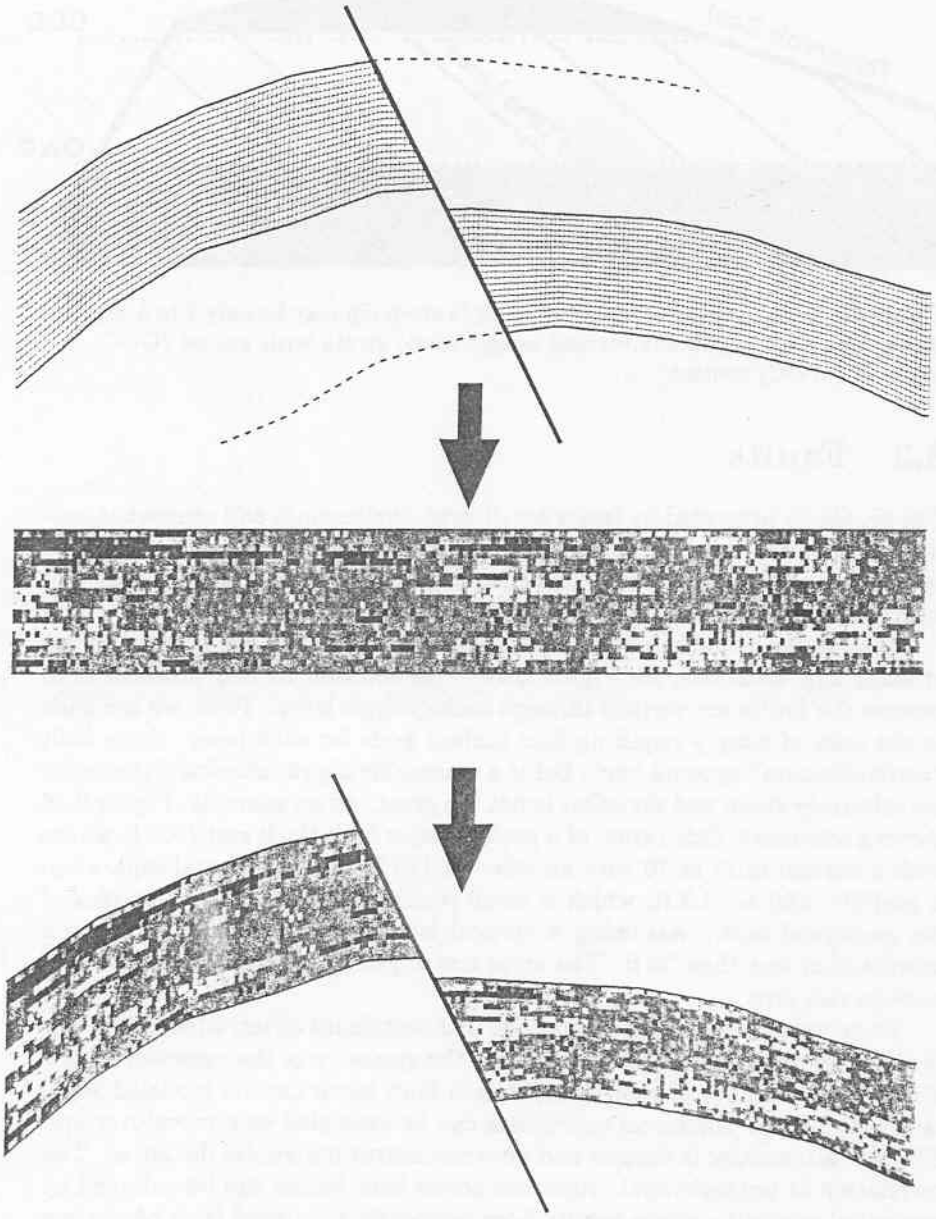


Figure 3.14: Modeling a faulted stratigraphic layer: (i) top figure shows the fault location and the existing surfaces extended as restored surfaces, (ii) the middle figure shows the restored grid with assigned petrophysical properties, and (iii) the bottom figure shows the properties within the existing faulted grids.

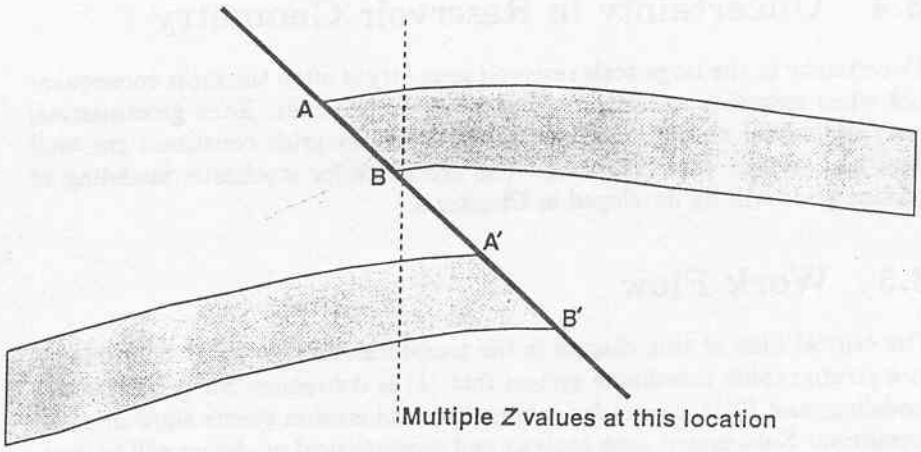


Figure 3.15: An example of a reverse fault illustrating the need for new horizontal coordinates; multiple z-coordinate values are needed for some locations.

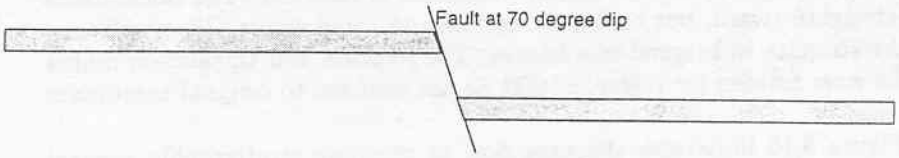


Figure 3.16: A faulted geological layer shown with no vertical exaggeration. Assuming the fault vertical would have little effect on the final geological model.

- Construct a geostatistical model of lithofacies, porosity, and permeability (as described in later chapters) for each fault block. Data from previously modeled fault blocks could be considered if correlated to the fault block currently being considered.

The two alternatives presented above will be used throughout this text, that is, (1) calculate Z_{rel} with modified restored grids or vertical faults, or (2) model each fault block separately.

It is important to note that geostatistical methods are *not* limited to Cartesian geometry and gridding. In fact, at the time of writing this book, research into the direct geostatistical modeling of facies and properties on unstructured grids is very active. The future of geostatistical reservoir modeling will be to model the complex geometry directly rather than with the coordinate "fixes" that have been introduced here.

3.4 Uncertainty in Reservoir Geometry

Uncertainty in the large-scale reservoir geometry is often the most consequential when assessing uncertainty in reservoir volumetrics. Each geostatistical reservoir model should consider stochastic surface grids consistent the well data and seismic data [121, 143]. The procedure for stochastic modeling of surface grids will be developed in Chapter 5.

3.5 Work Flow

The central idea of this chapter is the transformation of "real" coordinates to a stratigraphic coordinate system that (1) is convenient for geostatistical modeling, and (2) accounts for deformation and erosion events since original deposition. Subsequent data analysis and geostatistical modeling will be performed in the stratigraphic coordinates system. Finally, the model can be visualized and put to use back in the original coordinates. A number of work flows are related to this transformation.

Figure 3.17 illustrates the work flow to rotate and translate X, Y, Z coordinates to be aligned with the overall orientation of reservoir. The calculations are straightforward, but care must be taken to avoid errors. Most software would visualize in original coordinates. The rotation and translation makes model sizes smaller for reservoirs that do not conform to original coordinate systems.

Figure 3.18 illustrates the work flow to calculate stratigraphic vertical coordinates. Up to two existing surface grids and two restored stratigraphic grids are needed to get the data in the correct stratigraphic coordinate for modeling. Recall that the restored grids are often shifted, that is, vertically translated, existing grids.

Figure 3.19 illustrates the work flow to apply large-scale straightening function. Large-scale curvilinear continuity may need to be straightened.

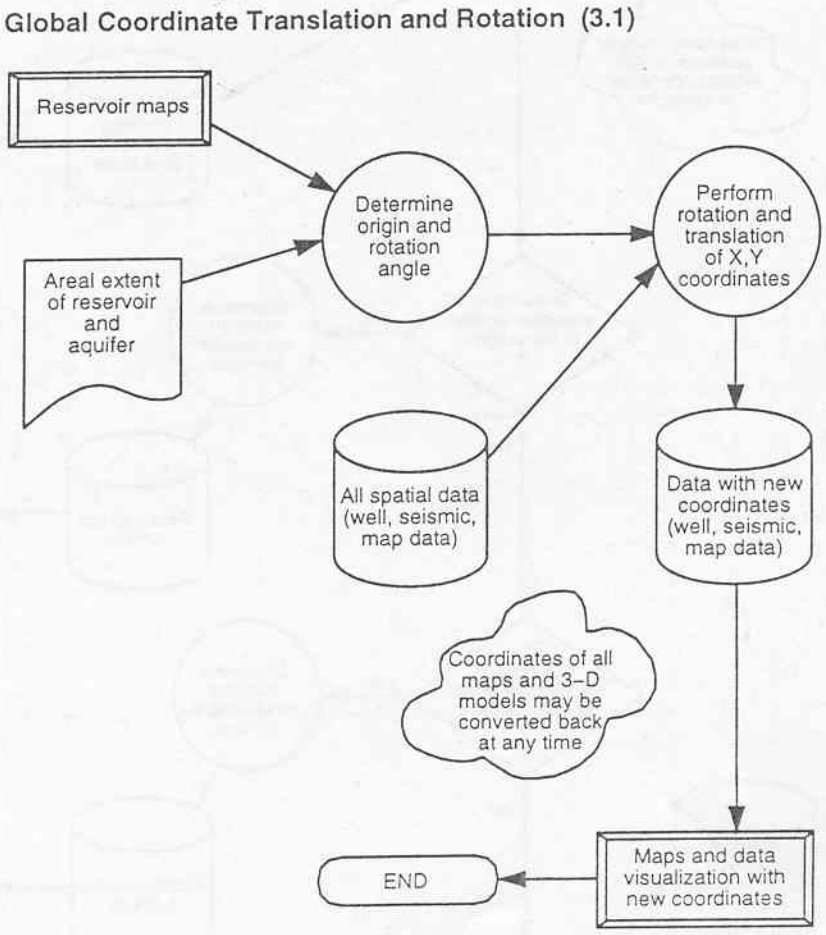


Figure 3.17: Work flow for global coordinate rotation and translation.

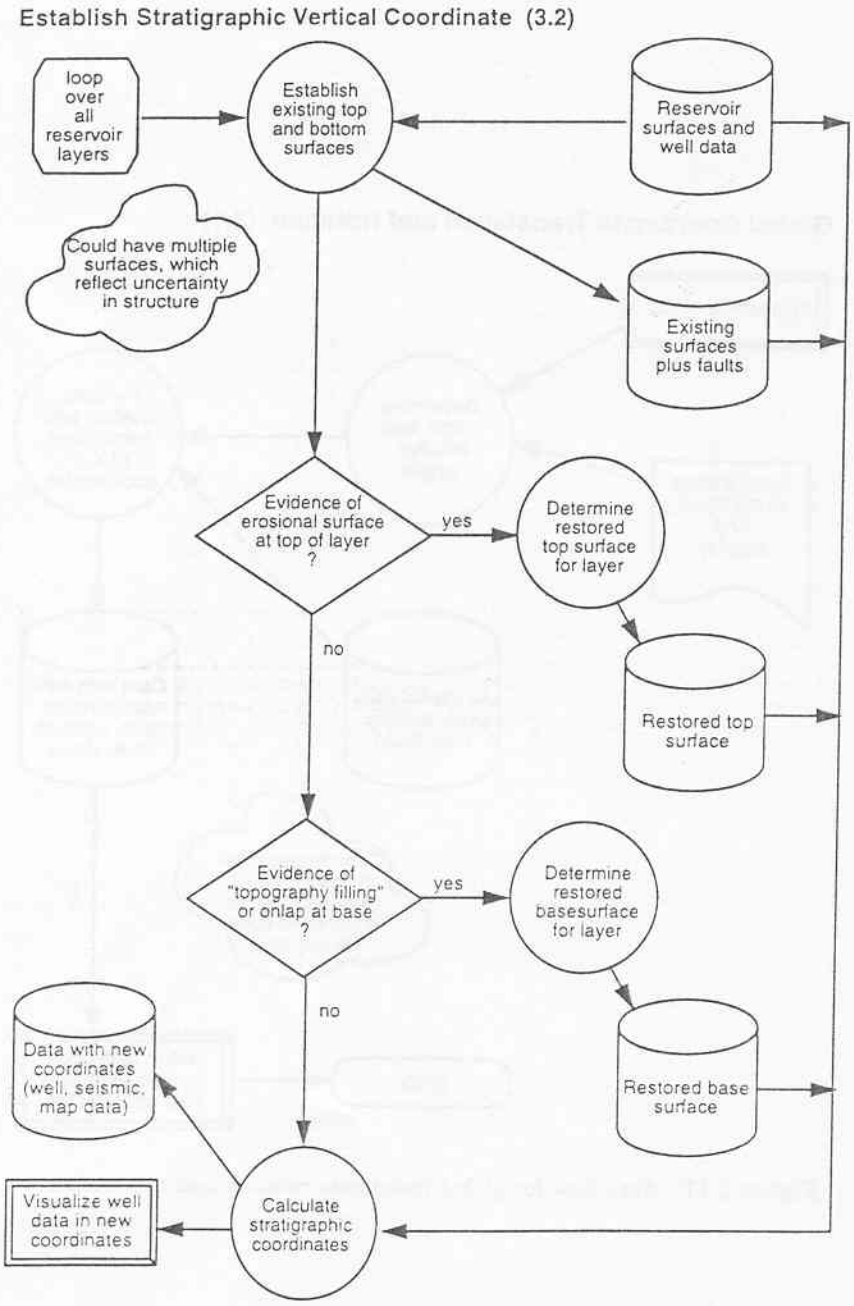


Figure 3.18: Work flow for calculation of vertical stratigraphic coordinate.

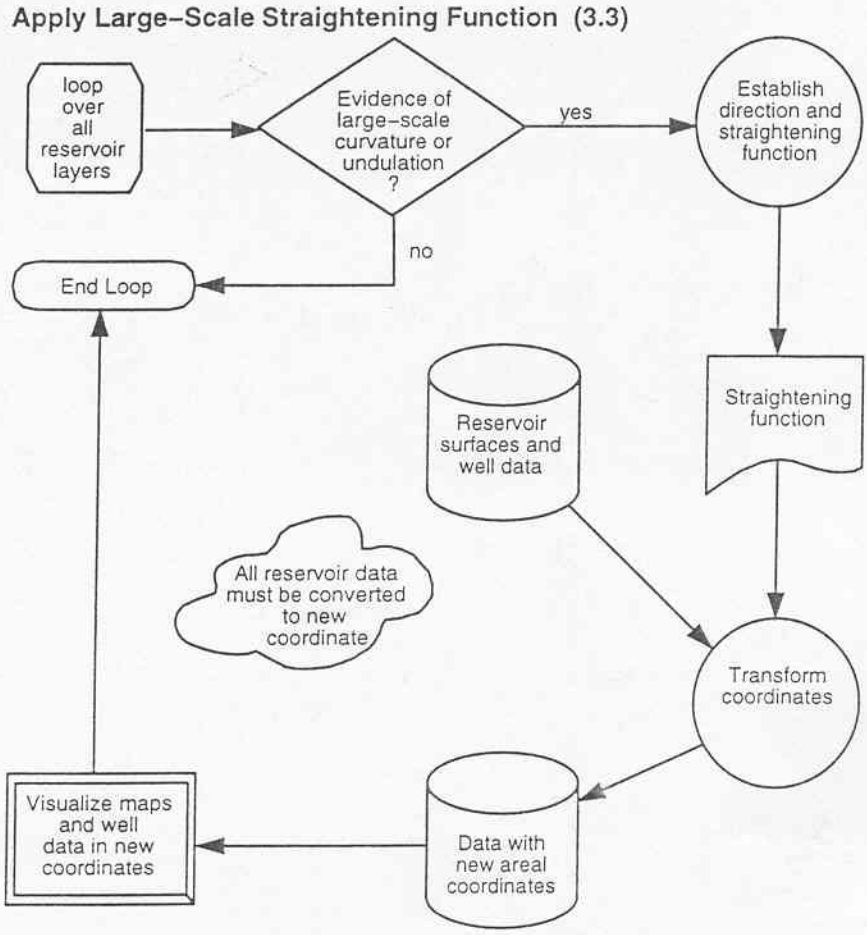


Figure 3.19: Work flow to apply large-scale straightening function.

Chapter 4

Quantifying Spatial Correlation

An essential aspect of geostatistical modeling is to establish quantitative measures of spatial correlation to be used for subsequent estimation and simulation. Spatial variability is different for each variable (facies indicators, porosity, and permeability) within each reservoir layer.

The spatial correlation for object-based modeling is quantified through object shapes, sizes, and relationships. These spatial measures are inseparable from the object-based modeling approach, see Chapter 7.

The variogram is the most commonly used measure of spatial correlation for cell-based facies modeling, porosity, and permeability modeling. The random function (RF) concept and theoretical background for variogram-based measures of spatial correlation are presented in Section 4.1. This section could be skipped at first reading. Section 4.2 describes the steps involved in variogram calculation. Careful selection of variogram calculation parameters is critical for obtaining a clean, interpretable, sample variogram.

Principles of variogram interpretation are described in Section 4.3. The effect of stratigraphic cyclicity, vertical and areal trends, and stratification are described in detail. Often, there are inadequate data to reliably calculate and interpret a reservoir-specific variogram. Section 4.4 presents analogue data and procedures to supplement limited well data by more abundant "soft" geological data. Modeling experimental variograms, consistent with soft geological data, is a necessary step prior to kriging-based techniques; Section 4.5 describes variogram modeling in detail.

Section 4.6 presents the calculation, interpretation, and modeling of cross variograms between multiple variables. Depending on software and data availability, inference and modeling of cross variograms can be tedious. Approximations as provided by the various Markov-type models and the model implicit to collocated cokriging are discussed.

Finally, Section 4.7 presents work flow diagrams for some operations related to variogram calculation and modeling.

4.1 The Random Function Concept

The material in this section is more completely covered in other geostatistical texts such as *Mining Geostatistics* [157], *An Introduction to Applied Geostatistics* [138], or *Geostatistics for Natural Resources Evaluation* [110]. Nevertheless, this section is included for completeness. It may be skipped at first reading. These details will become more relevant to answer questions such as "why are conventional geostatistical methods limited to two-point variogram statistics?"

The uncertainty about an unsampled value z is modeled through the probability distribution of a random variable (RV) Z. The probability distribution of Z after data conditioning is usually location dependent; hence, the notation $Z(\mathbf{u})$, with $\mathbf{u}$ being the coordinate location vector. A random function (RF) is a set of RVs defined over some field of interest, for example, $\{Z(\mathbf{u}), \mathbf{u} \in \text{ study area } A\}$ also denoted simply as $Z(\mathbf{u})$. Usually the RF definition is restricted to RVs related to the same attribute , say z, hence, another RF would be defined to model the spatial variability of a second attribute, say $\{Y(\mathbf{u}), \mathbf{u} \in \text{ study area}\}$.

Just as a RV $Z(\mathbf{u})$ is characterized by its cdf, a RF $Z(\mathbf{u})$ is characterized by the set of all its N-variate cdfs for any number N and any choice of the N locations $\mathbf{u}_i, i = 1, \ldots, N$ within the study area A:

$$F(\mathbf{u}_1, \ldots, \mathbf{u}_N; z_1, \ldots, z_N) = Prob\{Z(\mathbf{u}_1) \leq z_1, \ldots, Z(\mathbf{u}_N) \leq z_N\} \quad (4.1)$$

Just as the univariate cdf of the RV $Z(\mathbf{u})$ is used to characterize uncertainty about the value $z(\mathbf{u})$, the multivariate cdf (4.1) is used to characterize joint uncertainty about the N values $z(\mathbf{u}_1), \ldots, z(\mathbf{u}_N)$.

The bivariate ($N = 2$) cdf of any two RVs $Z(\mathbf{u}_1)$, $Z(\mathbf{u}_2)$, or more generally $Z(\mathbf{u}_1)$, $Y(\mathbf{u}_2)$, is particularly important since conventional geostatistical procedures are restricted to univariate ($F(\mathbf{u}; z)$) and bivariate distributions:

$$F(\mathbf{u}_1, \mathbf{u}_2; z_1, z_2) = Prob\{Z(\mathbf{u}_1) \leq z_1, Z(\mathbf{u}_2) \leq z_2\} \quad (4.2)$$

One important summary of the bivariate cdf $F(\mathbf{u}_1, \mathbf{u}_2; z_1, z_2)$ is the covariance function defined as:

$$C(\mathbf{u}_1, \mathbf{u}_2) = E\{Z(\mathbf{u}_1)Z(\mathbf{u}_2)\} - E\{Z(\mathbf{u}_1)\}E\{Z(\mathbf{u}_2)\} \quad (4.3)$$

However, when a more complete summary is needed, the bivariate cdf $F(\mathbf{u}_1, \mathbf{u}_2; z_1, z_2)$ can be described by considering binary *indicator* transforms of $Z(\mathbf{u})$ defined as

$$I(\mathbf{u}; z) = \begin{cases} 1, & \text{if } Z(\mathbf{u}) \leq z \\ 0, & \text{otherwise} \end{cases} \quad (4.4)$$

Then, the previous bivariate cdf (4.2) at various thresholds z_1 and z_2 appears as the non-centered covariance of the indicator variables:

$$F(\mathbf{u}_1, \mathbf{u}_2; z_1, z_2) = E\{I(\mathbf{u}_1; z_1) I(\mathbf{u}_2; z_2)\} \tag{4.5}$$

Relation (4.5) is the key to the indicator geostatistics formalism [145]: it shows that inference of bivariate cdfs can be done through sample indicator covariances.

The probability density function (pdf) representation is more relevant for categorical variables. For example,

$$\begin{aligned} f(\mathbf{u}_1, \mathbf{u}_2; k_1, k_2) &= Prob\{Z(\mathbf{u}_1) \in k_1, Z(\mathbf{u}_2) \in k_2,\} \\ & \quad k_1, k_2 = 1, \ldots, K \end{aligned} \tag{4.6}$$

is the bivariate or two-point distribution of $Z(\mathbf{u}_1)$ and $Z(\mathbf{u}_2)$. This two-point distribution, when established from experimental proportions, is also referred to as a *two-point histogram* [87].

Recall that a categorical variable $Z(\mathbf{u})$, which takes K outcome values $k = 1, \ldots, K$, may arise from a naturally occurring categorical variable or from a continuous variable discretized into K classes.

The purpose for conceptualizing a RF as $\{Z(\mathbf{u}), \mathbf{u} \in \text{study area } A\}$ is never to study the case where the variable Z is completely known. If all the $z(\mathbf{u})$'s were known for all $\mathbf{u} \in$ study area A, there would be neither any problem left nor any need for the concept of a random function. The ultimate goal of an RF model is to make some predictive statement about locations $\mathbf{u}$ where the outcome $z(\mathbf{u})$ is unknown.

Inference of any statistic requires some repetitive sampling. For example, repetitive sampling of the variable $z(\mathbf{u})$ is needed to evaluate the cdf

$$F(\mathbf{u}; z) = Prob\{Z(\mathbf{u}) \leq z\}$$

from experimental proportions. However, in many applications at most one sample is available at any single location $\mathbf{u}$ in which case $z(\mathbf{u})$ is known (ignoring sampling errors), and the need to consider the RV model $Z(\mathbf{u})$ vanishes. The paradigm underlying statistical inference processes is to trade the unavailable replication at location $\mathbf{u}$ for another replication available somewhere else in space and/or time. For example, the cdf $F(\mathbf{u}; z)$ may be inferred from the sampling distribution of z-samples collected at other locations, $\mathbf{u}_\alpha \neq \mathbf{u}$, within the same field.

This trade of replication corresponds to the decision of stationarity. Stationarity is a property of the RF model, not of the underlying physical spatial distribution. Thus, it cannot be checked from data. The decision to pool data into statistics across rock types is not refutable a priori from data; however, it can be shown inappropriate a posteriori if differentiation per rock type is critical to the undergoing study. For a more extensive discussion see [138, 146].

The RF $\{Z(\mathbf{u}), \mathbf{u} \in A\}$ is said to be stationary within the field A if its multivariate cdf (4.1) is invariant under any translation of the N coordinate vectors $\mathbf{u}_k$, that is:

$$F(\mathbf{u}_1, \ldots, \mathbf{u}_N; z_1, \ldots, z_N) = F(\mathbf{u}_1 + \mathbf{l}, \ldots, \mathbf{u}_n + \mathbf{l}; z_1, \ldots, z_n),$$
$$\forall \text{ translation vector } \mathbf{l}. \tag{4.7}$$

Invariance of the multivariate cdf entails invariance of any lower order cdf, including the univariate and bivariate cdfs, and invariance of all their moments, including all covariances of type (4.3) or (4.5). The decision of stationarity allows inference. For example, the unique stationary cdf

$$F(z) = F(\mathbf{u}; z), \forall\, \mathbf{u} \in A$$

can be inferred from the cumulative sample histogram of the z-data values available at various locations within A. The stationary mean and variance can then be calculated from that stationary cdf $F(z)$:

$$E\{Z(\mathbf{u})\} = \int z dF(z) = m, \quad \forall\, \mathbf{u}$$

$$E\{[Z(\mathbf{u}) - m]^2\} = \int [z - m]^2\, dF(z) = \sigma^2, \quad \forall\, \mathbf{u}$$

The decision of stationarity also allows inference of the stationary covariance

$$C(\mathbf{h}) = E\{Z(\mathbf{u}+\mathbf{h})Z(\mathbf{u})\} - [E\{Z(\mathbf{u})\}]^2$$
$$\forall \mathbf{u}, \mathbf{u}+\mathbf{h} \in A \tag{4.8}$$

from the sample covariance of all pairs of z-data values approximately separated by vector $\mathbf{h}$. At $\mathbf{h} = 0$ the stationary covariance $C(0)$ equals the stationary variance σ^2, that is,

$$\begin{aligned} C(0) &= E\{Z(\mathbf{u}+0)Z(\mathbf{u})\} - [E\{Z(\mathbf{u})\}]^2 \\ &= E\{Z(\mathbf{u})^2\} - [E\{Z(\mathbf{u})\}]^2 \\ &= Var Z(\mathbf{u}) = \sigma^2 \end{aligned}$$

The notation $C(0)$ will often be used for the variance.

In certain situations the standardized stationary correlogram is preferred

$$\rho(\mathbf{h}) = \frac{C(\mathbf{h})}{C(0)}$$

In other cases another second order (two-point) moment called the variogram is considered:

$$2\gamma(\mathbf{h}) = E\left\{[Z(\mathbf{u}+\mathbf{h}) - Z(\mathbf{u})]^2\right\}$$
$$\forall \mathbf{u}, \mathbf{u}+\mathbf{h} \in A \tag{4.9}$$

Under the decision of stationarity the covariance, correlogram, and variogram are equivalent tools for characterizing two-point correlation:

$$C(\mathbf{h}) = C(0) \cdot \rho(\mathbf{h}) = C(0) - \gamma(\mathbf{h}) \tag{4.10}$$

This relation depends on the stationarity decision implying that the mean and variance are constant and independent of location. These relations are the foundation of variogram interpretation. That is, (1) the sill plateau value of the stationary variogram is the variance, which is the variogram value that corresponds to zero correlation, (2) the correlation between $Z(\mathbf{u})$ and $Z(\mathbf{u}+\mathbf{h})$ is positive when the variogram value is less than the sill, and (3) the correlation between $Z(\mathbf{u})$ and $Z(\mathbf{u}+\mathbf{h})$ is negative when the variogram exceeds the sill.

These three principles depend on knowledge of the "model" variance σ^2, which must be finite. The variance of petrophysical properties in petroleum reservoirs is indeed finite. Moreover, the variance is known more precisely than any experimental variogram value; therefore, the experimental variance may be associated to the model variance. The bootstrap procedure of Chapter 2 could be considered to assess uncertainty in the variance.

h-scatterplots may be used to illustrate these principles. An h-scatterplot is a cross plot of the "head" value $Z(\mathbf{u})$ with the "tail" value $Z(\mathbf{u}+\mathbf{h})$. The use of the words "head" and "tail" relate to the vector $\mathbf{h}$. Figure 4.1 shows three h-scatterplots corresponding to three lag vectors on a typical semivariogram. The horizontal line at 1.0 on the semivariogram is at the model variance of 1.0.

The decision of stationarity is critical for the appropriateness and reliability of geostatistical simulation methods. Pooling data across geological facies may mask important geological differences; on the other hand, splitting data into too many subcategories may lead to unreliable statistics based on too few data per category. The rule in statistical inference is to pool the largest amount of *relevant* information to formulate predictive statements.

Stationarity is a property of the RF model; thus, the decision of stationarity may change if the scale of the study changes or if more data become available. If the goal of the study is global, then local details may be unimportant; conversely, the more data available the more statistically significant differentiation become possible.

4.2 Calculating Experimental Variograms

In probabilistic notation, the variogram is defined as the expected value (4.9):

$$2\gamma(\mathbf{h}) = E\left\{[Z(\mathbf{u}) - Z(\mathbf{u}+\mathbf{h})]^2\right\}$$

The variogram is $2\gamma(\mathbf{h})$. The *semi*variogram is one half of the variogram, that is, $\gamma(\mathbf{h})$. The word "variogram" is often mistakenly used for the semi-

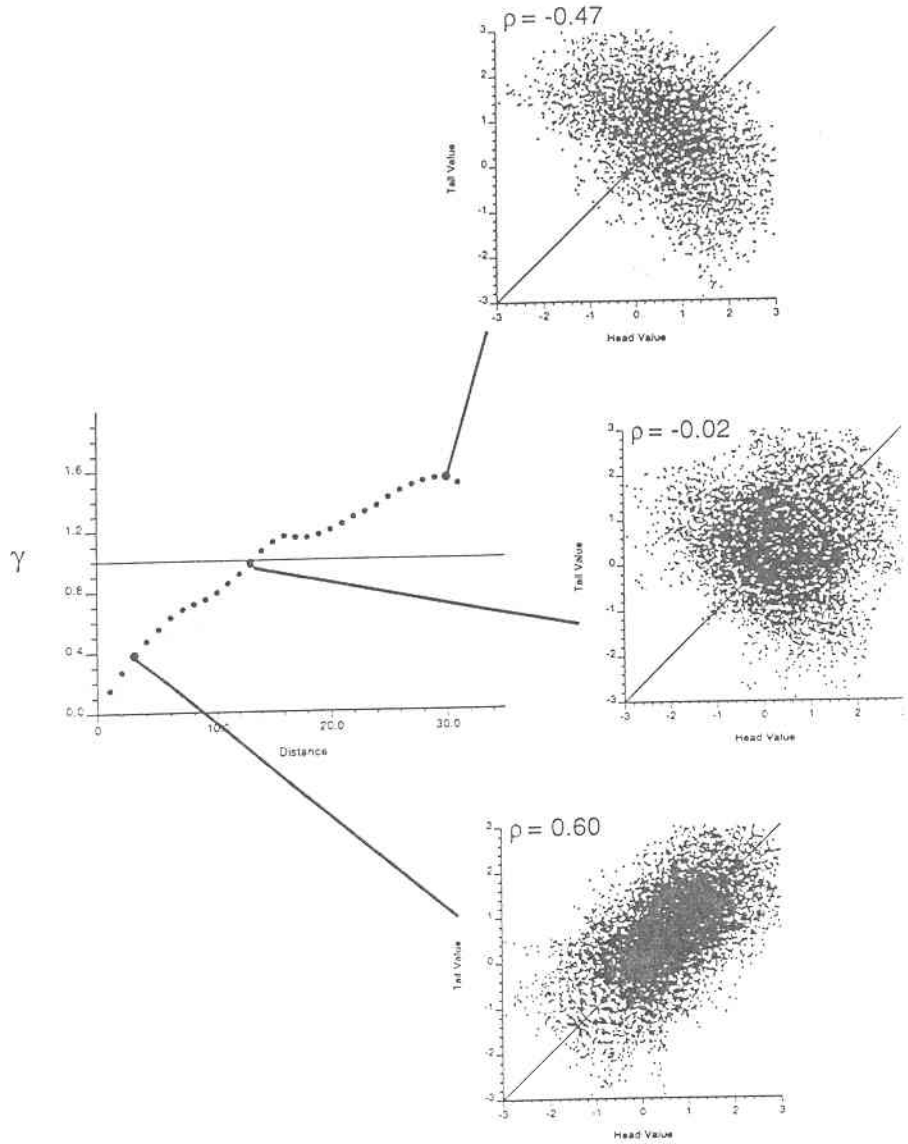


Figure 4.1: Semivariogram of the normal score of real porosity data with the h-scatterplots corresponding to three different lag distances. Note that the correlation on the h-scatterplot is positive when the semivariogram value is below the sill, zero when the semivariogram is at the sill, and negative when the semivariogram is above the sill.

variogram. Experimentally, the semivariogram for lag distance his defined as the average squared difference of values separated approximately by $\mathbf{h}$:

$$\hat{\gamma}(\mathbf{h}) = \frac{1}{2N(\mathbf{h})} \sum_{N(\mathbf{h})} [z(\mathbf{u}) - z(\mathbf{u}+\mathbf{h})]^2 \tag{4.11}$$

where $N(\mathbf{h})$ is the number of pairs for lag $\mathbf{h}$. Some questions that must be addressed before calculating experimental variograms:

- Does the data variable require transformation or the removal of an obvious trend? See Section 5.6, which follows.
- Do we have the correct geological or stratigraphic coordinate system for locations uand distance vectors $\mathbf{h}$? See below.
- What lag vectors hand associated tolerances should be considered? See below.

There are many different types of variogram-type measures of spatial variability that could be considered [15, 41, 62, 202, 227, 233]. The practice proposed here is use the traditional semivariogram with the correct data variable in the appropriate stratigraphic coordinate system. The theory behind kriging and simulation requires the use of either the covariance or the variogram. Some alternative variogram-type measures among the most robust to outlier values are only acceptable to help identify the range of correlation and anisotropy.

Establish the Correct Variable

Variogram calculation is preceded by selection of the Z variable to use in variogram calculation. The choice of the variable is evident in conventional kriging applications; however, data transformation is common in modern geostatistics. The use of Gaussian techniques requires a prior normal score transform of the data (see Section 2.4) and the variogram of those transformed data. Indicator techniques require an indicator coding of the data prior to variogram calculation (see Section 2.4).

The correct variable also depends on how trends are going to be handled in subsequent model building. Often, clear areal or vertical trends are removed prior to geostatistical modeling and then added to geostatistical models of the residual (original value minus trend). If this two-step modeling procedure is being considered, then the variogram of the residual data is required. There is a risk, however, of introducing artefact structures in the definition of the trend and residual data.

Another aspect of choosing the correct variable is outlier detection and removal. Extreme high and low data values have a large influence on the variogram since each pair is squared in variogram calculation. Erroneous data should be removed. Of greater concern are legitimate high values that may mask the spatial structure of the majority of the data. Logarithmic or normal

score transformation mitigates the effect of outliers, but are only suitable if an appropriate back transform is being considered in later geostatistical calculations.

Coordinate Transformation

The coordinate transformations presented in Chapter 3 are necessary before variogram calculation. In presence of vertical wells, the vertical variogram does not depend on the stratigraphic coordinate transform, as long as calculations are limited to data within the appropriate stratigraphic layer and facies type. The horizontal variogram, however, is very sensitive to the stratigraphic z_{rel} coordinate transform. Attempting variogram calculation prior to such stratigraphic coordinate transformation can lead the modeler to the erroneous conclusion that the data have no horizontal correlation. Data sparsity may also lead to the same erroneous conclusion.

A characteristic feature of geological phenomena is spatial correlation. Coordinate transform errors, sparse data, errors in calculation parameters, and many other factors could lead to the wrong conclusion that there is no spatial correlation. A pure-nugget model should not be adopted.

When the geological formation has been extensively folded, a more elaborate coordinate transform that allows following the curvilinear structure may be required [44]. The gOcad group has developed such elaborate structural unfolding schemes [180].

Data and coordinate transformation are necessary prerequisites to variogram calculation and interpretation. Once the data are prepared for variogram calculation, it is necessary to choose the distance lags, **h** values, to consider.

Choosing Variogram Directions and Lag Distances

Anisotropy

Variograms are rarely isotropic; geologic continuity and variogram continuity are direction dependent. In sedimentary structures continuity in the vertical direction is typically less than the horizontal direction. Moreover, horizontal continuity depends on the direction of deposition and subsequent diagenetic alteration. Directions of continuity are most often known from geological interpretation or preliminary contouring of the data. Rarely is it a good idea to "search" for the principal directions of continuity by calculating the variogram in many directions.

A critical first step is to identify the "vertical" direction. This direction is perpendicular to the time-stratigraphic correlation and often has the least continuity. This should normally be the Z_{rel} direction calculated with the considerations outlined in Chapter 3. One complication is when there are clinoform structures within the stratigraphic layer, see Figure 4.2. In this case

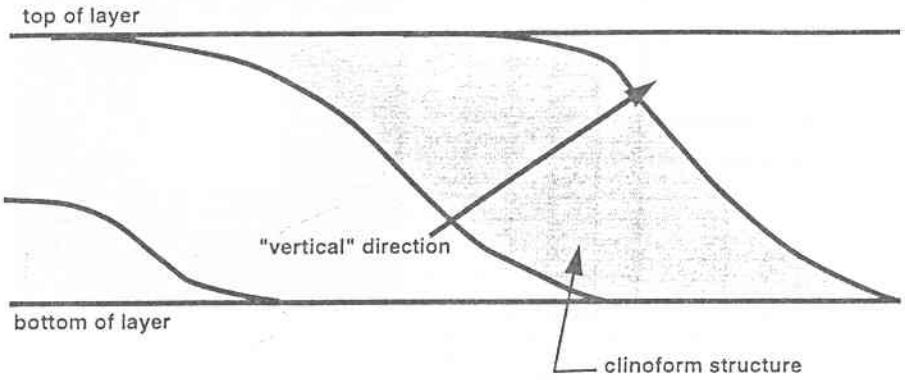


Figure 4.2: Flattened stratigraphic layer with clinoform structures. The geologic continuity follows the curvilinear directions within each clinoform.

the "vertical" directions will be at some azimuth and dip from the vertical direction.

Anisotropy or directional continuity in geostatistical calculations is typically *geometric*. Three angles define orthogonal x, y, and z coordinates and then the components of the distance vectors are scaled by three range parameters to determine the scalar distance, that is,

$$h = \sqrt{(h_x/a_x)^2 + (h_y/a_y)^2 + (h_z/a_z)^2}$$

where h_x, h_y, and h_z are the components of a vector $\mathbf{h}$ in 3-D coordinate space and a_x, a_y, and a_z are the scaling parameters in the principal directions. Contour lines of equal "distance" must follow ellipsoids. This concept of geometric anisotropy is covered in greater in Variogram Modeling (Section 4.5).

The anisotropy in petroleum reservoirs is defined by a single angle that identifies the "major" and "minor" horizontal directions of continuity. The vertical direction is then assumed perpendicular to the horizontal direction.

Most often, sound geological principles should be used to build a depositional and diagenetic understanding of the formation; then, the major and minor directions of continuity are evident. In case of ambiguity, the variogram may be calculated in a number of directions, but there are rarely enough data to do so.

The *variogram map* takes the idea of calculating the variogram in a number of directions to its logical extreme. The variogram is calculated for a large number of directions and distances; then, the variogram values are posted on a map where the center of the map is the lag distance of zero, see Figure 4.3. Selection of the lag spacing or the size of the pixels on a variogram map is dictated by the same considerations as discussed below. The variogram map

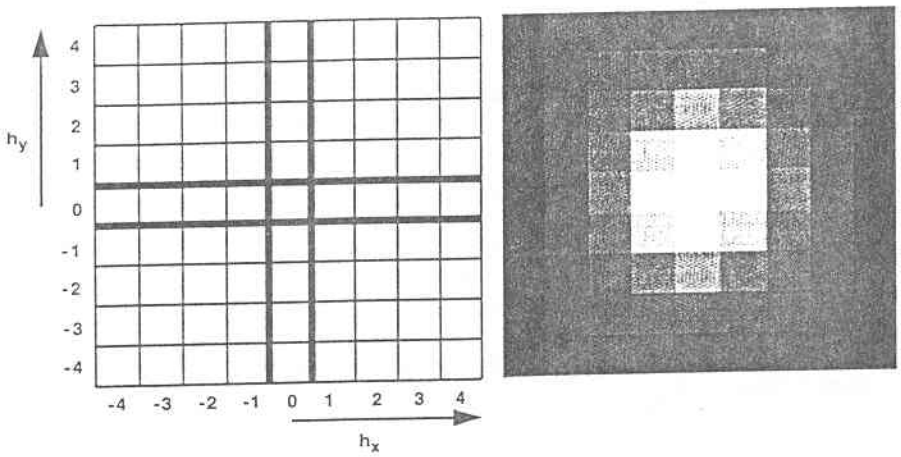


Figure 4.3: Illustration of a variogram map: a square lag template is shown to the left. The variogram is calculated for all $nx \cdot ny$ pairs and then plotted as a gray-scale map on the right.

is discussed extensively by Isaaks and Srivastava [138]. The varmap program in GSLIB [62] will calculate a variogram map with either gridded or scattered data. The main use of the variogram map is to detect the major and minor directions of continuity. Note that the variogram map will be very noisy and of little use in presence of sparse data, which is precisely the case when the directions of anisotropy are poorly understood.

Once the "vertical" and two "horizontal" directions are chosen, the next decision is the variogram lag distances to consider.

Lag Distance

Once the direction for variogram calculation has been chosen it is necessary to choose the distance lags. This is not a problem with regularly gridded data; the spacing between the grid nodes in the direction of interest is the distance lag. In presence of irregularly located data the specification becomes more complex.

For data not on a regular grid, distance and direction tolerances must be permitted to have enough data, that is, $N(\mathbf{h})$ in expression (4.11) should be large enough to permit reliable variogram calculation. The vertical direction is considered separately since its distance scale is significantly different from the horizontal. At the end, the vertical and horizontal directions will be modeled together; however, calculation proceeds separately.

There is an important trade-off in the selection of tolerance parameters for both the vertical and horizontal directions. We want to make the number of data in each lag, $N(\mathbf{h})$, as large as possible so that the variogram is as reliable as possible. At the same time, we would like to restrict the tolerance

parameters as much as possible so that the distance resolution and directional anisotropy are resolved in as much detail as possible. Inevitably, achieving the correct trade-off in any particular situation requires iterative refinement of the chosen tolerance parameters.

Vertical Lag Distance

Figure 4.4 shows the tolerance parameters for a vertical lag: a distance, h, a distance tolerance, h_{tol}, an angle tolerance, a_{tol}, and a bandwidth, b. Some guidance on the selection of these parameters:

- The lag separation distance h is usually chosen to coincide with the data spacing. Well log porosity values separated by 0.5 ft would naturally lead to a unit lag distance of $\mathbf{h} = 0.5$ and multiples of 0.5.
- The distance tolerance, h_{tol}, is often chosen at one half of the unit lag distance h. This could be reduced to, say, one quarter of the unit lag distance when there are many data on a nearly regular grid. The tolerance , h_{tol}, could also be increased to, say, three quarters the unit lag distance in presence of few data. Increasing the tolerance beyond one half the unit lag distance causes some data pairs to contribute to multiple lags; this is recommended when there are less than, say, 50 data pairs reporting to any given lag.
- The angle tolerance, a_{tol}, is needed when the wells are not truly vertical. A tolerance of 10 to 20 degrees is often used. Highly deviated and horizontal wells should not contribute to the calculation of the vertical variogram.
- A bandwidth parameter, b, is sometimes used to limit the maximum deviation from the vertical direction. The bandwidth should be considered with deviated wells.

The number of distance lags must also be chosen. The number of lags, $n_{\mathbf{h}}$, is chosen so that the total distance, $n_{\mathbf{h}} \cdot h$ is about one half of the reservoir size in the direction being considered. For example, 70 lags would be chosen with a lag separation distance of 0.5 ft in a reservoir 70 ft thick with data every 0.5 ft. Distance lags greater than one half of the reservoir size do not permit data in the center of the reservoir to be used. The variogram becomes erratic and not representative of the entire reservoir layer. Moreover, the variogram for such long distances is typically not needed for later geostatistical modeling.

Horizontal Lag Distance

Figure 4.5 shows the tolerance parameters for a horizontal lag: a distance, h. a distance tolerance, h_{tol}, a horizontal angle tolerance, a^h_{tol}, a horizontal bandwidth, b_{hor}, a vertical angle tolerance, a^v_{tol}, and a vertical bandwidth,

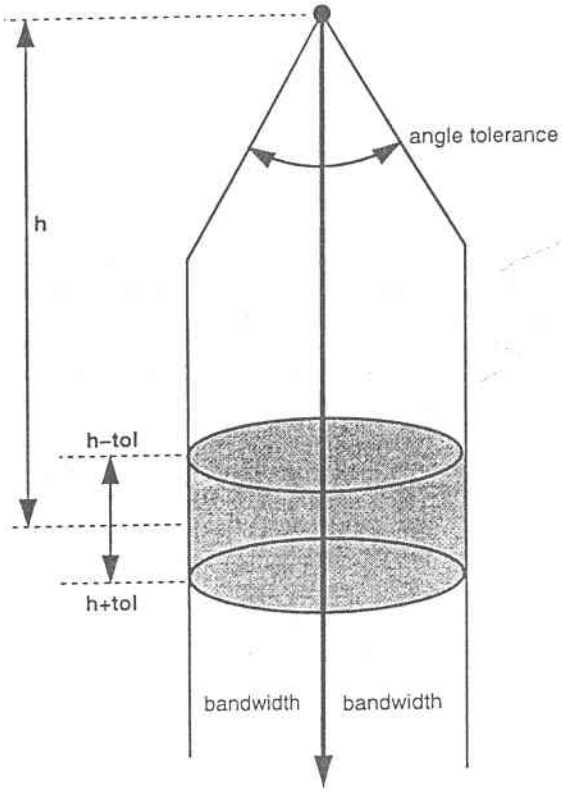


Figure 4.4: Illustration of a vertical lag distance specified by (1) a distance, h, (2) a distance tolerance, (3) an angle tolerance, and (4) a bandwidth.

b_{ver}. The guidance given above for the distance, distance tolerance, and number of lags applies here. Additional considerations:

- In presence of no horizontal anisotropy the parameter of the horizontal angle tolerance, a^h_{tol}, may be set to 90°or greater, which amounts to pool all horizontal directions. This yields an *omnidirectional* horizontal variogram.
- In presence of horizontal anisotropy, the tolerance a^h_{tol} must be restricted as much as possible. Too small and there are too few pairs for reliable variogram calculation; too large and the result is a *blurred* picture of the anisotropy. Starting at 22.5°and performing a sensitivity study is recommended.
- The bandwidth parameter, b_{hor}, is sometimes used to limit the maximum deviation from the horizontal direction. It is set to a large value for omnidirectional variogram calculation or when there are few data.

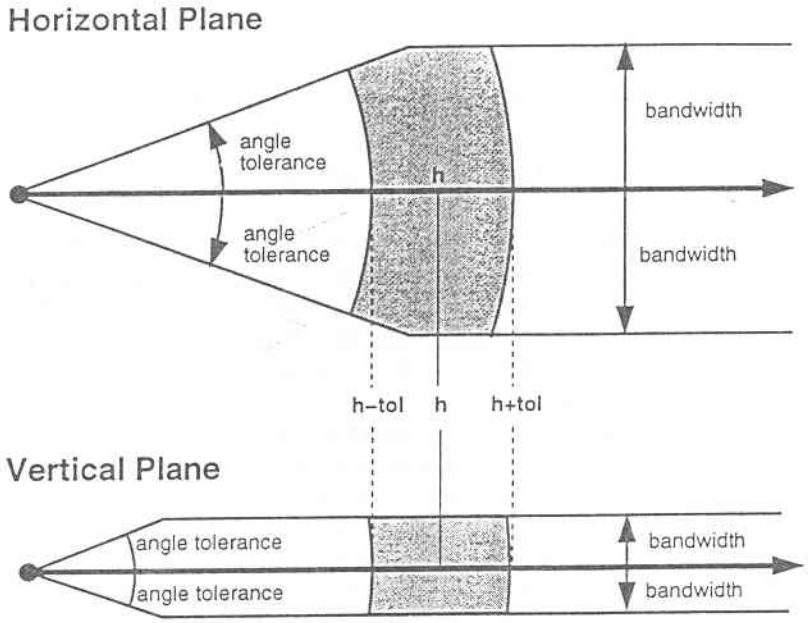


Figure 4.5: Illustration of a horizontal lag distance specified by (1) a distance, h, (2) a horizontal distance tolerance, (3) a horizontal angle tolerance, (4) a horizontal bandwidth, (5) a vertical angle tolerance, and (6) a vertical bandwidth.

For directional variograms in presence of sufficient data, it can be set small, say, to 1 to 3 times the unit lag distance.

- The vertical angle tolerance, a^v_{tol}, should be set quite small due to the much larger variability in the vertical direction. Normally a combination of a small angle tolerance, say $a^v_{tol} = 5^\circ$and a vertical bandwidth, b_{ver}, set to a small value, say $b_{ver} = 2$ ft is effective at limiting the calculation to data at approximately the same stratigraphic position.

As with the vertical variogram, the number of distance lags is chosen so that the total distance, $n_h \cdot h$ is about one half of the dimension of the area represented by the variogram. For example, 5 lags would be chosen with a lag separation distance of 1000 ft in a reservoir 10000 ft in size with wells about every 1000 ft.

The data spacing in the horizontal direction is often more irregular than the vertical direction, making it more difficult to establish the unit lag separation distance. A "lag histogram" may be a useful tool to help choose the lag distances in the horizontal direction.

A lag histogram shows the number of pairs falling into different distance classes. Figure 4.6 shows example lag histograms for two horizontal directions. In this case, natural lag spacing for variogram calculation are indicated

by the period of peaks in these histograms, for example, in the Y direction at 2000, 3000, 3800, 4500, 5600, and 7500 ft. Choosing a lag spacing of 1000 ft with a tolerance of 500 ft would result in a maximum of pairs in each lag. An arbitrary choice, usually too small, could lead to a noisy variogram.

4.3 Interpreting Experimental Variograms

Variogram interpretation is important. The calculated variogram points are not directly usable since (1) noisy results should be discounted, (2) geological interpretation should be used in the construction of the final variogram model, and (3) we need a licit variogram measure for *all* distances and directions. For these reasons, the experimental directional variogram must be understood and then modeled appropriately.

An apparently simple and widely misunderstood concept in variogram interpretation is the *sill*. In an ideal theoretical world, the sill is either the stationary infinite variance of the random function or the dispersion variance of data support volumes within the volume of the study area. In a naive practical world, the sill is the where the variogram appears to flatten off. Reservoirs are not infinite and variograms do not flatten off where they are supposed to. In practical geostatistical reservoir modeling, the sill is the equal-weighted variance of the data entering variogram calculation (1.0 if the data are normal scores with *no* declustering weights).

There is no agreement among geostatisticians regarding this definition of the sill. Some would maintain that the sill is where the variogram plateaus; however, large-scale "non-stationarities" such as trends and zonal anisotropy cause the variogram points at large distances to be very erratic. Although there are more calculated variogram values at large distance than the equal-weighted variance; nevertheless, the equal-weighted variance is always more reliable than trying to fit a plateau from experimental points. There is more discussion on this in [111].

The *range* is the distance at which the variogram reaches the sill. The range typically depends on direction, with horizontal directions showing greater continuity (larger ranges). At times, the continuity is so great that there is no apparent range. The range is merely one parameter of the variogram; the entire shape of the variogram is important.

Geological Variability

Variogram interpretation consists of explaining variability over different distance scales. The variogram is a chart of *variance versus distance* or *geological variability versus direction and Euclidean distance*. Different features are observed at different distance scales. The "nugget effect" is the behavior at distances less than the smallest experimental lag.

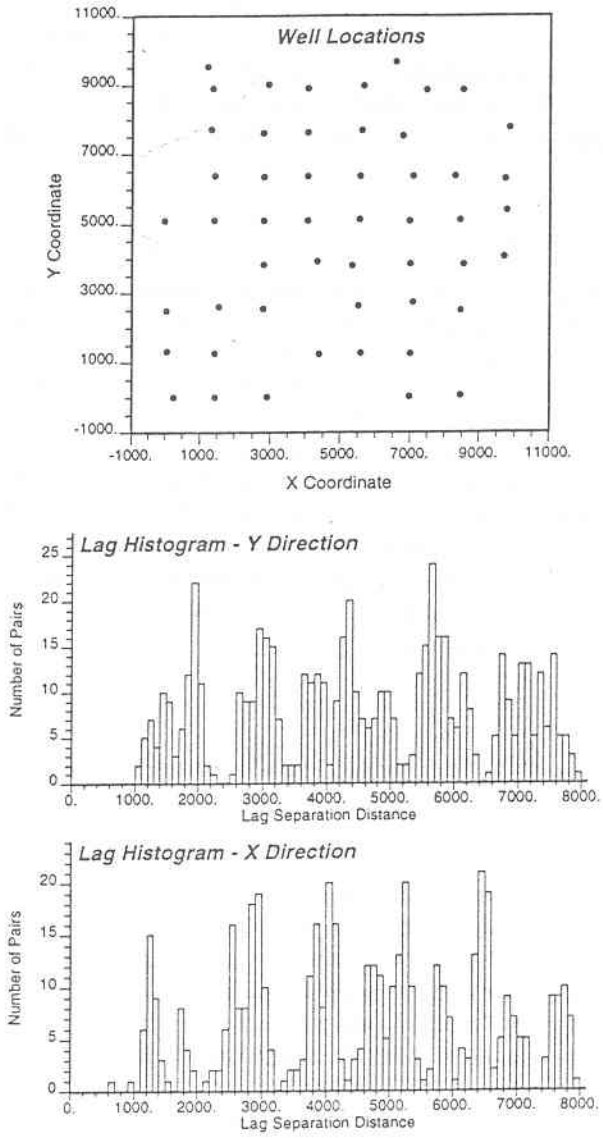


Figure 4.6: Location map of 51 wells and a histogram of the distance between well pairs in the X and Y directions. A 45°angle tolerance was considered in both cases. These histograms can be used to select the lag distances and tolerances for variogram calculation. This case is quite straightforward.

Nugget Effect

The nugget effect is the apparent discontinuity at the origin of the variogram. This discontinuity is the sum of measurement error and geological variability at scales smaller than the smallest experimental lag. This short scale variogram structure is important for scale-up and flow studies.

Any error in measurement values or the location assigned to the data translates to a higher nugget effect. Sparse data may also lead to an apparently high nugget effect. A "real" geological nugget effect of more than 30% is unusual.

Geometric Anisotropy

Most depositional processes impart spatial correlation to facies, porosity, and other petrophysical properties. The magnitude of spatial correlation decreases with separation distance until a distance at which no spatial correlation exists, the range of correlation. In all real depositional cases the length-scale or range of correlation depends on direction; typically, the vertical range of correlation is often much less than the horizontal range due to the larger lateral distance of deposition. Although the correlation range varies with direction, the nature of the decrease in correlation is often the same in different directions. The reasons for this similarity are the same reasons that underlie Walther's Law, that is, geological variations in the vertical direction are similar to those in the horizontal direction. This type of variogram behavior is called geometric anisotropy.

Cyclicity

Geological phenomena often occur repetitively over geologic time, leading to repetitive or cyclic variations in the facies and petrophysical properties. This imparts a cyclic behavior to the variogram, that is, the variogram will show positive correlation then going to negative correlation at the length-scale of the geologic cycles. These cyclic variations often dampen out over large distances as the size or length-scale of the geologic cycles is not perfectly constant. Such cyclicity is sometimes referred to as a hole effect, from the field of spectral analysis of the shape of the covariance.

Figure 4.7 shows a vertical semivariogram of permeability through a succession of fluvial channels. The high semivariogram value at a distance of 0.8m indicates an average channel thickness of about 0.8m. At a distance of 1.2 to 1.6m, the semivariogram, on average, is comparing values at similar stratigraphic positions in different channels.

Figure 4.8 shows a gray-scale image of an aeolian sandstone and the corresponding vertical and horizontal semivariograms. The semivariogram was calculated on the normal score transform of the gray-scale level (finer grained low-permeability sandstone appears darker). Note the cyclic behavior in the vertical direction.

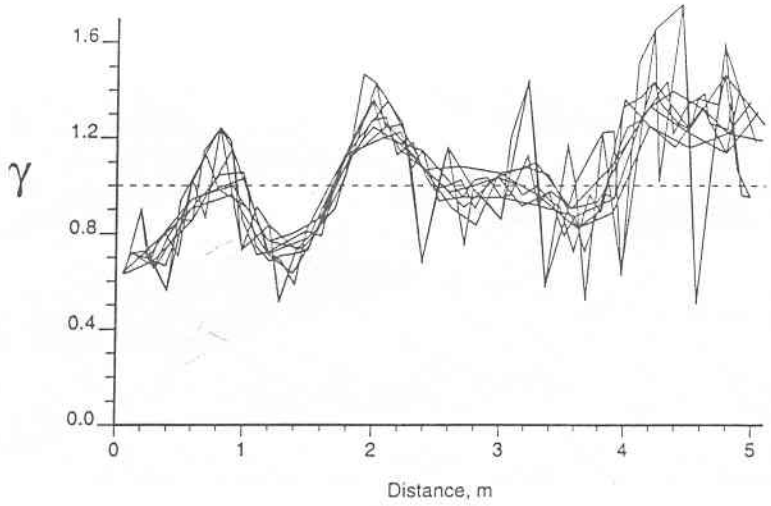


Figure 4.7: Semivariogram of permeability through an alternating succession of fluvial channels. The high semivariogram value at a distance of 0.8m indicates a channel thickness of about 0.8m. At a distance of 1.2 to 1.6 m the semivariogram, on average, is comparing values at similar stratigraphic positions in different channels.

Large-Scale Trends

Virtually all geological processes impart a trend in the petrophysical property distribution, for example, fining or coarsening upward vertical trends or the systematic decrease in reservoir quality from proximal to distal portions of the depositional system. Such trends cause the variogram to show a negative correlation at large distances. In a fining upward sedimentary package, the high porosity at the base of the unit is negatively correlated with low porosity at the top. The large-scale negative correlation indicative of a geologic trend shows up as a variogram that increases beyond the sill variance $C(0) = \sigma^2$. As we will see later, it may be appropriate to remove systematic trends prior to geostatistical modeling.[1] Then, we must calculate the residual variogram.

Figure 4.9 shows an example. A porosity well log (note scale) from a deltaic succession is shown on the left and the corresponding normal scores variogram is shown on the right. The fining upward trend in porosity is revealed by the variogram points climbing above the expected sill of 1.0. Semivariogram points above the sill imply a negative correlation at that distance, which is correct; a high porosity value at the bottom entails a low porosity at the top.

[1]The trend is of no consequence if at scales smaller than the modeling cells or if there are many data. The trend would be reproduced in the final models by the conditioning data. Most practical reservoir modeling is done in presence of sparse well data; therefore, trend modeling is important

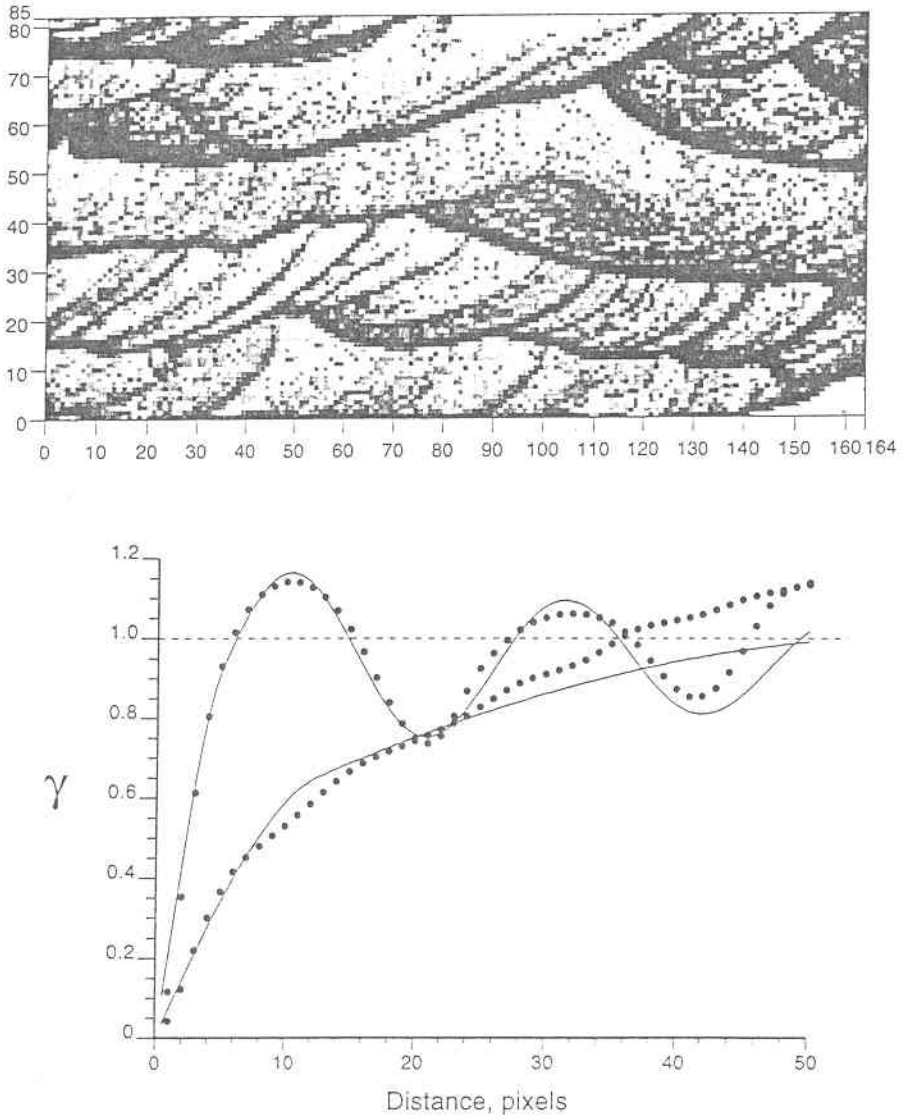


Figure 4.8: Gray-scale image of an aeolian sandstone and the corresponding vertical and horizontal semivariograms. The semivariogram was calculated on the normal score transform of the gray-scale level (finer grained low-permeability sandstone appears darker). Note the cyclic behavior in the vertical direction.

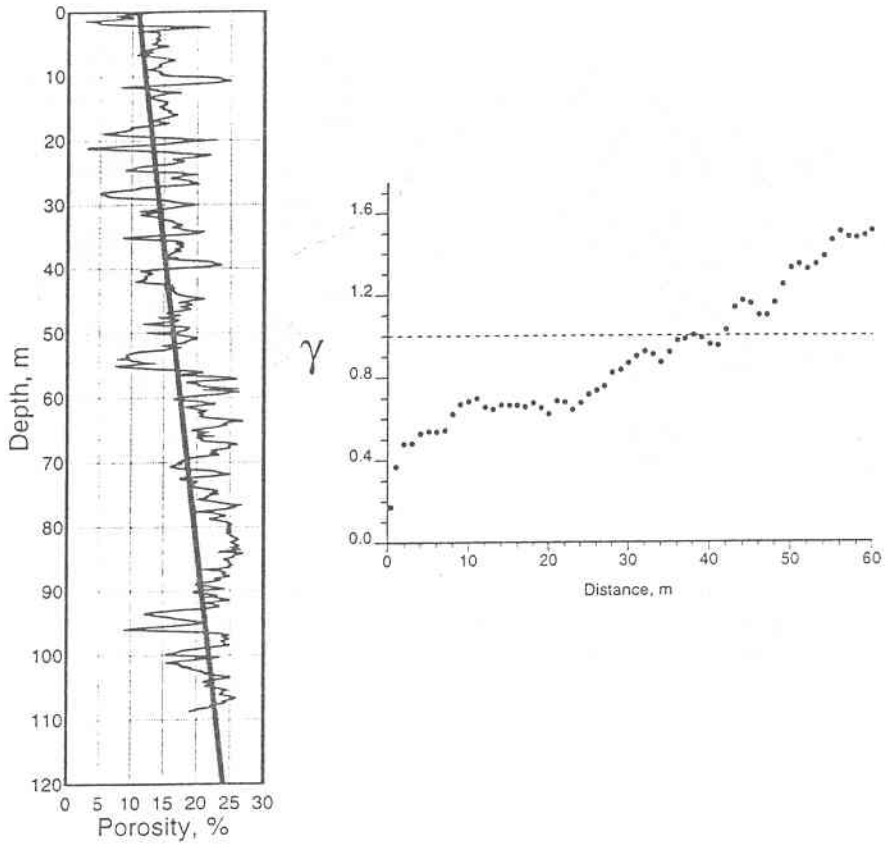


Figure 4.9: A porosity log (note scale) from a deltaic succession is shown on the left and the corresponding normal scores variogram is shown on the right. The fining upward trend in porosity is revealed by the variogram points climbing above the expected sill of 1.0.

Zonal Anisotropy

Zonal anisotropy is a limit case of geometric anisotropy where the range of correlation in one direction exceeds the field size, which leads to a directional variogram that appears not to reach the sill or variance. In practice, a zonal anisotropy is modeled as an extreme case of geometric anisotropy; however, there are two important special cases, which follow.

Areal Trends

Areal (horizontal) trends have an influence on the vertical semivariogram, in that the vertical semivariogram will not encounter the full variability of the petrophysical property. There will be positive correlation (semivariogram

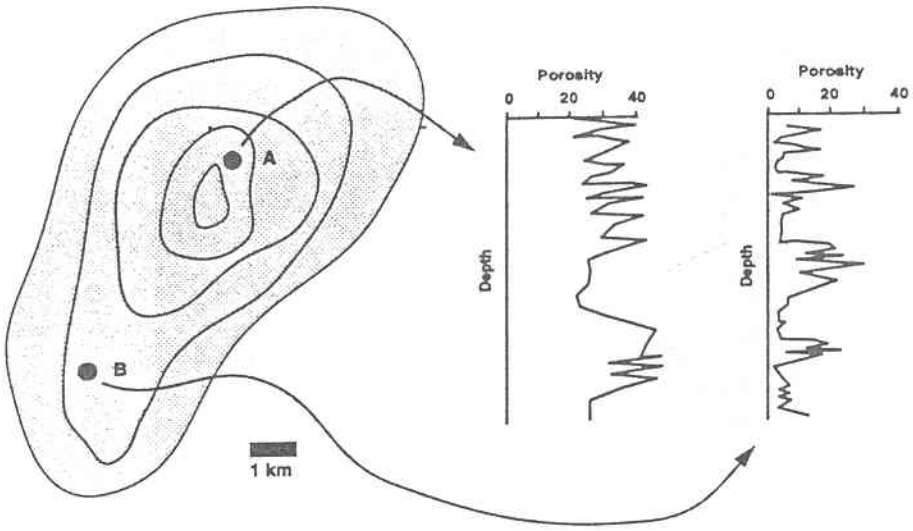


Figure 4.10: In presence of areal trends (illustrated at the left) each well will not "see" the full range of variability, that is, wells in the higher valued areas (for example, well A) encounter mostly high values whereas wells in the lower valued areas (for example, well B) encounter mostly low values. The vertical variogram in this case does not reach the total variability, that is, it shows a zonal anisotropy.

$\gamma(\mathbf{h})$ below the sill variance $C(0) = \sigma^2$) for large distances in the vertical direction. A schematic illustration of this is given in Figure 4.10. A vertical variogram for such a case is given in Figure 4.11.

Layering

There are often stratigraphic layer-like features or vertical trends that persist over the entire areal extent of the reservoir. This leads to positive correlation (semivariogram $\gamma(\mathbf{h})$ below the sill variance $C(0) = \sigma^2$) for large horizontal distances. Although large-scale geologic layers are handled explicitly in the modeling; there can exist layering and features at a smaller scale than cannot be handled conveniently by deterministic interpretation. This type of variogram behavior is also called zonal anisotropy because it is manifested when some directional variograms do not reach the expected sill variance.

Common Problems

The single biggest problem in variogram interpretation is a lack of data to calculate a reliable variogram; there is too little to interpret. Too few data for reliable variogram interpretation does not mean there is no variogram; it is just hidden or masked by data paucity. The use of analogue data and

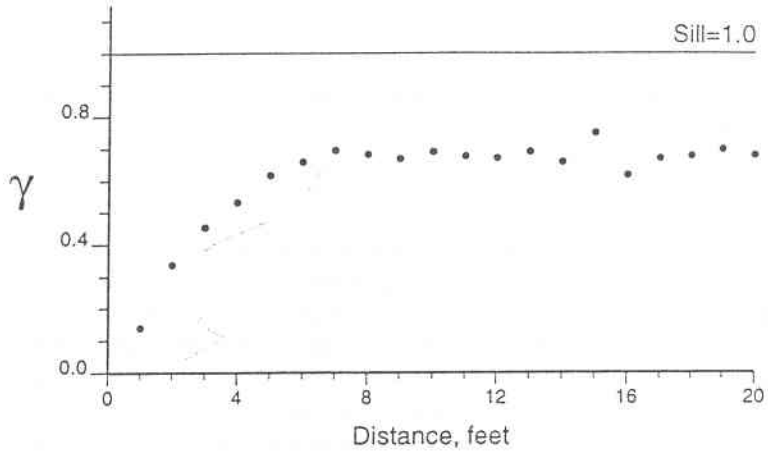


Figure 4.11: The vertical semivariogram of the normal scores of porosity showing a zonal anisotropy, that is, the semivariogram points flatten off at a variance different than the expected variance. This corresponds to the schematic interpretation given on Figure 4.10.

familiarity with other reservoirs of similar depositional setting is critical to make up for too few data.

Clustered data are often present in anomalous areas, for example, regions of high reservoir quality. Such data can cause the first variogram lags to be too high or too low, which could lead to a misinterpretation of the variogram structure. The most common misinterpretations would be (1) a too-high nugget effect, and (2) a short-scale cyclic characteristic. Solutions to this problem include (1) removing the clustered data, (2) discrediting unusually high variogram points for the short distances, or (3) using more robust two-point statistics such as the sample correlogram rather than the variogram. The sample correlogram is defined as:

$$\rho(\mathbf{h}) = C(\mathbf{h})/\sqrt{\sigma \cdot \sigma_{\mathbf{h}}}$$

where

$$C(\mathbf{h}) = \frac{1}{N(\mathbf{h})} \sum_{N(\mathbf{h})} [z(\mathbf{u}) \cdot z(\mathbf{u}+\mathbf{h})] - m \cdot m_{\mathbf{h}}$$

$$m = \frac{1}{N(\mathbf{h})} \sum_{N(\mathbf{h})} z(\mathbf{u}), \quad m_{\mathbf{h}} = \frac{1}{N(\mathbf{h})} \sum_{N(\mathbf{h})} z(\mathbf{u}+\mathbf{h}),$$

$$\sigma = \frac{1}{N(\mathbf{h})} \sum_{N(\mathbf{h})} [z(\mathbf{u}) - m]^2, \text{ and } \sigma_{\mathbf{h}} = \frac{1}{N(\mathbf{h})} \sum_{N(\mathbf{h})} [z(\mathbf{u}+\mathbf{h}) - m_{\mathbf{h}}]^2$$

This measure is robust because of the use of lag-specific mean and variance values.

Outlier values may cause the variogram to be noisy and difficult to interpret. An **h**-scatterplot, that is, a cross plot of the $(Z(\mathbf{u}),Z(\mathbf{u}+\mathbf{h}))$ pairs sometimes reveals problem data that can be removed from the data set for variogram calculation. Figure 4.1 shows three **h**-scatterplots corresponding to three different lag values.

Removing Trends

As mentioned above, the first important step in all geostatistical modeling exercises is to establish the correct variable to model and to make sure (inasmuch as possible) that this property can be modeled as stationary over the domain of the study. Indeed, if the data shows a systematic trend, this trend must be modeled and removed before variogram modeling and geostatistical simulation. Variogram analysis and all subsequent estimations or simulations are then performed on the residuals. The trend is added back to the estimated or simulated values at the end of the study.

There are problems associated with defining a reasonable trend model and removing the deterministic portion of the trend; however, it is essential to consider deterministic features such as large-scale trends deterministically. The presence of a significant trend makes the variable non-stationary, in particular, it is unreasonable to expect the mean value to be independent of location. Residuals even from some simple trend model are easier to consider stationary.

Trends in the data can be identified from the experimental variogram, which keeps increasing above the theoretical sill, see earlier discussion. In simple terms, this means that as distances between data pairs increase the differences between data values also systematically increase.

To illustrate the above, consider the porosity data shown in Figure 4.12, which exhibit a clear trend in the porosity profile along a well. Porosity increases with depth due to a fining-upwards of the sand sequence. The (normal-score) variogram corresponding to this porosity data is shown at the upper right of Figure 4.12. It shows a systematic increase significantly above the theoretical sill of 1.[2] A linear trend was fitted to the porosity profile and then removed from the data. The resulting residuals constitute the new property of interest and their profile is shown at the lower left of Figure 4.12. The (normal score) variogram of the residuals, shown on the right, now exhibits a clearer structure reaching the theoretical sill of 1 at about 7 distance units.

Examples

Real experimental variograms almost always reflect a combination of these different behaviors. Considering the three images and their associated vari-

[2] One could fit a "power" or "fractal" variogram model to the experimental variogram at the upper right of Figure 4.12 however, since these models do not have a sill value, they cannot be used in simulation algorithms such as sequential Gaussian simulation.

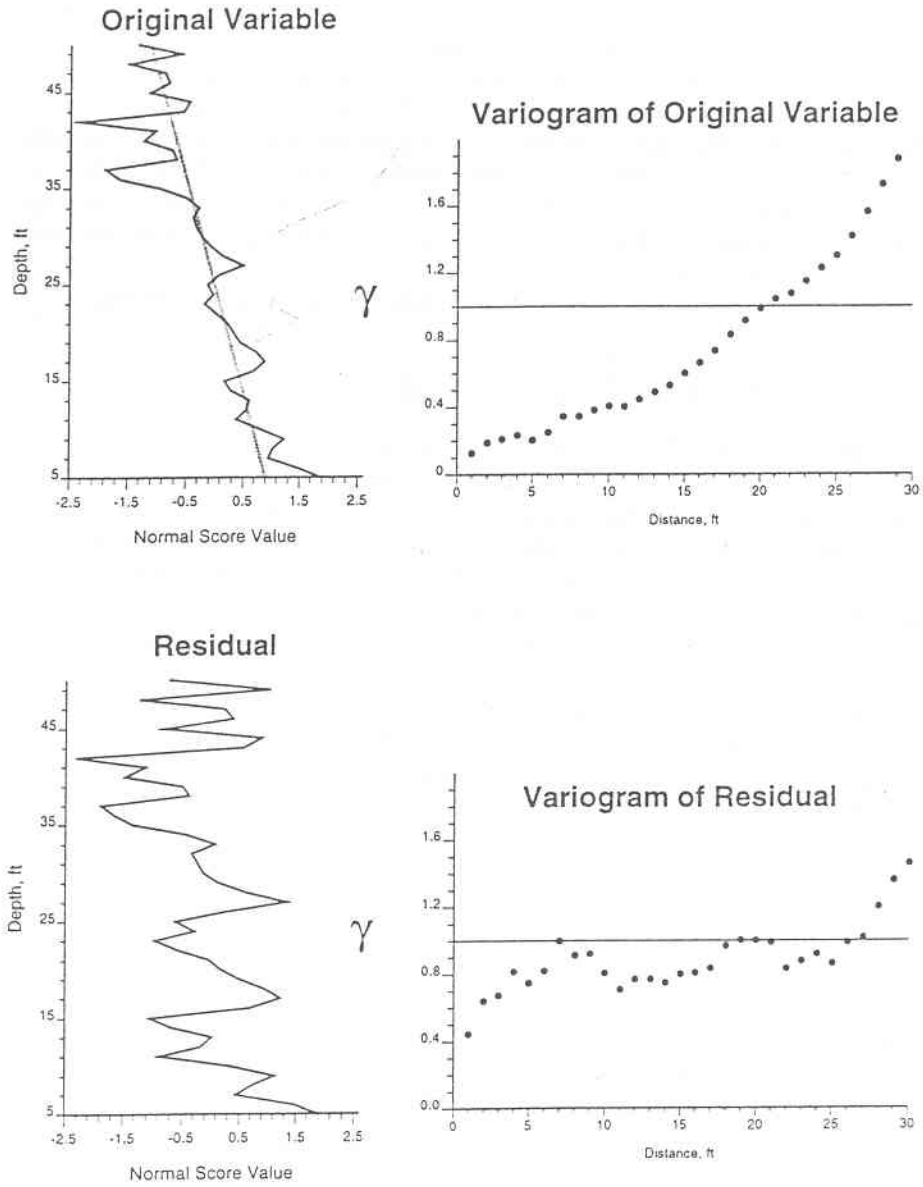


Figure 4.12: An example of removing a trend and the effect on the variogram: the upper left shows the normal score transform of the original porosity values, the upper right figure shows the variogram of these data, which clearly shows the trend, the lower left shows the normal score transform of the residual values (from the linear trend shown on the upper left), and the figure on the lower right shows the variogram of the residual values.

ograms presented in Figure 4.13, we see evidence of all the behaviors: nugget effect, geometric anisotropy and zonal anisotropy, a vertical trend in the middle example, and cyclicity most pronounced on the bottom image.

Figure 4.14 shows eight vertical normal scores variograms of well log porosity from eight different layers in a reservoir. The variogram for layer 4 appears *classical* in the sense that it shows a low nugget effect and rises to the theoretical sill of 1.0. All other cases show combinations of cyclicity (layers 1 and 8), vertical trend (layers 2 and 7), and areal trends (layers 3, 5, and 6). Figure 4.15 shows well log data from 3 layers illustrating cyclicity, vertical trend, and an areal trend.

More examples are shown after the section on variogram modeling.

4.4 Horizontal Variograms

Geostatistical reservoir modeling faces a unique problem. Most wells (particularly exploration wells) are vertical. This makes it straightforward to infer the vertical variogram, but difficult to infer a reliable horizontal variogram. Given the overwhelming noise content in sample horizontal variograms, one evident error is to adopt a pure nugget model, which appears to closely fit the experimental variogram values. This is a convenient, but unrealistic, alternative. Our goal is to infer the best parameters for the underlying phenomenon; it is not to obtain a best fit to unreliable experimental statistics. We must consider secondary information in the form of horizontal wells, seismic data, conceptual geological models, and analogue data. In all cases, however, expert judgment is needed to integrate global information from analogue data with sparse local data.

In presence of sparse horizontal data, there are two critical steps to synthesize a horizontal variogram:

1. Determine what fraction of the variance may be explained by zonal anisotropy, that is, stratification that leads to persistent positive correlation in the horizontal direction.

2. Establish the horizontal-to-vertical anisotropy ratio based on secondary data; the synthetic horizontal variogram consists of the zonal anisotropy (step one) and the scaled vertical variogram.

Figure 4.16 shows a schematic illustration of these two parameters. Both the zonal anisotropy contribution and the horizontal-to-vertical anisotropy ratio will have considerable uncertainty. A sensitivity study should be performed in subsequent geostatistical simulation.

Zonal Anisotropy

There may be persistent stratification that would make the horizontal variogram to plateau different than the sill variance, that is, show zonal anisotropy.

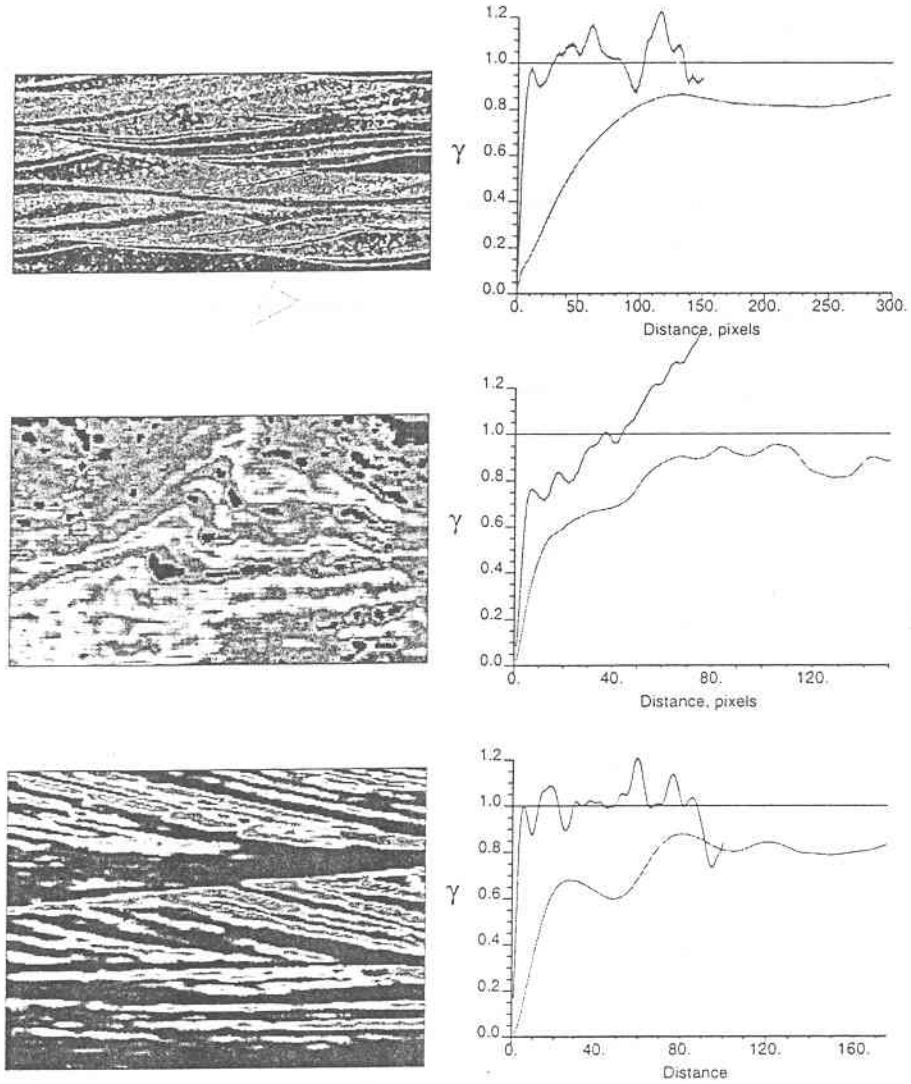


Figure 4.13: Three different geologic images with the corresponding directional variograms. Note the cyclicity, trends, geometric anisotropy, and zonal anisotropy. The top image is an example of migrating ripples in a man-made aeolian sandstone (from the U.S. Wind Tunnel Laboratory), the central image is an example of convoluted and deformed laminations from a fluvial environment. The original core photograph was taken from page 131 of Sandstone Depositional Environments, Scholle and Spearing, 1982. The bottom image is a real example of large-scale cross laminations from a deltaic environment. The original photograph was copied from page 162 of Sandstone Depositional Environments, Scholle and Spearing, 1982.

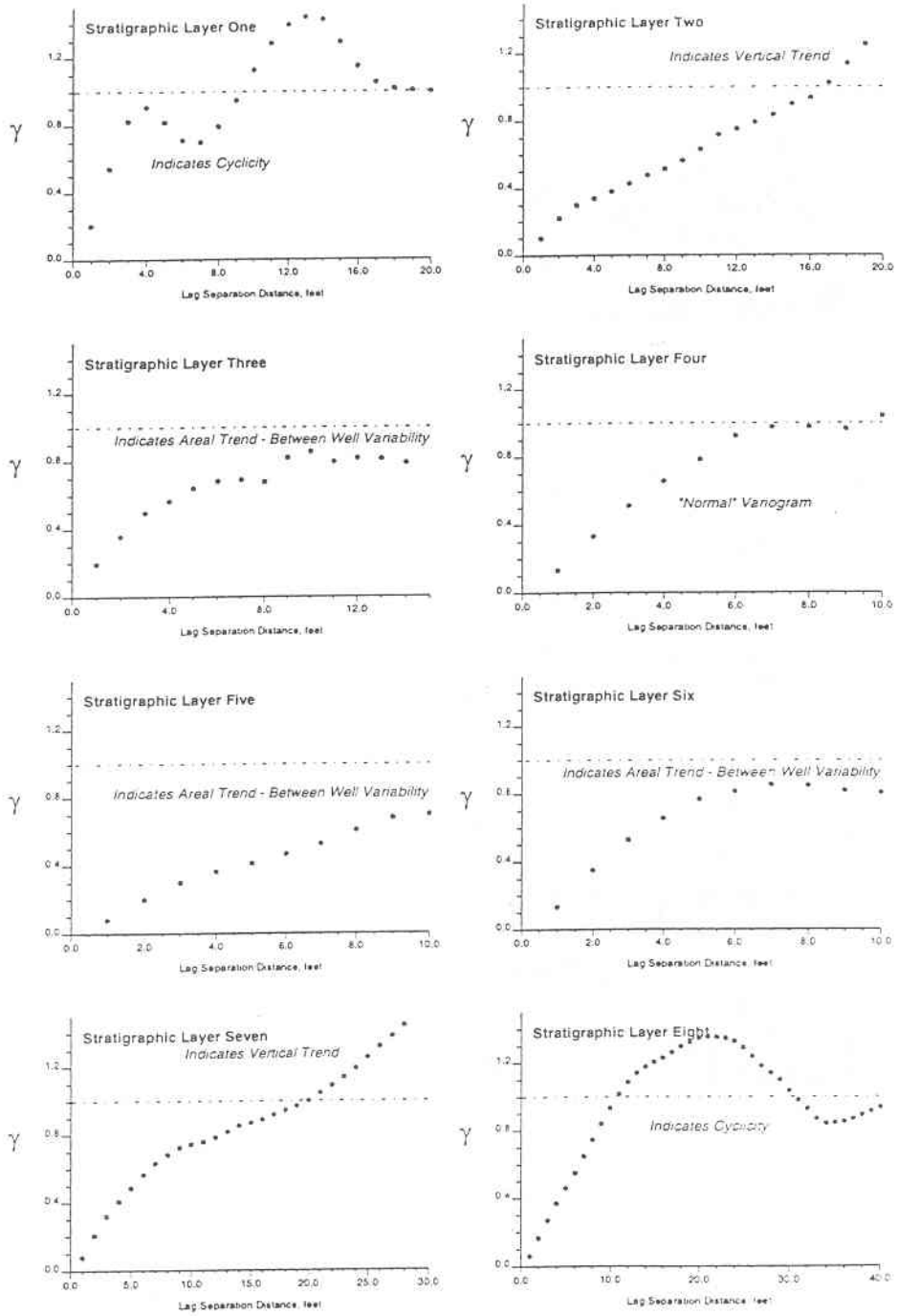


Figure 4.14: Eight variograms, of the normal score transform of porosity, for eight different stratigraphic layers within the same reservoir.

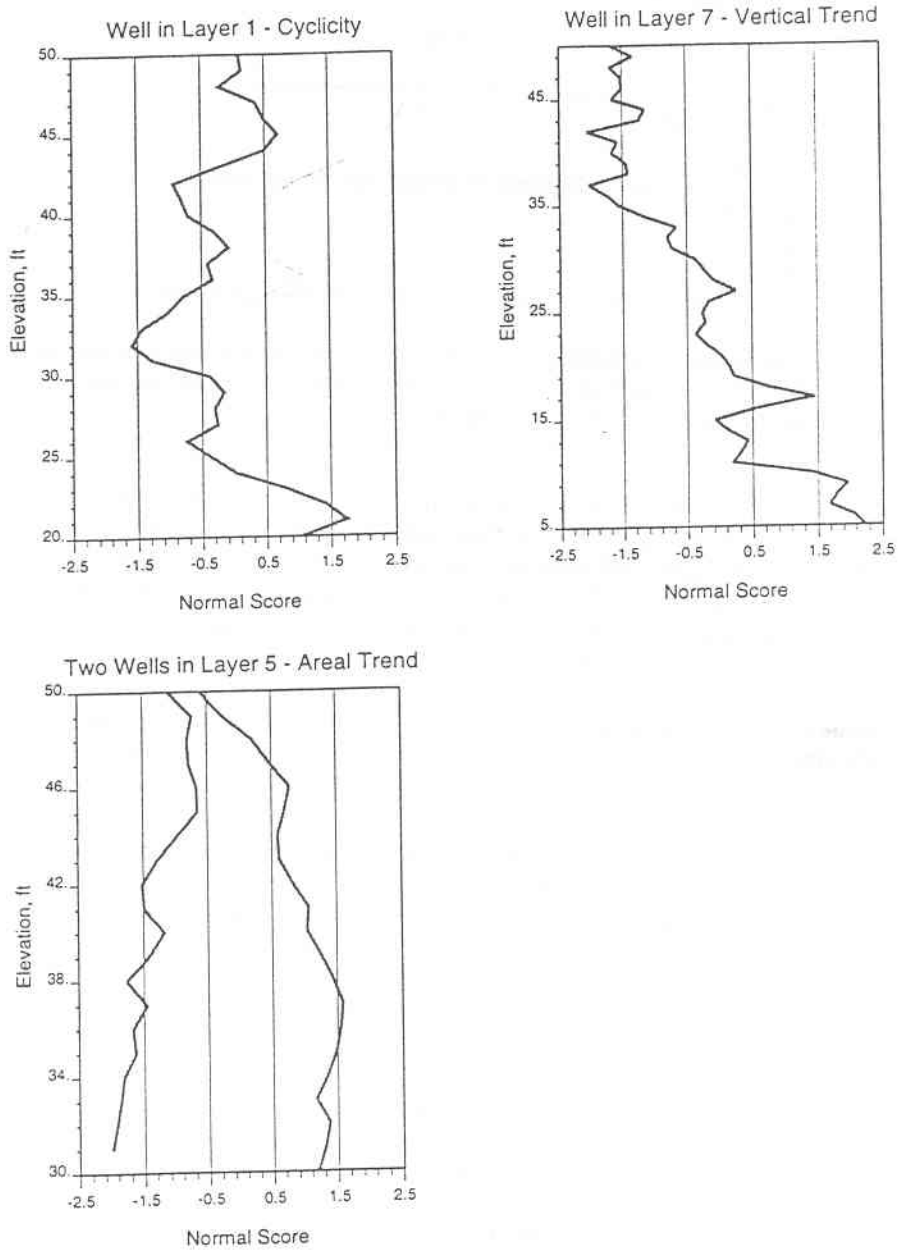


Figure 4.15: Well log data from three layers illustrating cyclicity, vertical trend, and an areal trend.

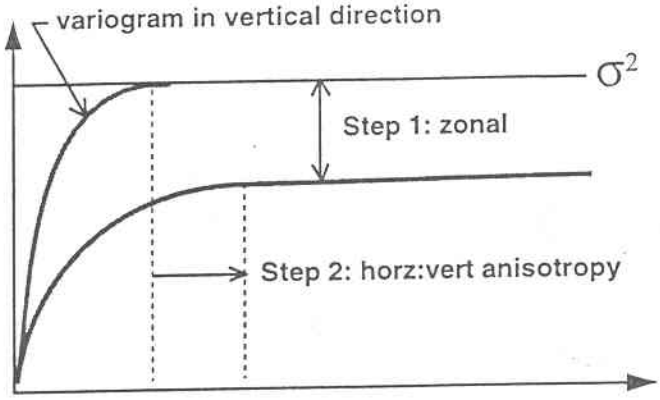


Figure 4.16: Schematic illustration of the two decisions required for horizontal variogram inference in presence of sparse data: (1) the amount of zonal anisotropy, and (2) the horizontal-to-vertical anisotropy ratio.

In presence of sufficient data, this could be observed on directional variograms. In presence of sparse data, that zonal anisotropy must be inferred either from the available well data or from a conceptual geologic model.

A *conceptual geologic model* is essentially an expert opinion or judgment regarding the spatial variability. No reservoir stands in complete isolation. There may be reservoirs in the same sedimentary basin, reservoirs of similar type in different parts of the world, or modern analogues that give some indication of patterns of spatial variation. The "zonal anisotropy" considered here is a specification of the fraction of the variability explained by reservoir-wide stratification. Normally, this is between 0 to 30% of the total variability.

The *available data* may be too few and too widely spaced to calculate a reliable horizontal variogram. Nevertheless, zonal anisotropy consists of reservoir-wide patterns of variation and may be observed with as few as two wells. A horizontal **h**-scatterplot of values paired at the same stratigraphic vertical coordinate can be constructed; the correlation coefficient is an estimate of the relative magnitude of the zonal anisotropy; a large correlation coefficient indicates a large zonal anisotropy. A correlation coefficient of zero would indicate no zonal anisotropy. This value cannot be used blindly. Outliers or insufficient data would be reason to discredit this approach. Figure 4.17 shows an example with two vertical wells.

Horizontal-to-Vertical Anisotropy Ratio

A horizontal variogram model could consist of the vertical variogram points scaled (1) to show the correct plateau, determined previously, and (2) to show the correct distance scale. The distance scale is determined by establishing a horizontal-to-vertical anisotropy ratio. Secondary information such as hor-

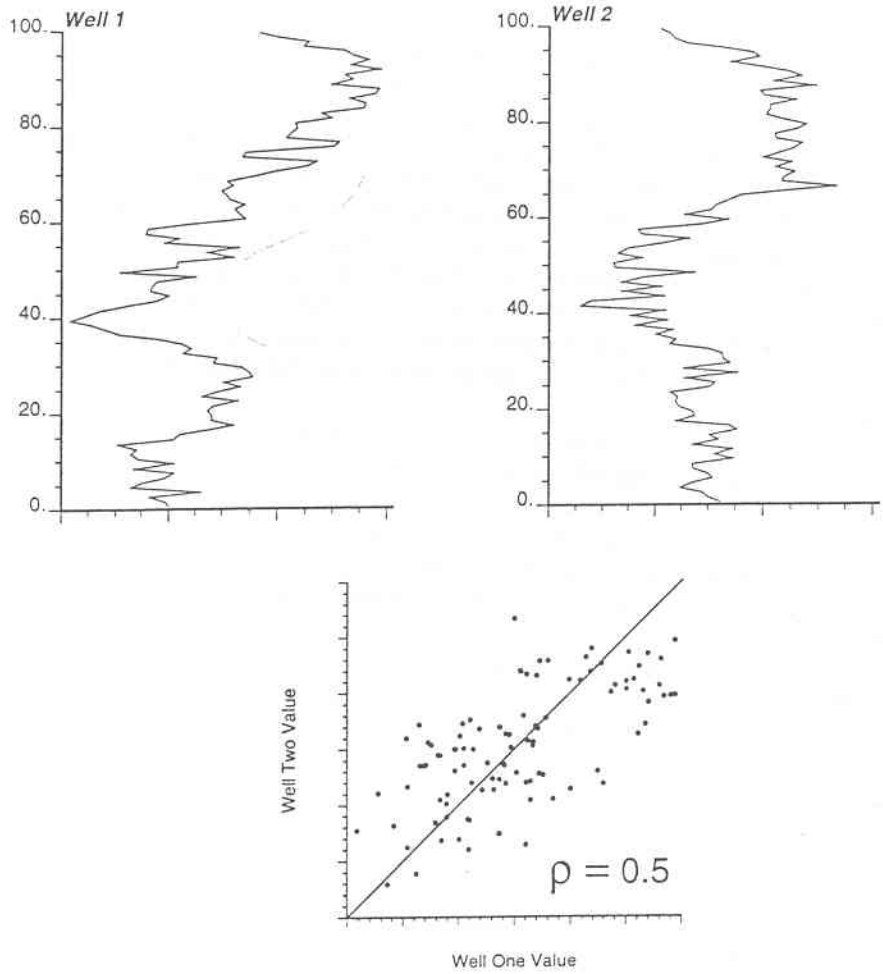


Figure 4.17: Porosity values at two vertical wells and a cross plot of the matching (same stratigraphic position) porosity values from one well to the other.

izontal wells, seismic data, conceptual geological models, and analogue data are considered to establish the required anisotropy ratio.

The increasing popularity of horizontal wells has not significantly helped with horizontal variogram inference. Horizontal wells rarely track the stratigraphic "time lines" or stratigraphic coordinates described in Chapter 3. When wells are close to "stratigraphic" horizontal, experimental horizontal variograms should be calculated. Note that even small departures from horizontal can lead to drastically reduced correlation, that is, an underestimation of the horizontal continuity.

Seismic does not directly measure the facies, porosity, or permeability variables being considered in variogram analysis. Nevertheless, the acoustic responses of seismic are related to these petrophysical properties. One could consider the seismic data as a non-linear average of the underlying petrophysical variables with additional error. The volume-scale difference leads to predictable changes to the variogram behavior if the averaging is linear(see Chapter 11); the range of correlation is increased by the scale of averaging. This is very important for the vertical variogram since the vertical scale of averaging is often many times larger than the vertical range. The large-scale nature of seismic is not as critical in the horizontal direction since the horizontal range of correlation is many times larger than the horizontal scale of averaging. The idea, then, is to use the horizontal range of correlation from seismic.

The extensive lateral coverage of seismic permits calculation of a clear and easy-to-interpret horizontal variogram and horizontal range of correlation. The horizontal-to-vertical anisotropy ratio is calculated from the horizontal seismic variogram and the vertical variogram for the variable being studied.

Conceptual geological models consist of an expert opinion regarding the depositional system and the consequent spatial correlation. Such conceptual models are usually based on analogue data, which may come from (1) other, more extensively sampled, reservoirs, (2) geological process simulation, or (3) outcrop measurements. In all cases, expert judgment is needed to integrate global information from analogues with sparse local data. There are published sources of such data, for example, the compilation of Kupfersberger and Deutsch [167].

Geological-based process models simulate physical processes such as sediment transport, deposition, and erosion to characterize property variations over the time scale in which a reservoir evolved. Such process models are not yet viable reservoir modeling tools; however, we expect them to show the general character of reservoir heterogeneity [50, 262]. Essential features such as the existence of preferential flowpaths or sedimentological organization may be revealed [11]. Geological process models are constrained by the geological variables that create the stratigraphic record such as amount and type of sediment supply. Even if the observed location and geometry of facies units cannot be reproduced exactly, the general character may be appropriate for purposes of horizontal variogram inference [213]. They could be used to construct a "training image" from which horizontal variograms are extracted. This would require significant effort.

Experimental variograms are calculated. Analogue data may be considered to supplement inadequate data. These variograms must be consistent with the geological model of the reservoir. Lastly, the variograms must be modeled.

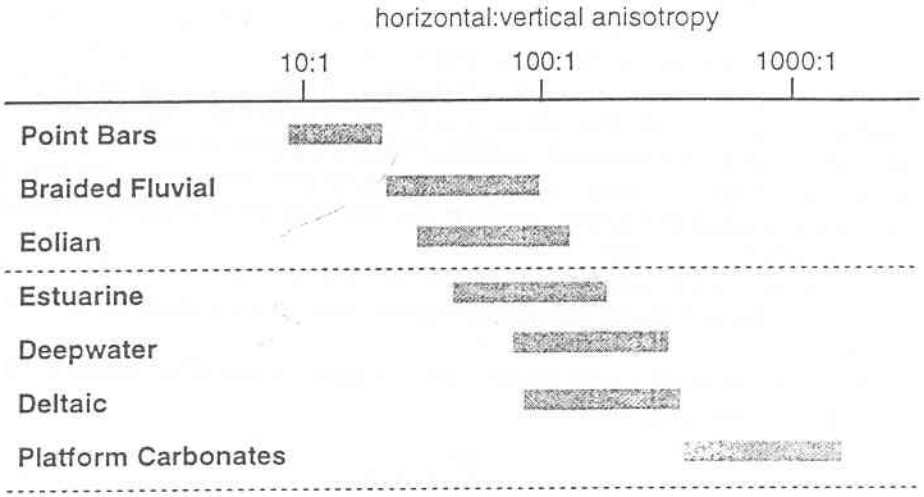


Figure 4.18: Some typical horizontal-to-vertical anisotropy ratio *conceptualized* from available literature and experience. Such generalizations can be used to verify actual calculations and supplement very sparse data.

4.5 Variogram Modeling

The experimental variogram points are not used directly in subsequent geostatistical steps such as kriging and simulation; a parametric variogram model is fitted to the experimental points. There are two reasons why experimental variograms must be modeled:

1. The variogram function is required for all distance and direction vectors $\mathbf{h}$ within the search neighborhood of subsequent geostatistical calculations; however, we only calculate the variogram for specific distance lags and directions (often, only along the principal directions of continuity). There is a need to interpolate the variogram function for $\mathbf{h}$ values where too few or no experimental data pairs are available. In particular, the variogram is often calculated in the horizontal and vertical directions, but geostatistical simulation programs require the variogram in off-diagonal directions where the distance vector simultaneously contains contributions from the horizontal and vertical directions.

2. The variogram measure $-\gamma(\mathbf{h})$ must have the mathematical property of "positive definiteness" for the corresponding covariance model, that is, we must be able to use the variogram and its covariance counterpart in kriging and stochastic simulation. A positive definite model ensures that the kriging equations can be solved and that the kriging variance is positive.

For these reasons, geostatisticians have fit sample variograms with specific known positive definite functions like the spherical, exponential, Gaussian, and hole effect variogram models. It should be mentioned that any positive definite variogram function can be used, including tabulated variogram or covariance values, [276]. The use of an arbitrary function or non-parametric table of variogram values would require a preliminary check to ensure positive definiteness [196]. In general, the result will not be positive definite and some iterative procedure is required to adjust the values until the requirement for positive definiteness is met.

Traditional parametric models permit all geological behaviors discussed above to be fit and allows straightforward transfer to existing geostatistical simulation codes.

A variogram model can be constructed as a positive sum of known positive definite licit variogram functions:

$$\gamma(\mathbf{h}) = \sum_{i=1}^{nst} C_i \Gamma_i(\mathbf{h}) \tag{4.12}$$

where $\gamma(\mathbf{h})$ is the variogram model for any distance and direction vector $\mathbf{h}$, nst is the number of variogram functions or "nested structures," $C_i, i = 1, \ldots, nst$ are the variance contributions for each nested structure, and $\Gamma_i, i = 1, \ldots, nst$ are the elementary licit variogram functions. The variance or "sill" of each elementary structure $\Gamma_i, i = 1, \ldots, nst$ is one; the sum of the variance contributions $\sum_{i=1}^{nst} C_i$ is then the variance implicit to the variogram model.

In general, variograms are systematically fit to the theoretical sill. The commonly used licit variogram models for each nested structure will be presented and then some details about how to combine them together into a final variogram model.

Licit Variogram Models

Figures 4.19 shows six variogram models commonly used in practice. The nugget effect, spherical, exponential, and Gaussian models have a distinct plateau or sill. Each of these structures will be multiplied by a variance C_j value as in 4.12.

Each variogram model has the distance h as a parameter. This scalar distance is calculated by decomposing the distance vector $\mathbf{h}$ into its three principal components, h_{vert}, $h_{h-major}$, and $h_{h-minor}$. By convention *vert* is the vertical direction or that direction aligned with the stratigraphic Z_{rel}coordinate, *h-major* is the horizontal direction of greatest continuity aligned perpendicular to the Z_{rel}coordinate, and *h-minor* is the horizontal direction of least continuity aligned perpendicular to the Z_{rel}coordinate. Then, the distance is calculated by as:

$$h = \sqrt{\left(\frac{h_{vert}}{a_{vert}}\right)^2 + \left(\frac{h_{h-major}}{a_{h-major}}\right)^2 + \left(\frac{h_{h-minor}}{a_{h-minor}}\right)^2} \tag{4.13}$$

In all generality, three angles are needed to define the three directions in (4.13); however, we usually only need one azimuth angle to define the *major* and *minor* directions since the vertical direction is defined stratigraphically. The distance *range* parameters, a_{vert}, $a_{h-major}$, and $a_{h-minor}$ may be different for each nested structure. They are calculated or iteratively adjusted so that all directional sample variograms are fit correctly. Considering all range distance parameters in the calculation of h (4.13) amounts to scale h to be dimensionless with a range of 1. The variance contributions multiply the nested structures (4.12); therefore, the following nested structures also have a dimensionless sill value of 1:

Nugget effect: a constant of 1 except at h equal to zero, where the variance is zero:

$$\Gamma(h) = \begin{cases} 0 & \text{if } h = 0 \\ 1 & \text{if } h > 0 \end{cases}$$

The nugget effect is due to measurement error and geological small-scale structures, that is, features occurring at a scale smaller than the smallest data separation distance.

Spherical: a commonly encountered variogram shape, which increases in a linear fashion and then curves to sill of 1 at a distance of 1:

$$\Gamma(h) = Sph(h) = \begin{cases} \left[1.5h - 0.5h^3\right], & \text{if } h \leq 1 \\ 1 & \text{if } h \geq 1 \end{cases}$$

Exponential: a variogram shape similar to the spherical shape, that is, linear near the origin and curving toward a sill of 1. The main differences are that it rises more steeply that the spherical shape and reaches the sill value asymptotically:

$$\Gamma(h) = Exp(h) = 1 - e^{-h}$$

Gaussian model: the squared term in the exponential causes this variogram structure to have a parabolic shape at short distances:

$$\Gamma(h) = Gau(h) = 1 - e^{-h^2}$$

The implicit continuity at short distances is typical of continuous phenomena such as structural surfaces and isochores.

Hole effect model: a periodic shape:

$$H_a(h) = 1.0 - cos\,(h \cdot \pi)$$

A hole effect variogram can act in one direction only, that is, two of the three range parameters are set to (essentially) infinity ∞.

Dampened hole effect model: is the product of the exponential covariance and the hole effect model:

$$DH_{d,a}(h) = 1.0 - exp\left(\frac{-h \cdot a}{d}\right) \cdot cos\,(h \cdot a)$$

Because geological periodicity is rarely at the same distance scale, the dampened hole effect model is more commonly used than the hole effect model. A dampened hole effect variogram can also act in one direction only.

There are many other variogram models. The *power-law* variogram model, $\Gamma(h) = h^{\omega}$ with $0 < \omega \leq 2$, is one example that is characteristic of trends or fractal-type behavior.

Summary of Variogram Modeling

Variogram modeling is facilitated by interactive software. Whether or not such software is available, the principles and procedures for variogram modeling are the same if the results are to be transferred to conventional geostatistical modeling software.

1. Determine the number of nested structures, that is, the nst of equation (4.12). This number should be as small as possible while retaining the flexibility to capture significant behaviors in different variogram directions.

2. For each nested structure, $i = 1, \ldots, nst$, choose (1) the variogram model type, that is, the Γ_i elementary model, (2) the variance contribution, C_i for this nested structure, and (3) the anisotropy parameters to define distance, equation (4.13), that is, horizontal azimuth rotation angle and a_{vert}, $a_{h-major}$, and $a_{h-minor}$.

3. Display the fitted model with the calculated points in all directions. Iteratively refine parameters as necessary to achieve a "good" fit. The short-scale behavior and significant anisotropy must be fit well for a good fit. Slight periodicity, minor zonal anisotropy, and long range trends are often of little consequence.

The presentation of final variogram model fits in the following examples does not reveal the iterative nature of variogram fitting.

Examples

Figure 4.20 shows horizontal and vertical variograms from a Canadian reservoir. These variograms are indicator variograms standardized to a sill of 1.0. The horizontal distance scale is in meters; the vertical distance scale is relative to the thickness. Five variance regions can be defined for the horizontal

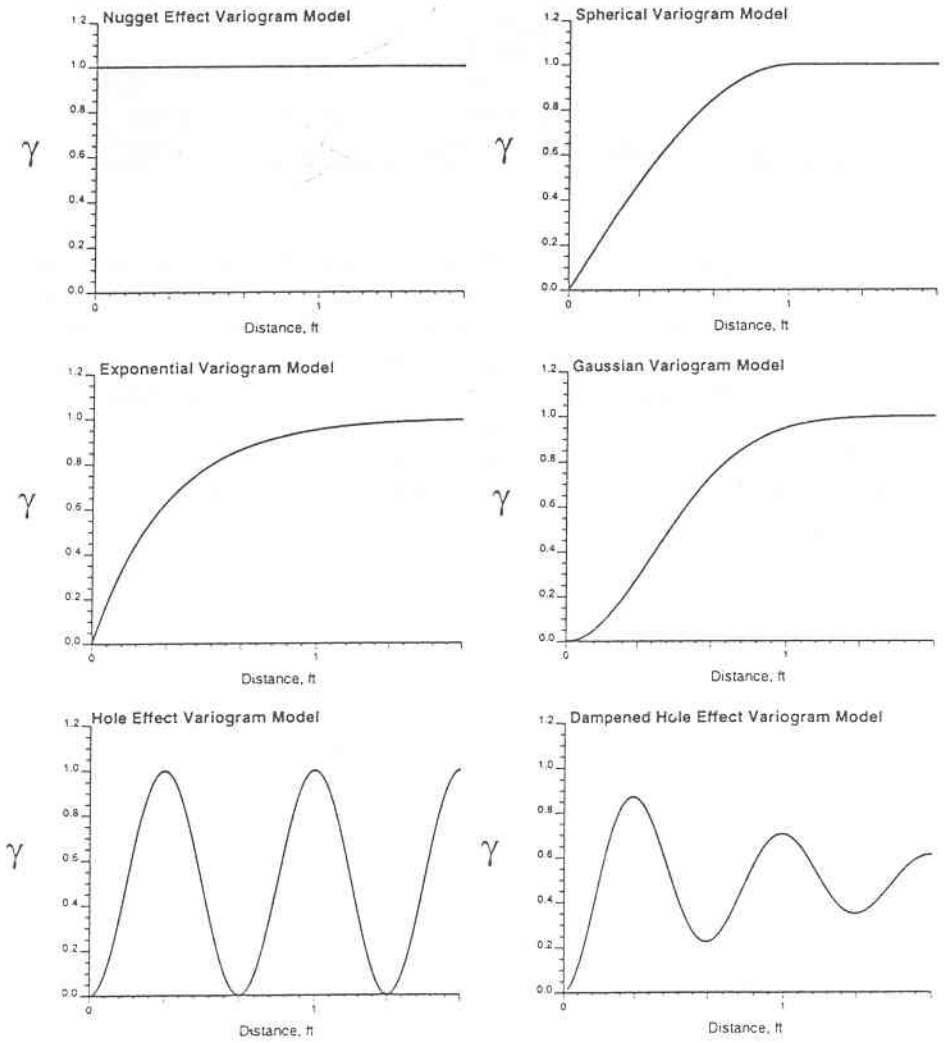


Figure 4.19: Six commonly used variogram models: nugget effect, spherical, exponential, Gaussian model, hole effect model, and dampened hole effect model.

Table 4.1: Variogram parameters corresponding to variogram shown on Figure 4.20.

Variance Contribution	Type of Variogram	Horizontal Range, m	Vertical Range, m
0.05	Nugget		
0.29	Exponential	100.0	0.015
0.46	Exponential	175.0	0.450
0.20	Exponential	100.0	0.500
0.20	Dampened hole effect	∞	0.060

and vertical variograms, see Table 4.1 and Figure 4.20. The first small component (5% of variance) is an isotropic nugget effect. The next three are exponential variogram structures with different range parameters. Three exponential structures are required to capture the inflection point at a variance value of about 0.3 on the vertical variogram, and the long-range structure in the vertical variogram in the variance region 0.8 to 1.0. The fifth dampened hole effect variogram structure only applies in the vertical direction and adds no net contribution to the variance. The dampening component is five times the vertical range of 0.06. The dampened hole effect is largely aesthetic, since it would have little affect on subsequent geostatistical calculations.

Figure 4.21 shows a horizontal and vertical variogram of the normal score transform of porosity for a West Texas reservoir. This data is the "Amoco" data provided to Stanford University for testing and developing geostatistical methods. The horizontal variogram has a range of 5000 ft. The vertical variogram shows a clear additional variance component (zonal anisotropy), which is due to the systematic areal variations in average porosity. Three variance regions could be defined for the horizontal and vertical variogram of Figure 4.21, see Table 4.2. The first two components are spherical variogram structures with geometric anisotropy. The last spherical structure captures the zonal anisotropy in the vertical direction.

Table 4.2: Variogram parameters corresponding to variogram shown on Figure 4.21.

Variance Contribution	Type of Variogram	Horizontal Range, m	Vertical Range, m
0.50	Spherical	750.0	6.0
0.40	Spherical	2000.0	50.0
0.10	Spherical	7000.0	∞

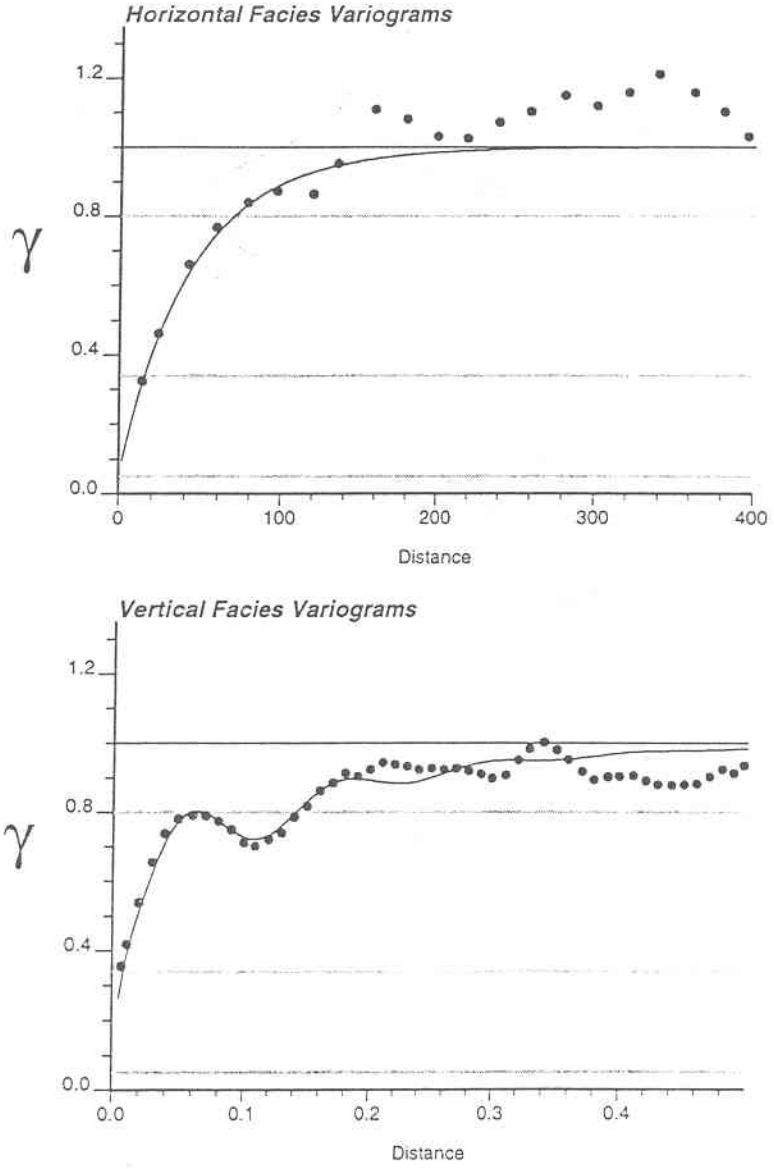


Figure 4.20: Horizontal and vertical indicator variogram for presence/absence of sand in Canadian Reservoir. Note the excellent horizontal variogram that could be calculated with eight horizontal wells.

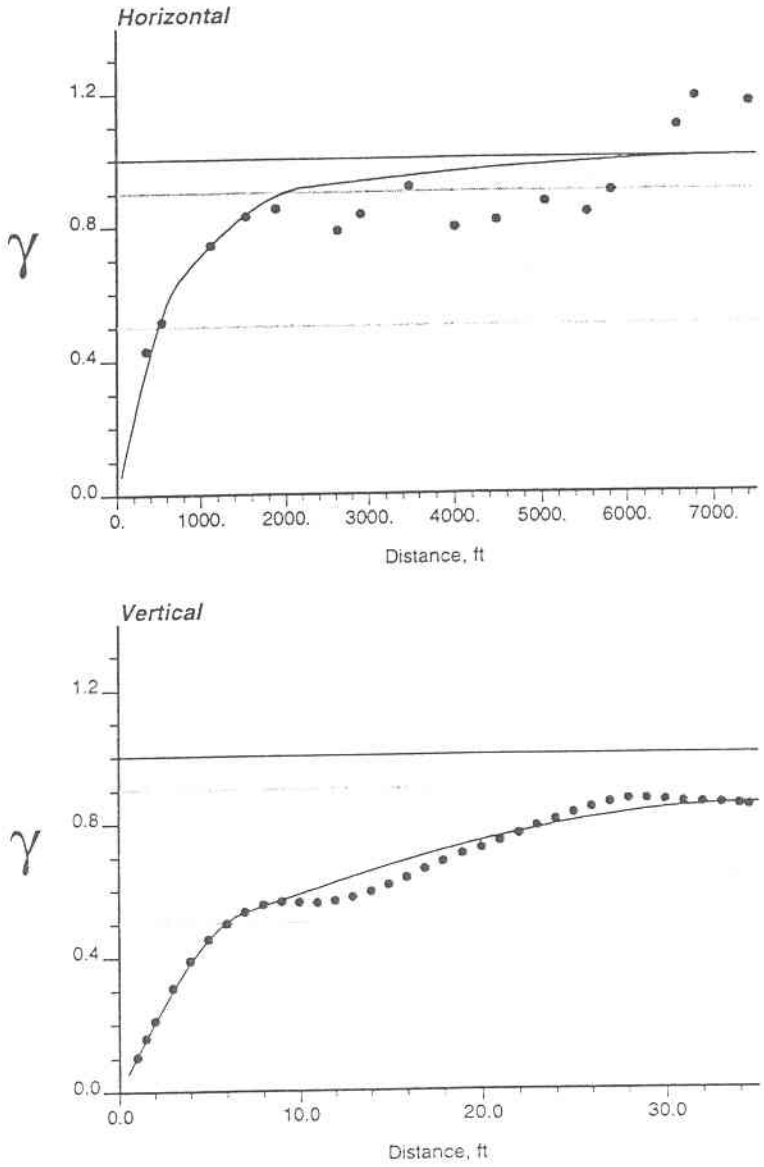


Figure 4.21: Horizontal and vertical variogram of normal score of porosity for West Texas reservoir. This data is the "Amoco" data provided to Stanford University for testing and developing geostatistical methods.

Table 4.3: Variogram parameters corresponding to variogram shown on Figure 4.22.

Variance Contribution	Type of Variogram	Horizontal Range NW-SE, ft	Horizontal Range NE-SW, ft	Vertical Range rel.
0.10	Exponential	150.0	400.0	0.8
0.14	Exponential	2500.0	4000.0	1.2

Figure 4.22 and 4.23 show facies (presence of limestone coded as 1, dolomite and anhydrite coded as zero) and porosity variograms calculated from a major Saudi Arabian reservoir, see [22] for the original variograms. Note the interpretable horizontal variograms and the consistent vertical and horizontal variograms. We also note the presence of a zonal anisotropy in the case of porosity, but not for the facies indicator variogram.

Two variance regions were identified for the facies variogram on Figure 4.22, see Table 4.3. Note that the sill in this case is 0.24 (related to the relative proportion of limestone and dolomite). Both are anisotropic exponential structures. Three variance regions were identified for the porosity variogram on Figure 4.23, see Table 4.4. The third region defines the zonal anisotropy in the vertical direction.

4.6 Cross Variograms

In virtually all reservoir modeling cases there is a need to model the joint distribution of multiple variables, for example, seismic impedance, facies, porosity, permeability, residual water saturation. Conventional practice is to model them in a sequential fashion considering pair-wise correlation:

1. Establish the *best* seismic attribute possibly as a complex combination of the different seismic attributes,

Table 4.4: Variogram parameters corresponding to variogram shown on Figure 4.23.

Variance Contribution	Type of Variogram	Horizontal Range NW-SE, ft	Horizontal Range NE-SW, ft	Vertical Range rel.
0.35	Exponential	400.0	500.0	0.6
0.40	Exponential	2000.0	4000.0	0.6
0.25	Exponential	12000.0	40000.0	∞

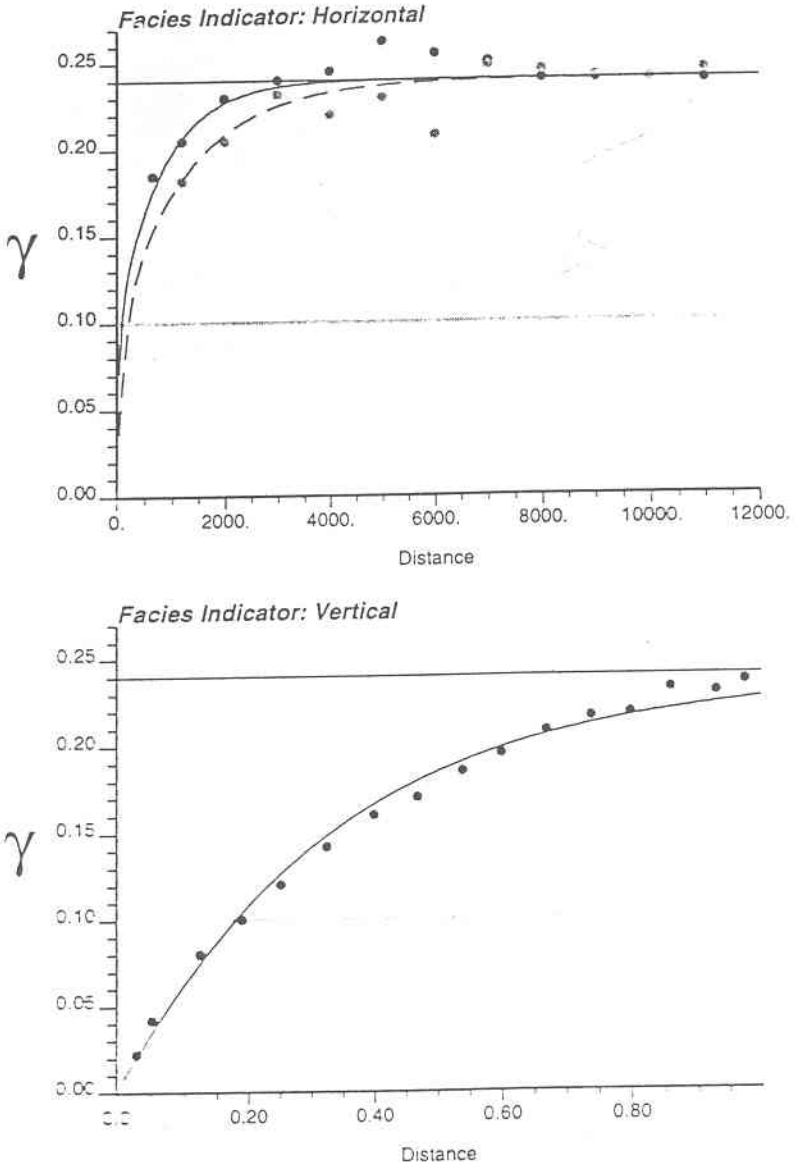


Figure 4.22: Horizontal and vertical facies variograms from a "major Arabian reservoir"; the top figure shows two directional horizontal variograms (see Table 4.3) and the bottom figure is the vertical variogram.

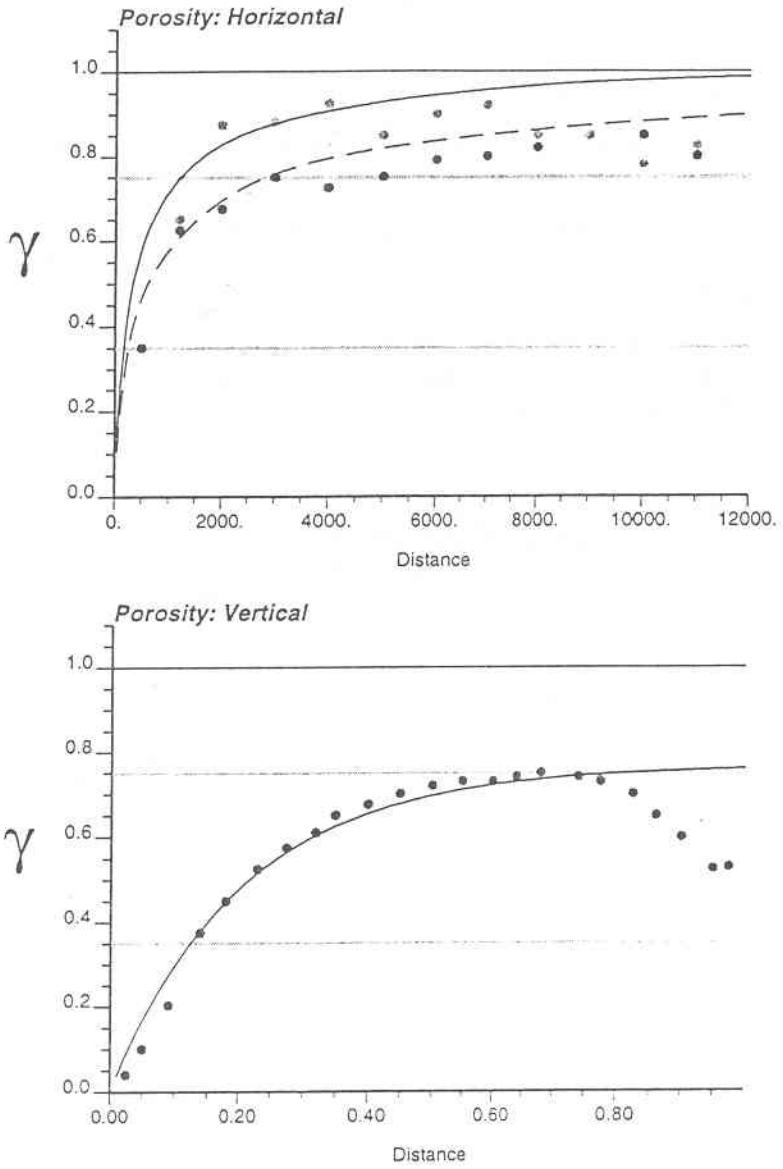


Figure 4.23: Horizontal and vertical porosity (normal score) variograms from the limestone facies in a "major Arabian reservoir"; the top figure shows two directional horizontal variograms (see Table 4.4) and the bottom figure is the vertical variogram. Note the zonal anisotropy in the vertical direction, which indicates the presence of areal trends.

2. Create a geostatistical model of facies considering the seismic data as a secondary variable,
3. Model porosity within each facies considering the seismic data as a secondary variable,
4. Model permeability within each facies considering porosity as a secondary variable, and
5. Continue with other variables such as residual water saturation using the "best" secondary variable, that is, the one most correlated with the variable being modeled.

Geostatistical models have been devised for any number of variables; however, they are difficult to apply in practice. We only consider the spatial correlation between two variables Z and Y.

The cross plot of collocated Z and Y values provides the first assessment of the correlation between the two variables. In general, the relationship between collocated or paired values $z(\mathbf{u}_i), y(\mathbf{u}_i), i = 1, \ldots, n$ does not suffice to capture the full spatial relationship between two variables. The cross spatial relationship between pairs of value separated by some lag distance $z(\mathbf{u}_i), y(\mathbf{u}_i + \mathbf{h}), i = 1, \ldots, n(\mathbf{h})$ must also be considered. The cross variogram, defined below, measures this cross spatial dependence between two variables.

It is possible to avoid cross variograms. A model of cross spatial relationship can be built from the correlation of collocated values. These collocated or Markov models [279] have gained wide popularity because of their simplicity. The implicit assumptions underlying such models are reasonable in presence of an exhaustively sampled secondary variable from either measurement (seismic) or prior geostatistical simulation. These models will be developed more fully in following chapters; here, the more general cross variogram is defined, interpretation principles are discussed, and modeling with the linear model of coregionalization is presented.

Here is classical definition of the semivariogram for one variable (Z):

$$\begin{aligned}\gamma_{Z,Z}(\mathbf{h}) &= \frac{1}{2}E\left\{[Z(\mathbf{u}) - Z(\mathbf{u}+\mathbf{h})]^2\right\} \\ &= \frac{1}{2}E\left\{(Z(\mathbf{u}) - Z(\mathbf{u}+\mathbf{h}))\,(Z(\mathbf{u}) - Z(\mathbf{u}+\mathbf{h}))\right\}\end{aligned}$$

This is extended to the cross semivariogram for two variables (Z and Y):

$$\gamma_{Z,Y}(\mathbf{h}) = \frac{1}{2}E\left\{(Z(\mathbf{u}) - Z(\mathbf{u}+\mathbf{h}))\,(Y(\mathbf{u}) - Y(\mathbf{u}+\mathbf{h}))\right\} \tag{4.14}$$

The variogram is the average product of the difference $z(\mathbf{u}) - z(\mathbf{u}-\mathbf{h})$ multiplied by itself. The cross variogram is the average product of the z-difference, $z(\mathbf{u}) - z(\mathbf{u}+\mathbf{h})$, and the y-difference for the same location and lag distance, $y(\mathbf{u}) - y(\mathbf{u}+\mathbf{h})$. The cross variogram is a straightforward extension of the variogram.

To better understand the cross variogram, consider standardized variables (mean and variance are 0 and 1, respectively) and expand the squared term:

$$\gamma_{Z,Y}(\mathbf{h}) = \frac{1}{2}E\left\{Z(\mathbf{u})\cdot Y(\mathbf{u}) - Z(\mathbf{u}+\mathbf{h})\cdot Y(\mathbf{u}) - \right. \tag{4.15}$$

$$\left. Z(\mathbf{u})\cdot Y(\mathbf{u}+\mathbf{h}) + Z(\mathbf{u}+\mathbf{h})\cdot Y(\mathbf{u}+\mathbf{h})\right\}$$

Recall the definition of the covariance and correlation coefficient for normalized variables from Section 4.1:

$$\begin{aligned}
\gamma_{Z,Y}(\mathbf{h}) &= \frac{1}{2}\left[C_{ZY}(0) + C_{ZY}(0) - C_{ZY}(\mathbf{h}) - C_{YZ}(\mathbf{h})\right] \\
&= C_{ZY}(0) - \frac{1}{2}\left[C_{ZY}(\mathbf{h}) + C_{YZ}(\mathbf{h})\right] \\
&= \rho_{ZY}(0) - \frac{1}{2}\left[\rho_{ZY}(\mathbf{h}) + \rho_{YZ}(\mathbf{h})\right] \\
&= \rho_{ZY}(0) - \rho_{ZY}(\mathbf{h}) \qquad (4.16)
\end{aligned}$$

The replacement of covariances with correlation coefficients, for example., $C_{ZY}(\mathbf{h})$ with $\rho_{ZY}(\mathbf{h})$ is possible only with standardized variables where $\sigma_Z^2 = \sigma_Y^2 = 1$. Setting $\rho_{ZY}(\mathbf{h}) = \rho_{YZ}(\mathbf{h})$ amounts to assume that the Z and Y covariances are "symmetric."[3]

The cross semivariogram at $\mathbf{h} = 0$ is equal to 0, just like the semivariogram, because $\rho_{ZY}(\mathbf{h} = 0) = \rho_{ZY}(0)$, see equation (4.16). At large lag distances the Z and Y variables often show no correlation, that is, $\rho_{ZY}(\mathbf{h})$ approaches zero. Therefore, the sill of the cross semivariogram is the correlation coefficient of collocated Z and Y values, $\rho_{ZY}(0)$, see equation (4.16). This assumes that the Z and Y variables have been standardized. The sill of a semivariogram is the variance; the sill of a cross semivariogram is the covariance (correlation coefficient for standardized variables) at lag 0.

The covariance or correlation at lag 0 can be negative and so can the cross variogram. The relation in equation (4.16) can be used to interpret the cross-variogram. Like the variogram, values above the sill imply negative correlation and values below the sill imply positive correlation. In fact, all of the interpretation principles discussed above including geometric anisotropy, trends, cyclicity, and zonal anisotropy are applicable to cross variograms.

Example

As an example of correlated variables, consider 1396 measurements of % bitumen and % fines from an oil sands operation in Northern Alberta. These data are distributed in the 2-D horizontal plane over an area about 10 km

[3] Non-symmetric cross covariances are encountered when correlated Z and Y variables are somehow "offset" from one another. This sometimes happens with well logs. Fitting and using non-symmetric cross covariances is not conveniently handled by any software; therefore, the data should be depth shifted and modeled with conventional symmetric cross covariances or cross variograms.

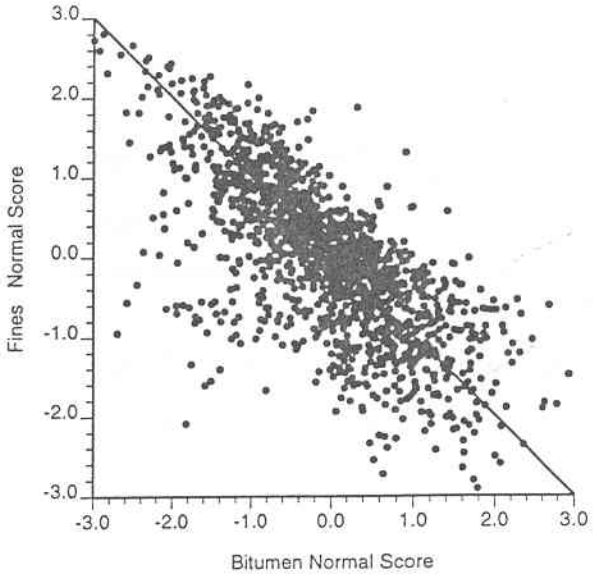


Figure 4.24: Cross plot of % fines (normal score transform) versus % bitumen (normal score transform). The correlation $\rho_{ZY}(0)$ is about -0.73.

square. % bitumen and % fines are the two critical factors affecting recovery of hydrocarbon in oil sands operations. Evaluating their spatial variability is important for process control in the extraction plant. Consider the normal score transforms of each variable. The cross plot of the collocated measurements is shown in Figure 4.24. Note the correlation $\rho_{ZY}(0)$ is -0.73, which implies that the sill of the cross variogram model should be about negative 0.73. This negative correlation is geologically reasonable, that is, the more fines there are the less space there is for bitumen.

The variogram of the normal score transform of % bitumen is shown in Figure 4.25, the variogram of the normal score transform of % fines is shown in Figure 4.26, and the cross variogram between the two is shown in Figure 4.27.

One cannot fit variograms independently since the corresponding (cross) covariances must be jointly positive definite. The Linear Model of Coregionalization (LMC) is used almost exclusively for modeling variograms of two or more variables.[4] The linear model of coregionalization (LMC) takes the following form:

$$\gamma_{Z,Z}(\mathbf{h}) = b^0_{Z,Z} + b^1_{Z,Z} \cdot \Gamma^1(\mathbf{h}) + b^2_{Z,Z} \cdot \Gamma^2(\mathbf{h}) \ldots$$

$$\gamma_{Z,Y}(\mathbf{h}) = b^0_{Z,Y} + b^1_{Z,Y} \cdot \Gamma^1(\mathbf{h}) + b^2_{Z,Y} \cdot \Gamma^2(\mathbf{h}) \ldots$$

$$\gamma_{Y,Y}(\mathbf{h}) = b^0_{Y,Y} + b^1_{Y,Y} \cdot \Gamma^1(\mathbf{h}) + b^2_{Y,Y} \cdot \Gamma^2(\mathbf{h}) \ldots$$

[4] Other models of coregionalization related to the collocated cokriging and other such models are but limit cases of this linear model of coregionalization.

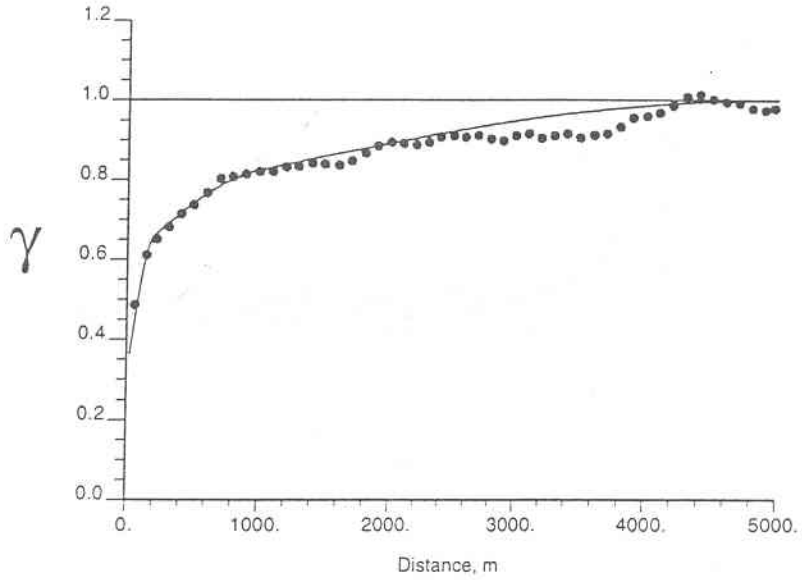


Figure 4.25: Variogram of % bitumen (normal score transform).

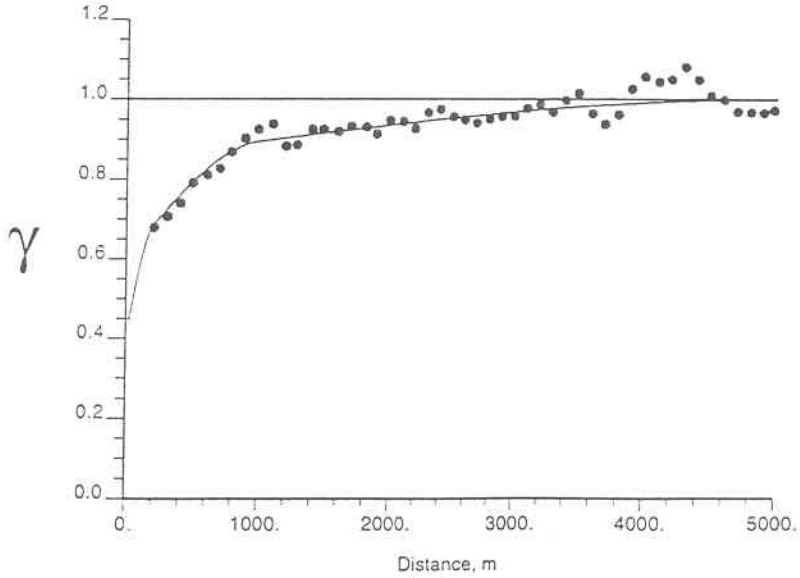


Figure 4.26: Variogram of % fines (normal score transform).

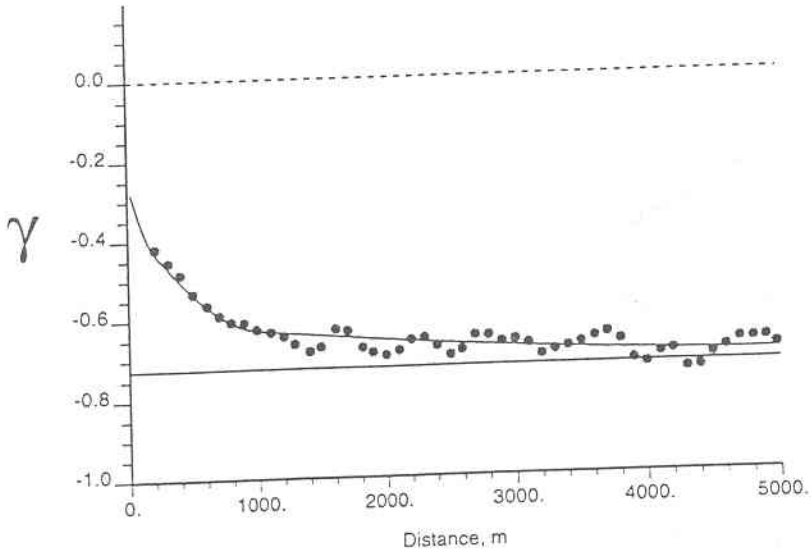


Figure 4.27: Cross variogram of % fines (normal score transform) and % bitumen (normal score transform).

where the $\Gamma^i, i = 1, \ldots, n_{st}$s are nested structures made up of common variogram models, as presented earlier. So, the LMC amounts to model each direct and cross variogram with the same variogram nested structures. The sill parameters (the b-values) are allowed to change within the following constraints:

$$\left.\begin{array}{c} b^i_{Z,Z} > 0 \\ b^i_{Y,Y} > 0 \\ b^i_{Z,Z} \cdot b^i_{Y,Y} \geq b^i_{Z,Y} \cdot b^i_{Z,Y} \end{array}\right\} \forall\, i$$

The first step in fitting a LMC is to choose the pool of nested structures: $\Gamma^i, i = 1, \ldots, n_{st}$. Some considerations are as follows:

- Each nested structure is defined by a variogram type (shape) and anisotropic ranges (angles defining directions of anisotropy and respective ranges in each direction).
- The nested structures should be chosen so that all of the *deemed* significant features on the experimental variograms can be modeled.

An automatic fitting algorithm could be considered, see Goovaerts [110] for more details.

The fitted variogram curves in Figures 4.25, 4.26, and 4.27 are from the following model:

$$\begin{aligned}
\gamma_Z(\mathbf{h}) &= 0.3 + 0.3 \cdot \Gamma^1(\mathbf{h}) + 0.25 \cdot \Gamma^2(\mathbf{h}) + 0.15 \cdot \Gamma^3(\mathbf{h}) \\
\gamma_{Z,Y}(\mathbf{h}) &= -0.25 - 0.1 \cdot \Gamma^1(\mathbf{h}) - 0.25 \cdot \Gamma^2(\mathbf{h}) - 0.1 \cdot \Gamma^3(\mathbf{h}) \\
\gamma_Y(\mathbf{h}) &= 0.4 + 0.2 \cdot \Gamma^1(\mathbf{h}) + 0.25 \cdot \Gamma^2(\mathbf{h}) + 0.15 \cdot \Gamma^3(\mathbf{h})
\end{aligned}$$

where $\Gamma^1(\mathbf{h})$ is spherical with range 200 m, $\Gamma^2(\mathbf{h})$ is spherical with range 1000 m, and $\Gamma^3(\mathbf{h})$ is spherical with range 5000 m. This is a licit model of coregionalization since $0.3 \cdot 0.4 \geq (-0.25)^2$, $0.3 \cdot 0.2 \geq (-0.1)^2$, $0.25 \cdot 0.25 \geq (-0.25)^2$, and $0.15 \cdot 0.15 \geq (-0.1)^2$.

The linear model of coregionalization could be extended to three or more variables; however, that is rarely done in practice. As mentioned at the start of this section, variables are typically considered pair wise.

Markov Models of Coregionalization

Many people consider the linear model of coregionalization cumbersome to apply. A simpler alternative consists of modeling the cross variogram with a single correlation coefficient parameter. This simplification is reasonable when the cross variogram cannot be fit reliably from the data. The simplification results from a Markov-type assumption whereby one data type prevails completely over collocated hard and soft data. The classical Markov model, sometimes called Markov Model I, assumes that hard Z data prevails over soft Y data. The cross variogram is given by:

$$C_{ZY}(\mathbf{h}) = B \cdot C_Z(\mathbf{h}), \quad \forall \mathbf{h} \tag{4.17}$$

where $B = \sqrt{\sigma_Z^2/\sigma_Y^2} \cdot \rho_{ZY}(0)$, σ_Z^2, σ_Y^2 are the variances of Z and Y, and $\rho_{ZY}(0)$ is the correlation coefficient of collocated $z - y$ data. Application of this model of coregionalization requires the covariance of the Z variable $(\sigma_Z^2 - \gamma_Z(\mathbf{h}))$ and the correlation coefficient of collocated data.

A second Markov model, sometimes called Markov Model II [150, 222], assumes that soft Y data prevail over hard Z data. The cross variogram is given by:

$$C_{ZY}(\mathbf{h}) = B \cdot C_Y(\mathbf{h}), \quad \forall \mathbf{h} \tag{4.18}$$

where B is the same as above. Application of this model of coregionalization requires the covariance of the Y variable $(\sigma_Y^2 - \gamma_Y(\mathbf{h}))$ and the correlation coefficient of collocated data.

The full LMC gives greater flexibility and better results that either Markov model; however, at the expense of greater effort. The choice of Markov model depends on whether the cross variogram looks more like the Z variogram or Y variogram. When the Y data is at a larger volume support, for example, coming from seismic, then the second Markov Model II is likely most appropriate.

There are implementation details to be addressed in the Markov approach including the *other* variogram, that is, the Y variogram if Markov Model I

is used or the Z variogram if Markov Model II is used. There is no problem if a single collocated datum is retained for estimation. These and other application details are addressed in subsequent chapters.

4.7 Work Flow

The central idea of this chapter is to quantify spatial correlation to be used in subsequent geostatistical modeling. A number of work flows are related to this quantification.

Figure 4.28 illustrates the work flow to prepare data for variogram calculation. A categorical data variable must be coded as an indicator. A continuous variable must have deterministic trends removed, normal score transformed for subsequent Gaussian simulation, and indicator transformed for subsequent indicator simulation of continuous variables.

Figure 4.29 illustrates the work flow to calculate the variogram of scattered data. The common case of vertical wells permits the calculation of a vertical variogram. Horizontal variograms must be inferred from analogue data when there are few wells. In presence of sufficient data we can calculate the omnidirectional horizontal variogram and then directional variograms.

Figure 4.30 illustrates the work flow to create a 3-D variogram model. A 3-D model is assembled once the three principal directional variograms have been determined. The range parameters in the three directions can be changed to achieve good fits in all directions.

Figure 4.31 illustrates the work flow to fit a linear model of coregionalization. Collocated cosimulation does not require a linear model of coregionalization, but full cokriging (or cosimulation) requires such a model. The same pool of nested structures is used with coefficients that must meet the criteria for positive definiteness.

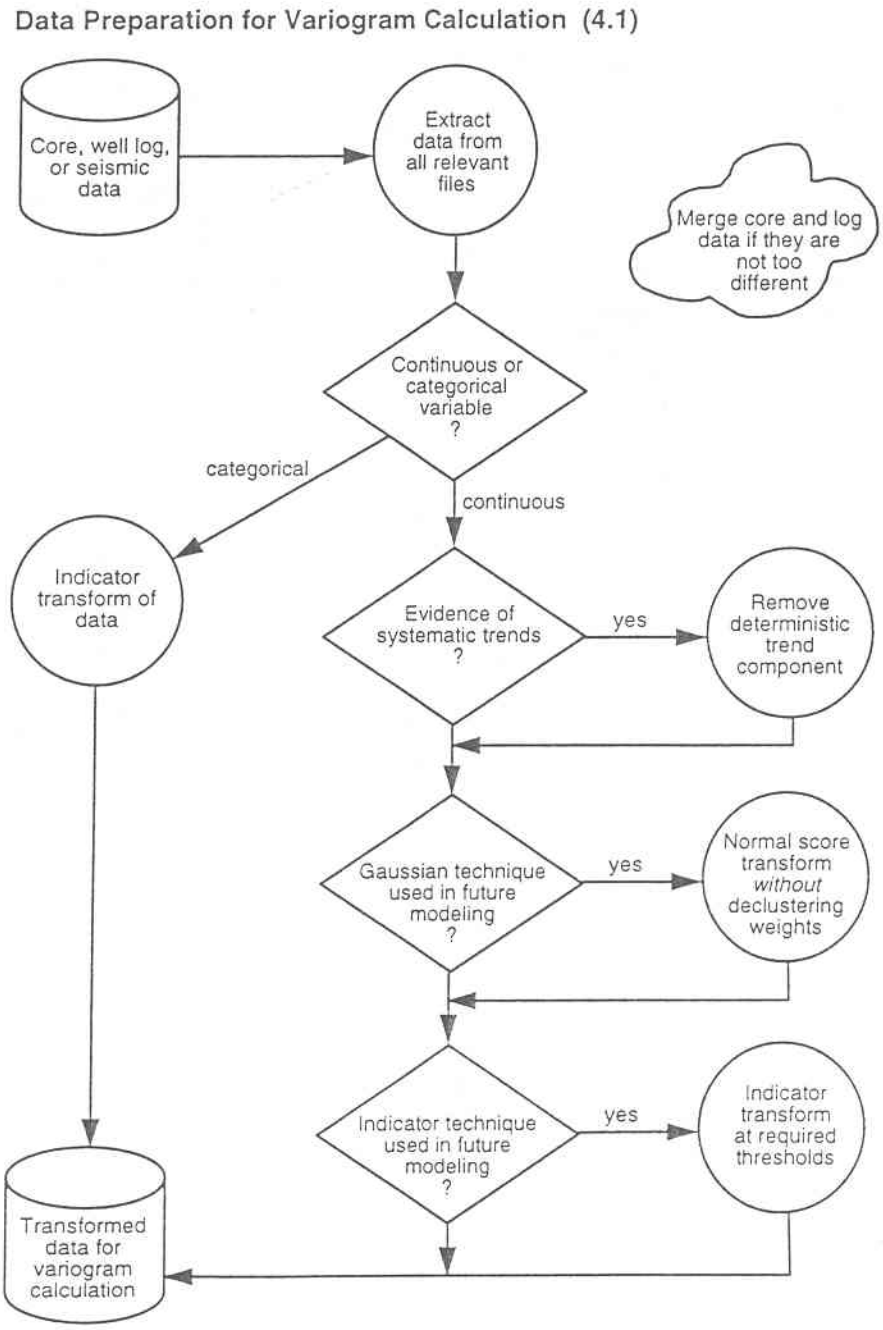


Figure 4.28: Work flow for data extraction and transformation for variogram calculation.

Variogram Calculation of Scattered Data (4.2)

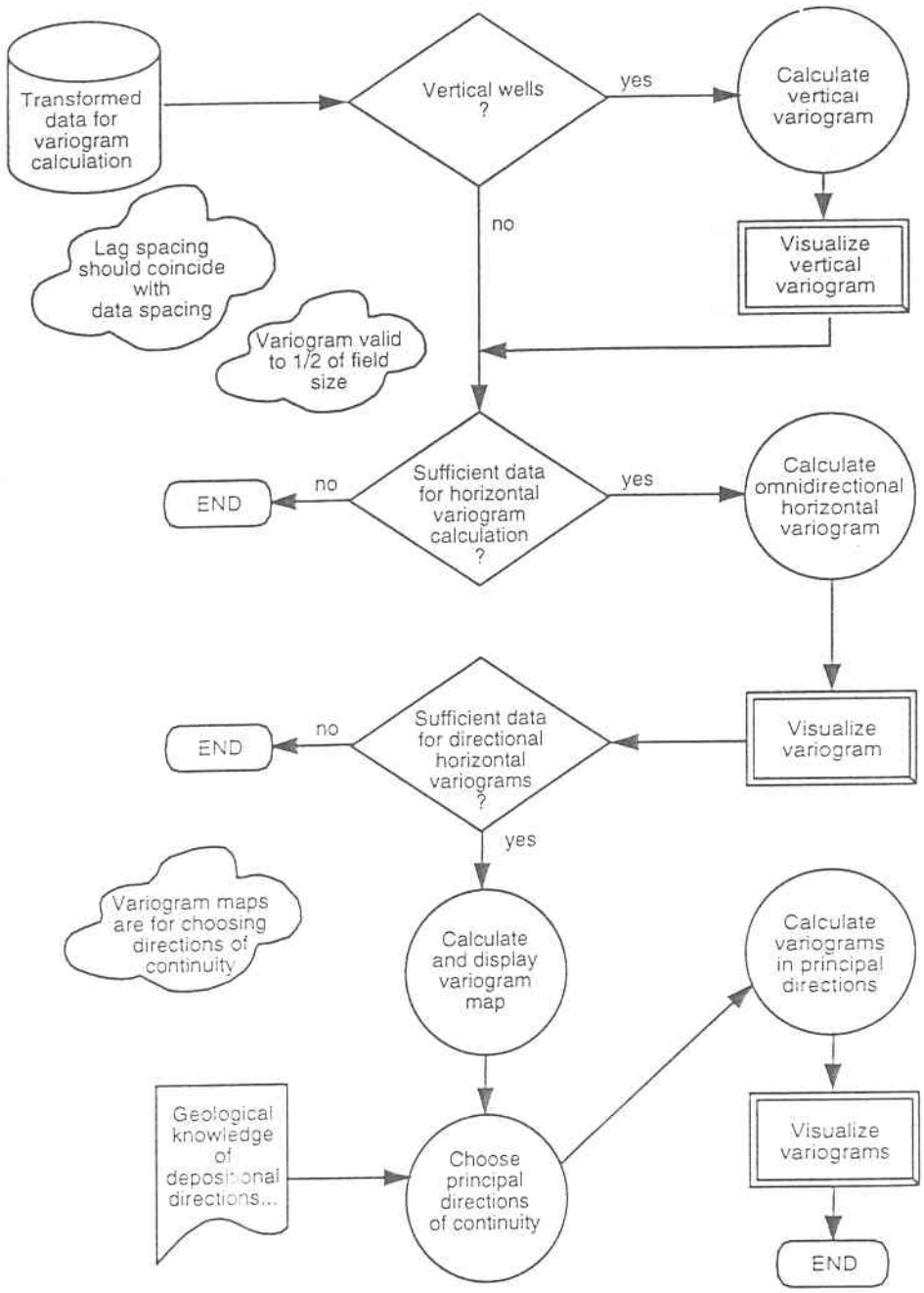


Figure 4.29: Work flow for calculation of variograms of scattered data.

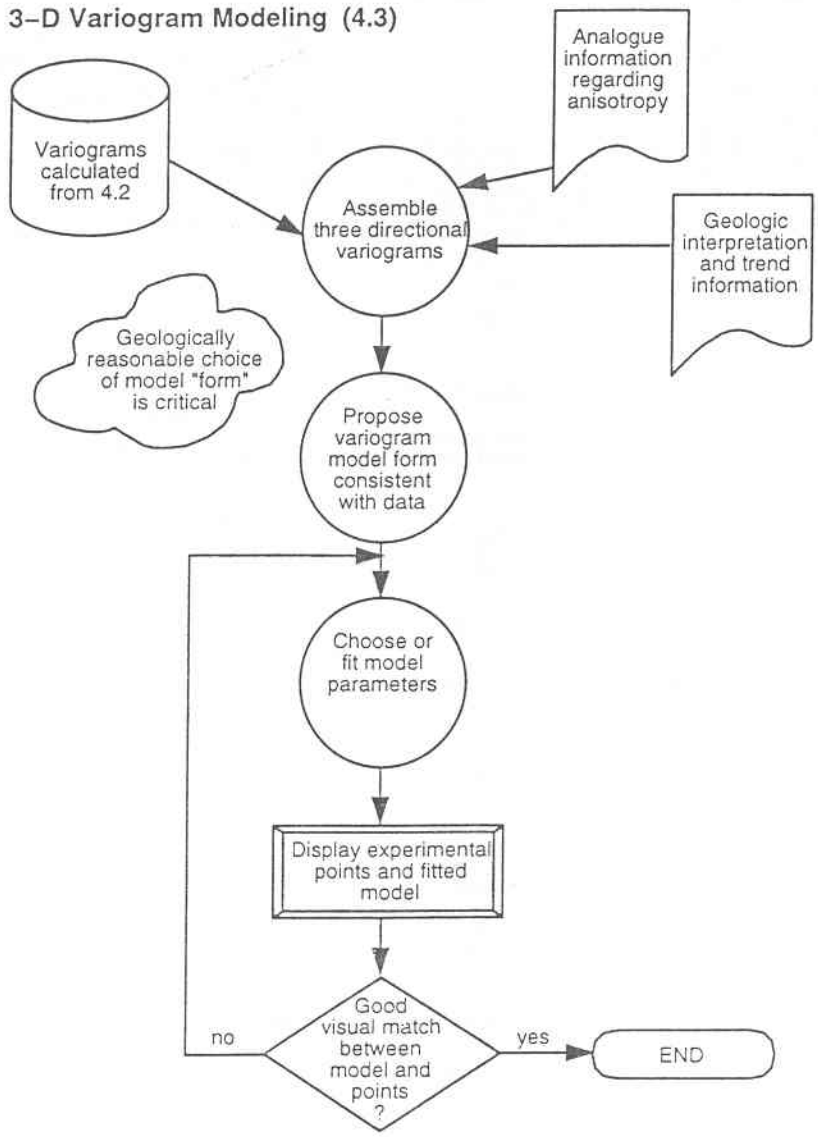


Figure 4.30: Work flow for 3-D variogram modeling.

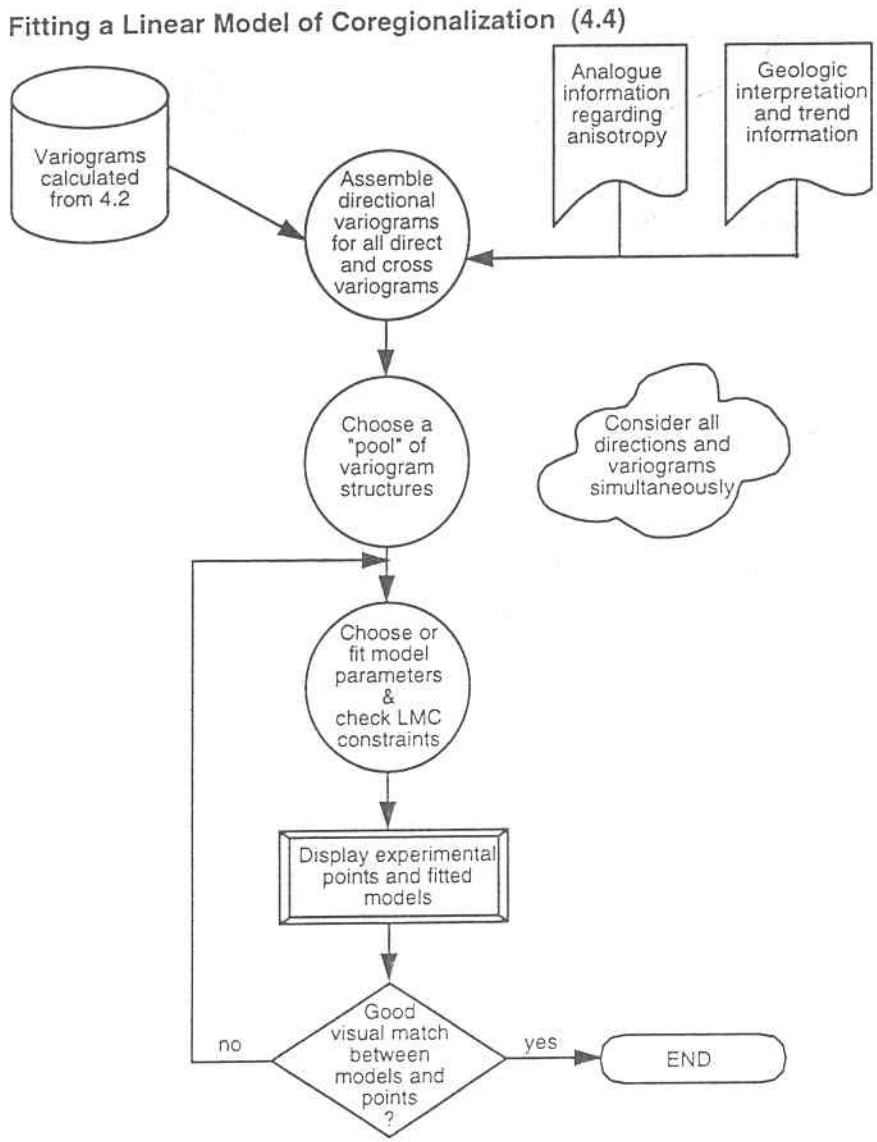


Figure 4.31: Work flow for fitting a linear model of coregionalization.

Chapter 5

Preliminary Mapping Concepts

Geostatistics is concerned with a wide variety of techniques for spatial data analysis, estimation, simulation, and decision making. The field of geostatistics is extensive; the goal here is a pragmatic presentation of the concepts used in reservoir modeling. Kriging, cokriging, and simulation concepts are condensed into a single reservoir-oriented chapter.

Kriging is the workhorse of traditional mapping applications and an essential component of geostatistical simulation methods. For that reason, Section 5.1 gives a detailed description of estimation variance, the simple kriging equations, and cokriging. Although this section could be skipped at first reading, an understanding of this material is required for an understanding of the commonly used simulation methods presented in Chapters 6 and 8.

Section 5.2 presents sequential Gaussian simulation, which has become a most popular algorithm for continuous property modeling. The popularity of this algorithm is due to its relative simplicity and effectiveness at creating numerical models with correct spatial statistics.

Section 5.3 extends the concept to direct sequential simulation, that is, sequential simulation that does not depend on a prior Gaussian transformation and back transformation afterwards. Transformation, however, is still required: implementation details are presented.

Section 5.4 presents indicator methods for uncertainty assessment and sequential simulation. Probability or "p"-field simulation methods have gained popularity because of their flexibility in accounting for secondary information such as seismic data; Section 5.5 presents the p-field methods.

Mapping geological variables becomes complex due to many special considerations such as discontinuities, trends, and changing directions of continuity. Section 5.6 presents a number of these special considerations. In many respects, this section is one of the most important of the book. The considerations discussed here are relevant even *before* statistical analysis and

spatial statistics. The need, however, to discuss trends or locally varying means comes up after a presentation of key kriging and simulation concepts.

The mathematics presented in this chapter are not difficult, but they may appear odd to the newcomer to geostatistics or statistical inference. Little of this is required from a practical perspective; however, some of the implementation details and limitations of geostatistical reservoir models can only be appreciated after working through these basics. Start with Section 5.6 if you wish, but come back to the rest for more in-depth understanding.

5.1 Kriging

Background

A common problem in earth sciences is creating a map of a regionalized variable from limited sample data. This common problem was initially addressed by hand-contouring, which provides insight into trends and the uncertainty in the data.[1] Early machine contouring algorithms evolved from principles of hand-contouring. The goal of such algorithms was to create smooth maps that would reveal the same geological trends as hand-contouring. Such algorithms remain popular.

There were others, primarily mathematicians and engineers, who wanted to create maps fit for a particular purpose. That is, they considered that the mapped values should be optimal in some objective sense. One measure of optimality is *unbiasedness*, that is, taken all together, the map should have the correct average. It turns out that global unbiasedness is straightforward to achieve; however, with simplistic estimation methods global unbiasedness is achieved at the expense of overestimating low values and underestimating high values. The estimates are said to have *conditional bias*. Daniel Krige, a South African mining engineer, was interested in correcting this bias because it was unacceptable to have mine stopes with predicted high grades systematically have lower true grades. The interpolation method of kriging developed by Georges Matheron (1961/62), was named for Daniel Krige's pioneering work in the early 1950s [166].

Least squares optimization has been used for more than 200 years. The idea proposed by early workers in geostatistics [103, 184] was to construct an estimator $Z^*(\mathbf{u})$ that is optimum in a minimum expected squared error sense, that is, minimize the average squared difference between the true value $z(\mathbf{u})$ and the estimator:

$$SE = [z(\mathbf{u}) - z^*(\mathbf{u})]^2 \tag{5.1}$$

Of course, the true values are only known at the data locations and *not* at the locations being estimated. Therefore, as is classical in statistics, the squared

[1] Hand-contouring important variables such as the thickness of net pay should always be performed as a means to investigate trends, anisotropy, and other essential features of the data.

error SE is minimized in expected value over all locations $\mathbf{u}$ with the same data configuration within a study area A, deemed stationary.

Most estimators in geostatistics are linear. The original data may have been transformed in a non-linear fashion into polynomial coefficients, indicators, or power-law transforms; however, the estimates of the transformed variables remain linear, that is,

$$z^*(\mathbf{u}) - m(\mathbf{u}) = \sum_{\alpha=1}^{n} \lambda_\alpha \cdot [z(\mathbf{u}_\alpha) - m(\mathbf{u}_\alpha)] \tag{5.2}$$

where $z^*(\mathbf{u})$ is the estimate at unsampled location $\mathbf{u}$, $m(\mathbf{u})$ is the prior mean value at unsampled location $\mathbf{u}$, $\lambda_\alpha, \alpha = 1, \ldots, n$ are weights applied to the n data. $z(\mathbf{u}_\alpha), \alpha = 1, \ldots, n$ are the n data values, and $m(\mathbf{u}_\alpha), \alpha = 1, \ldots, n$ are the n prior mean values at the data locations. All prior mean values could be set to a constant mean $m(\mathbf{u}) = m(\mathbf{u}_\alpha) = m$ if no prior information on trends is available.

From a classical regression estimator perspective we could imagine adding terms involving jointly 2 or more data such as the product $[z(\mathbf{u}_\alpha) - m(\mathbf{u}_\alpha)]$ $[z(\mathbf{u}_\beta) - m(\mathbf{u}_\beta)]$; however, this is not done in practice. The most important reason to limit ourselves to linear estimates is for simplicity. Incorporation of many data simultaneously would require measure of correlation between these multiple data and the property being estimated, which is a multiple point measure of correlation and not just the two-point variogram or covariance we have considered thus far.

The following classical regression approach to arrive at the weights $\lambda_\alpha, \alpha = 1, \ldots, n$ is known as kriging for historical reasons.

Estimation Variance

A principle of numerical modeling is that known geological and geometrical constraints should be accounted for deterministically. For that reason, we systematically consider residuals from known trends, that is, we work with residual data values:

$$Y(\mathbf{u}_\alpha) = Z(\mathbf{u}_\alpha) - m(\mathbf{u}_\alpha), \quad \alpha = 1, \ldots, n \tag{5.3}$$

The subject of trend modeling, that is, determining the trend or prior mean $m(\mathbf{u})$ for all locations $\mathbf{u}$ in the reservoir, is contained in Section 5.6 at the end of this chapter. When there are no large-scale predictable trends, the mean could be considered known and fixed at the global mean: $m(\mathbf{u}) = m$ for all locations in the area or reservoir $\mathbf{u} \in A$.

Regardless of the decision for the mean m the residual variable Y is deemed stationary, that is, $E\{Y\mathbf{u})\} = 0$, $\forall\ \mathbf{u} \in A$ with stationary covariance $C_Y(\mathbf{h})$ and variogram $2\gamma_Y(\mathbf{h})$. The adoption of Gaussian techniques would involve a normal score transform of the Y residual data variable. The decision of stationarity would apply to the normal score variable, and the variogram would be calculated after normal score transformation. As mentioned

before, stationarity assumes that the statistics apply over the study area, for example, the porosity within a particular facies in a particular reservoir layer.

Although it is somewhat repetitive, there is value in reviewing the needed statistics. The stationary mean or expected value of the Y variable is 0, that is, $E\{Y\} = 0$. The stationary variance is $Var\{Y\} = E\{Y^2\} = C(0) = \sigma^2$. The variogram of the stationary residuals is defined as:

$$2\gamma(\mathbf{h}) = E\left\{[Y(\mathbf{u}) - Y(\mathbf{u}+\mathbf{h})]^2\right\} \tag{5.4}$$

and the covariance is defined as:

$$C(\mathbf{h}) = E\left\{Y(\mathbf{u}) \cdot Y(\mathbf{u}+\mathbf{h})\right\} \tag{5.5}$$

The link between the variogram and the covariance is:

$$C(\mathbf{h}) = C(0) - \gamma(\mathbf{h}) \tag{5.6}$$

This result is important since it permits us the flexibility of calculating and modeling experimental variograms and then using the covariance in subsequent mathematical calculations such as kriging.

Now, consider a linear estimator at a location where no data is available:

$$Y^*(\mathbf{u}) = \sum_{\alpha=1}^{n} \lambda_\alpha \cdot Y(\mathbf{u}_\alpha) \tag{5.7}$$

where $Y^*(\mathbf{u})$ is the estimated value, $\lambda_\alpha, \alpha = 1, \ldots, n$ are weights applied to the n data values $Y(\mathbf{u}_\alpha), \alpha = 1, \ldots, n$. The error variance for this estimator may be defined as:

$$E\left\{[Y^*(\mathbf{u}) - Y(\mathbf{u})]^2\right\} = \sigma_E^2(\mathbf{u}) \tag{5.8}$$

Although we do not know the true value $y(\mathbf{u})$, it is possible to expand this error variance term using expected values,[2] the error variance is expanded as:

$$\begin{aligned}
&= E\left\{[Y^*(\mathbf{u})]^2\right\} - 2 \cdot E\left\{Y^*(\mathbf{u}) \cdot Y(\mathbf{u})\right\} + E\left\{[Y(\mathbf{u})]^2\right\} \\
&= \sum_{\alpha=1}^{n}\sum_{\beta=1}^{n} \lambda_\alpha \lambda_\beta E\left\{Y(\mathbf{u}_\beta) \cdot Y(\mathbf{u}_\alpha)\right\} - 2 \cdot \sum_{\alpha=1}^{n} \lambda_\alpha E\left\{Y(\mathbf{u}) \cdot Y(\mathbf{u}_\alpha)\right\} + C(0) \\
&= \underbrace{\sum_{\alpha=1}^{n}\sum_{\beta=1}^{n} \lambda_\alpha \lambda_\beta C(\mathbf{u}_\beta - \mathbf{u}_\alpha)}_{\text{redundancy}} - 2 \cdot \underbrace{\sum_{\alpha=1}^{n} \lambda_\alpha C(\mathbf{u} - \mathbf{u}_\alpha)}_{\text{closeness}} + \underbrace{C(0)}_{\text{variance}}
\end{aligned} \tag{5.9}$$

[2]This expression for the estimation variance is derived using (1) the linearity of the expected value operator, that is, $E\{A+B\} = E\{A\} + E\{B\}$, and (2) the implicit assumption of stationarity, for example, $E\{Y(\mathbf{u}_\beta) \cdot Y(\mathbf{u}_\alpha)\} = C(\mathbf{h} = \mathbf{u}_\beta - \mathbf{u}_\alpha)$.

This final equation for the error variance is very interesting; it is a mathematical expression for the error variance for any set of weights $\lambda_\alpha, \alpha = 1, \ldots, n$. Clearly, the estimation variance depends on the model of covariance or variogram. There are three terms in the equation for the estimation variance:

1. *Redundancy:* the more redundant the data (the covariances between them, $C(\mathbf{u}_\alpha, \mathbf{u}_\beta)$, will be higher) the larger the estimation variance.
2. *Closeness:* the closer the data to the location being estimated (the covariances $C(\mathbf{u}-\mathbf{u}_\alpha)$ will be higher), the smaller the estimation variance.
3. *Variance:* the estimation variance is equal to the variance or $C(0)$ when all data are too distant to receive any weight, that is, $\lambda_\alpha = 0, \alpha = 1, \ldots, n$ and the estimate is the local mean $Y^*(\mathbf{u}) = m(\mathbf{u})$ assumed known.

Equation (5.9) allows us to calculate the estimation variance for any set of weights $\lambda_\alpha, \alpha = 1, \ldots, n$; however, our goal is to calculate the weights that minimize this estimation variance. A classical minimization procedure is followed.

Simple Kriging

The partial derivatives of expression (5.9) with respect to each of the weights $\lambda_\alpha, \alpha = 1, \ldots, n$ are calculated and set to zero. This leads to equations that can be solved for the weights that minimize the estimation variance.[3] This procedure is classical in mathematics. The partial derivatives of expression (5.9) with respect to the weights $\lambda_\alpha, \alpha = 1, \ldots, n$ are calculated as:

$$\frac{\partial \left[\sigma_E^2(\mathbf{u})\right]}{\partial \lambda_\alpha} = 2 \cdot \sum_{\beta=1}^{n} \lambda_\beta C(\mathbf{u}_\beta - \mathbf{u}_\alpha) - 2 \cdot C(\mathbf{u} - \mathbf{u}_\alpha), \quad \alpha = 1, \ldots, n \tag{5.10}$$

These may be set to zero to calculate the weights that minimize the estimation variance:

$$\sum_{\beta=1}^{n} \lambda_\beta C(\mathbf{u}_\beta - \mathbf{u}_\alpha) = C(\mathbf{u} - \mathbf{u}_\alpha), \quad \alpha = 1, \ldots, n \tag{5.11}$$

This system of n equations with n unknown weights is known as the simple kriging (SK) system. These equations are also called the normal equations in optimization theory [176].

As an example, consider a case with three data values:

$$\begin{aligned}
C(\mathbf{u}_1 - \mathbf{u}_1) \cdot \lambda_1 + C(\mathbf{u}_1 - \mathbf{u}_2) \cdot \lambda_2 + C(\mathbf{u}_1 - \mathbf{u}_3) \cdot \lambda_3 &= C(\mathbf{u} - \mathbf{u}_1) \\
C(\mathbf{u}_2 - \mathbf{u}_1) \cdot \lambda_1 + C(\mathbf{u}_2 - \mathbf{u}_2) \cdot \lambda_2 + C(\mathbf{u}_2 - \mathbf{u}_3) \cdot \lambda_3 &= C(\mathbf{u} - \mathbf{u}_2) \\
C(\mathbf{u}_3 - \mathbf{u}_1) \cdot \lambda_1 + C(\mathbf{u}_3 - \mathbf{u}_2) \cdot \lambda_2 + C(\mathbf{u}_3 - \mathbf{u}_3) \cdot \lambda_3 &= C(\mathbf{u} - \mathbf{u}_3)
\end{aligned}$$

[3] The second derivative could be taken to show that we have a minimum; however, the maximum estimation variance is infinity, so any solution to kriging must be a minimum.

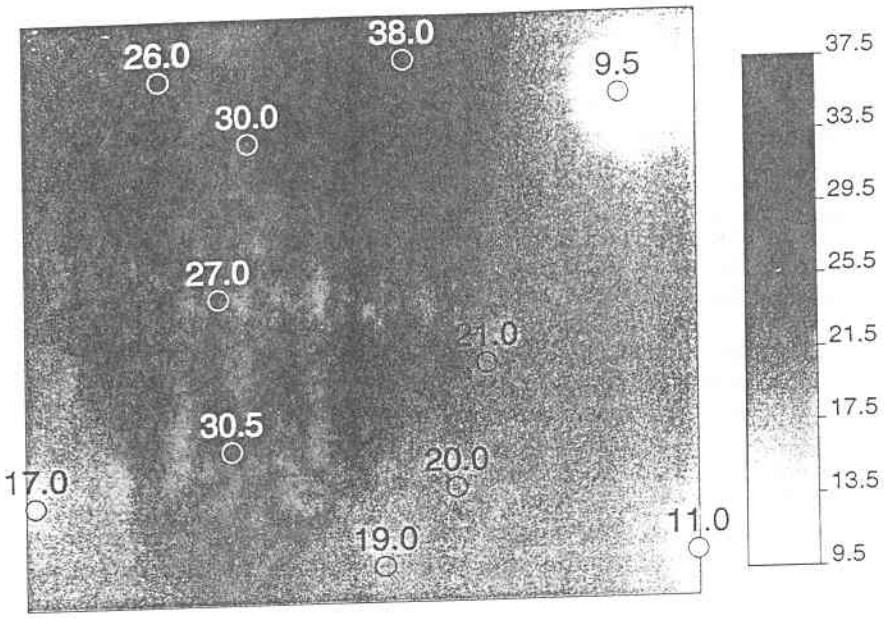


Figure 5.1: A kriged map of thickness using a Gaussian variogram. Note the smooth interpolation between the data points.

which may be written in matrix notation as:

$$\begin{bmatrix} C(\mathbf{u}_1-\mathbf{u}_1) & C(\mathbf{u}_1-\mathbf{u}_2) & C(\mathbf{u}_1-\mathbf{u}_3) \\ C(\mathbf{u}_2-\mathbf{u}_1) & C(\mathbf{u}_2-\mathbf{u}_2) & C(\mathbf{u}_2-\mathbf{u}_3) \\ C(\mathbf{u}_3-\mathbf{u}_1) & C(\mathbf{u}_3-\mathbf{u}_2) & C(\mathbf{u}_3-\mathbf{u}_3) \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{bmatrix} = \begin{bmatrix} C(\mathbf{u}-\mathbf{u}_1) \\ C(\mathbf{u}-\mathbf{u}_2) \\ C(\mathbf{u}-\mathbf{u}_3) \end{bmatrix}$$

$$\mathbf{C}\lambda = \mathbf{c}$$

The weights are obtained by inverting the left-hand-side $\mathbf{C}$ matrix and multiplying it to the right-hand-side $\mathbf{c}$ vector. Note that the left-hand-side $\mathbf{C}$ matrix contains all of the information related to redundancy in the data and the right-hand-side $\mathbf{c}$ vector contains all of the information related to closeness of the data to the location being estimated.

Kriging solves this system of linear equations. The matrix $\mathbf{C}$ is symmetric. The matrix is positive semidefinite, provided the covariance (or the corresponding variogram) is fit with a positive definite function and no two data of the same support are co-located.

Figure 5.1 shows a map of kriged estimates using 11 data. The variogram in this case is a Gaussian variogram with no nugget effect and no anisotropy.

The Kriging Variance

The equation for the estimation variance may be simplified if simple kriging is used to calculate the weights, that is,

$$\sigma_E^2(\mathbf{u}) = \sum_{\alpha=1}^{n} \lambda_\alpha \underbrace{\sum_{\beta=1}^{n} \lambda_\beta C(\mathbf{u}_\beta - \mathbf{u}_\alpha)}_{\text{to be replaced}} - 2 \cdot \sum_{\alpha=1}^{n} \lambda_\alpha C(\mathbf{u} - \mathbf{u}_\alpha) + C(0)$$

which simplifies to

$$\sigma_E^2(\mathbf{u}) = C(0) - \sum_{\alpha=1}^{n} \lambda_\alpha C(\mathbf{u} - \mathbf{u}_\alpha) \tag{5.12}$$

when the "to be replaced" portion is replaced with the right-hand side of the kriging equations, equation (5.11).

Kriging may be considered a spatial regression, thus it creates an estimate that is often smooth. The combination of n data by a weighted linear sum tends away from low and high estimates; there are more near the mean. Kriging accounts for redundancy between the data and a measure of distance specific to the data considered, that is, the variogram model. Another significant advantage of kriging is that it provides a quantification of how smooth the estimates are, that is, the variance of the kriging estimate may be calculated as:

$$\begin{aligned} Var\{Y^*(\mathbf{u})\} &= E\left\{[Y^*(\mathbf{u})]^2\right\} - E\left\{y^*(\mathbf{u})\right\}^2 \\ &= \sum_{\alpha=1}^{n}\sum_{\beta=1}^{n} \lambda_\alpha \lambda_\beta E\left\{Y(\mathbf{u}_\beta) \cdot Y(\mathbf{u}_\alpha)\right\} - 0^2 \\ &= \sum_{\alpha=1}^{n} \lambda_\alpha C(\mathbf{u} - \mathbf{u}_\alpha) \end{aligned} \tag{5.13}$$

Once again, the last substitution in equation (5.13) is the right-hand side of the kriging equations (5.11). Equation (5.13) tells us the variance of the kriging estimator $Y^*(\mathbf{u})$ at any location $\mathbf{u}$. This variance is the stationary variance $C(0)$ when kriging at a data location since $\lambda_\alpha = 1$ for the data location $\mathbf{u} = \mathbf{u}_\alpha$, all other weights are zero, and $C(\mathbf{u} - \mathbf{u}_\alpha) = C(0)$, the stationary variance $C(0) = \sigma^2$. The variance far from local data is zero since no data are given any weight, that is, all weights are zero.

A map of kriged estimates is too smooth. This smoothing is particularly noticeable away from the data locations. The smoothing of kriging is directly proportional to the kriging variance; there is no smoothing at data locations where the kriging variance is zero; the estimate is completely smooth far away from data where all estimates are equal to the mean and the kriging variance

is the stationary variance $C(0)$. We can write the "missing" variance or the smoothing effect as:

$$\text{Missing Variance} = C(0) - \sum_{\alpha=1}^{n} \lambda_\alpha C(\mathbf{u}, \mathbf{u}_\alpha) \tag{5.14}$$

This rather unusual concept is important when we describe the theoretical background behind simulation.

Other Forms of Kriging

There are other versions of kriging that make different assumptions regarding the mean, see [62, 110, 157] for more information. Simple kriging of residuals from a geologically interpreted trend is the preferred approach because the theory of simulation requires simple kriging. Other types of kriging may be used for estimation or in special cases.

Ordinary kriging (OK) estimates a constant mean value m from the data used in the kriging neighborhood; then, simple kriging is done with residuals from this implicitly estimated mean. Ordinary kriging has seen wide application in map making because it is robust with respect to trends. There is a need, however, to have sufficient data to reliably estimate the mean at each location. This is not the case early in a reservoir lifecycle.

Universal kriging (UK) or kriging with a trend model (KT) assumes that the mean follows a particular polynomial shape, for example, linear in the N30E direction. That mean "surface" is fitted as a scaling of the polynomial trend. Then, simple kriging is performed with residuals from the implicitly calculated mean values. Universal kriging is unstable in extrapolation because the coefficients of the fitted mean surface become unstable. As with ordinary kriging, universal kriging requires sufficient data to reliably estimate the mean.

Kriging with an external drift considers a non-parametric trend shape that could come from a secondary variable such as seismic. The mean "surface" is fitted as a linear scaling of the external drift variable. Then, simple kriging is performed with residuals from the implicitly calculated mean values.

Cokriging

The term kriging is traditionally reserved for linear regression using data with the same attribute as that being estimated. For example, an unsampled porosity value $z(\mathbf{u})$ is estimated from neighboring porosity sample values defined on the same volume support.

The term cokriging is reserved for linear regression that also uses data defined on different attributes. For example, the porosity value $z(\mathbf{u})$ may be estimated from a combination of porosity samples and related acoustic impedance values.

In the case of a single secondary variable (Y_2), the simple cokriging estimator of $Y(\mathbf{u})$ is written:

$$Y^*_{COK}(\mathbf{u}) = \sum_{\alpha_1=1}^{n_1} \lambda_{\alpha_1} Y(\mathbf{u}_{\alpha_1}) + \sum_{\alpha_2=1}^{n_2} \lambda'_{\alpha_2} Y_2(\mathbf{u}'_{\alpha_2}) \tag{5.15}$$

where the λ_{α_1}s are the weights applied to the n_1 Y samples and the λ'_{α_2}s are the weights applied to the n_2 Y_2 samples. Expression (5.15) is written as simple cokriging for standardized variables, that is, the means of Y and Y_2 are assumed known and zero.

Kriging requires a model for the Y covariance. Cokriging requires a *joint* model for the matrix of covariance functions including the Y covariance $C_Y(\mathbf{h})$, the Y_2 covariance $C_Y2(\mathbf{h})$, the cross Y-Y_2 covariance $C_{YY_2}(\mathbf{h})$, and the cross Y_2-Y covariance $C_{Y_2Y}(\mathbf{h})$.

When K different variables are considered, the covariance matrix requires in all generality K^2 covariance functions. Such inference becomes extremely demanding in terms of data and the subsequent joint modeling is particularly tedious [110]. This is the main reason why cokriging has not been extensively used in practice. Algorithms such as collocated cokriging and Markov models (see hereafter) have been developed to shortcut the tedious inference and modeling process required by cokriging.

Other than tedious variogram or covariance inference, cokriging is the same as kriging. The cokriging equivalents of OK and UK exist where the mean values are implicitly estimated from the neighborhood data. The reader is referred to [34, 68, 110, 157, 195, 194, 259] for details.

Collocated Cokriging

A reduced form of cokriging consists of retaining only the collocated secondary data $y_2(\mathbf{u})$ or relocated data $y_2(\mathbf{u}')$ to the nearest node $\mathbf{u}$ being estimated [68, 274]. This is not a problem if the distance $|\mathbf{u} - \mathbf{u}'|$ is small with respect to the volume of influence of $y_2(\mathbf{u})$. The cokriging estimator is written:

$$Y^*_{COK}(\mathbf{u}) = \sum_{\alpha_1}^{n_1} \lambda_{\alpha_1} Y(\mathbf{u}_{\alpha_1}) + \lambda' Y_2(\mathbf{u}) \tag{5.16}$$

The corresponding cokriging system requires knowledge of the Y covariance $C_Y(\mathbf{h})$ and the Y-Y_2 cross covariance $C_{YY_2}(\mathbf{h})$. The latter can be approximated through the following model:

$$C_{YY_2}(\mathbf{h}) = B \cdot C_Z(\mathbf{h}), \quad \forall \mathbf{h} \tag{5.17}$$

where $B = \sqrt{C_Y(0)/C_{Y_2}(0)} \cdot \rho_{YY_2}(0)$, $C_Y(0)$, $C_{Y_2}(0)$ are the variances of Y and Y_2, and $\rho_{YY_2}(0)$ is the linear coefficient of correlation of collocated $y - y_2$ data.

The approximation consisting of retaining only the collocated secondary datum does not affect the estimate (close-by secondary data are typically

very similar in values), but it may affect the resulting cokriging estimation variance: that variance is overestimated, sometimes significantly. In a kriging (estimation) context this is not a problem, because kriging variances are of little use. In a simulation context, see Chapter 8.2, where the kriging variance defines the spread of the conditional distribution from which simulated values are drawn, this may be a problem. The collocated cokriging variance should then be reduced by a factor (assumed constant $\forall \mathbf{u}$) to be determined by trial and error.

5.2 Sequential Gaussian Simulation

There are many algorithms that can be devised to create stochastic simulations: (1) matrix approaches (LU Decomposition), which are not extensively used because of size restrictions (an N x N matrix must be solved where N, the number of locations, could be in the millions for reservoir applications), (2) turning bands methods where the variable is simulated on 1-D lines and then combined into a 3-D model; not commonly used because of artifacts, (3) spectral methods using FFTs can be CPU fast, but honoring conditioning data requires an expensive kriging step, (4) fractals, which are not used extensively because of the restrictive assumption of self-similarity, and (5) moving average methods, which are infrequently used due to CPU requirements.

The common approach adopted in recent times for reservoir modeling applications is the sequential Gaussian simulation (SGS) approach [137]. This method is simple, flexible, and reasonably efficient. Let's review the theory underlying SGS. Recall the simple kriging estimator:

$$Y^*(\mathbf{u}) = \sum_{\beta=1}^{n} \lambda_\beta \cdot Y(\mathbf{u}_\beta)$$

and the corresponding simple kriging system:

$$\sum_{\beta=1}^{n} \lambda_\beta C(\mathbf{u}_\alpha - \mathbf{u}_\beta) = C(\mathbf{u} - \mathbf{u}_\alpha), \;\; \mathbf{u}_\alpha = 1, \ldots, n$$

The covariance between the kriged estimate and one of the data values can be written as:

$$\begin{aligned} Cov\{Y^*(\mathbf{u}), Y(\mathbf{u}_\alpha)\} &= E\{Y^*(\mathbf{u}), Y(\mathbf{u}_\alpha)\} \\ &= E\left\{\left[\sum_{\beta=1}^{n} \lambda_\beta \cdot Y(\mathbf{u}_\beta)\right] \cdot Y(\mathbf{u}_\alpha)\right\} \\ &= \sum_{\beta=1}^{n} \lambda_\beta \cdot E\{Y(\mathbf{u}_\beta) \cdot Y(\mathbf{u}_\alpha)\} \end{aligned}$$

$$
\begin{aligned}
&= \sum_{\beta=1}^{n} \lambda_\beta C(\mathbf{u}_\alpha - \mathbf{u}_\beta) \\
&= C(\mathbf{u} - \mathbf{u}_\alpha)
\end{aligned}
$$

The covariance is correct! Note that the last substitution comes from the kriging equations shown above. The kriging equations force the covariance between the data values and the kriging estimate to be correct; however, the variance is too small, and the covariance between the kriged estimates themselves is incorrect. We can imagine fixing the covariance between the kriged estimates by proceeding sequentially, that is, to use previously predicted values in subsequent predictions.

Although the covariance between the kriged estimates and the data is correct, the variance is too small. The variance of the stationary random function should be $\sigma^2 = C(0)$. This stationary model variance should be constant everywhere:

$$\sigma^2(\mathbf{u}) = \sigma^2, \ \forall \ \mathbf{u} \in A$$

The smoothing effect of kriging amounts to reduce this variance, particularly at locations far away from data values. Another interesting property of kriging is that the variance of the kriged estimate is known:

$$Var\{Y^*(\mathbf{u})\} = C(0) - \sigma^2_{SK}(\mathbf{u})$$

This tells us how much variance is *missing*: the kriging variance $\sigma^2_{SK}(\mathbf{u})$. This missing variance must be added back in without changing the covariance reproduction properties of kriging.

An independent component with zero mean and the correct variance can be added to the kriged estimate:

$$Y_s(\mathbf{u}) = Y^*(\mathbf{u}) + R(\mathbf{u})$$

The covariance between the simulated value $Y_s(\mathbf{u})$ and one of the data values used in the estimation may be calculated:

$$
\begin{aligned}
Cov\{Y_s(\mathbf{u}), Y(\mathbf{u}_\alpha)\} &= E\{Y_s(\mathbf{u}), Y(\mathbf{u}_\alpha)\} \\
&= E\left\{\left[\sum_{\beta=1}^{n} \lambda_\beta \cdot Y(\mathbf{u}_\beta) + R(\mathbf{u})\right] \cdot Y(\mathbf{u}_\alpha)\right\} \\
&= \sum_{\beta=1}^{n} \lambda_\beta \cdot E\{Y(\mathbf{u}_\beta) \cdot Y(\mathbf{u}_\alpha)\} + E\{R(\mathbf{u}) \cdot Y(\mathbf{u}_\alpha)\}
\end{aligned}
$$

note that $E\{R(\mathbf{u}) \cdot Y(\mathbf{u}_\alpha)\} = E\{R(\mathbf{u})\} \cdot E\{Y(\mathbf{u}_\alpha)\}$, since $R(\mathbf{u})$ is independent of any data value. The expected value of the residual $E\{R(\mathbf{u})\} = 0$, therefore, $E\{R(\mathbf{u}) \cdot Y(\mathbf{u}_\alpha)\} = 0$.

Thus the covariance between the simulated value and all data values is correct, that is, $Cov\{Y_s(\mathbf{u}), Y(\mathbf{u}_\alpha)\} = Cov\{Y^*(\mathbf{u}), Y(\mathbf{u}_\alpha)\} = C(\mathbf{u}), \mathbf{u}_\alpha\}$.

In light of these two features (1) the covariance reproduction of kriging, and (2) unchanged covariance by adding an independent residual, the SGS algorithm is as follows:

- Transform the original Z data to a standard normal distribution (all work will be done in "normal" space). We will see later why this is commonly performed.

- Go to a location $\mathbf{u}$ and perform kriging to obtain kriged estimate and the corresponding kriging variance:

$$Y^*(\mathbf{u}) = \sum_{\beta=1}^{n} \lambda_\beta \cdot Y(\mathbf{u}_\beta)$$

$$\sigma^2_{SK}(\mathbf{u}) = C(0) - \sum_{\alpha=1}^{n} \lambda_\alpha C(\mathbf{u}, \mathbf{u}_\alpha)$$

- Draw a random residual $R(\mathbf{u})$ that follows a normal distribution with mean of 0.0 and variance of $\sigma^2_{SK}(\mathbf{u})$.

- Add the kriged estimate and residual to get the simulated value:

$$Y_s(\mathbf{u}) = Y^*(\mathbf{u}) + R(\mathbf{u})$$

  Note that $Y_s(\mathbf{u})$ could be equivalently obtained by drawing from a normal distribution with mean $Y^*(\mathbf{u})$ and variance $\sigma^2_{SK}(\mathbf{u})$. The independent residual $R(\mathbf{u})$ is drawn with classical Monte Carlo simulation. Any reputable pseudo-random number generator may be used. It is good practice to retain the "seed" to the random number generator to permit reconstruction of the model if required.

- Add $Y_s(\mathbf{u})$ to the set of data to ensure that the covariance with this value and all future predictions is correct. As stated above, this is the key idea of sequential simulation, that is, to consider previously simulated values as data so that we reproduce the covariance between *all* of the simulated values.

- Visit all locations in a random order. There is no theoretical requirement for a random order or path; however, practice has shown [137] that a regular path can induce artifacts.

- Back transform all data values and simulated values when model is populated.

- Create any number of realizations by repeating with different random number seeds. A different seed leads to a different sequence of random numbers and, as a consequence, a different random path and different residuals for each simulated node. Each realization is "equally likely to be drawn" and often called equiprobable.

Why Gaussian?

The key mathematical properties that make sequential Gaussian simulation work are not limited to the Gaussian distribution. The covariance reproduction of kriging holds regardless of the data distribution; the correction of variance by adding a random residual works regardless of shape of the residual distribution (the mean must be zero and the variance equal to the kriging variance).

There is one very good reason why the Gaussian distribution is used: the use of any other distribution does not lead to a correct global distribution of simulated values. The mean may be correct, the variance is correct, the variogram of the values taken all together is correct, but the "shape" will not be. This is not a problem in "Gaussian space" since all distributions are Gaussian [13] including that of the final simulated values.

There is a second, less important, reason why the Gaussian distribution is used: the central limit theorem tells us that the sequential addition of random residuals to obtain simulated values leads to a Gaussian distribution.

The Cost of Gaussianity

The cost of the mathematical simplicity of Gaussian techniques is the character of maximum entropy [154], which implies maximum spatial disorder beyond the imposed variogram correlation. One consequence is that there is no potential to simulate connected paths of extreme values. This will have an affect on the results of flow simulation, but it is impossible to know how significant this will be without performing flow simulation.

The consequence of maximum entropy is that an SGS realization has less spatial structure than the real reservoir. This may be important for variables like permeability since the continuity of high and low permeability values has a significant impact on fluid flow predictions. Other simulation methods such as indicator geostatistics or simulated annealing could be considered when additional knowledge of the spatial structure can be quantified.

Notwithstanding the consequences of Gaussianity, the sequential Gaussian simulation algorithm is widely used. Figure 5.2 illustrates two SGS realizations of reservoir thickness. The spatial correlation matches the input variogram; high and low values are grouped together.

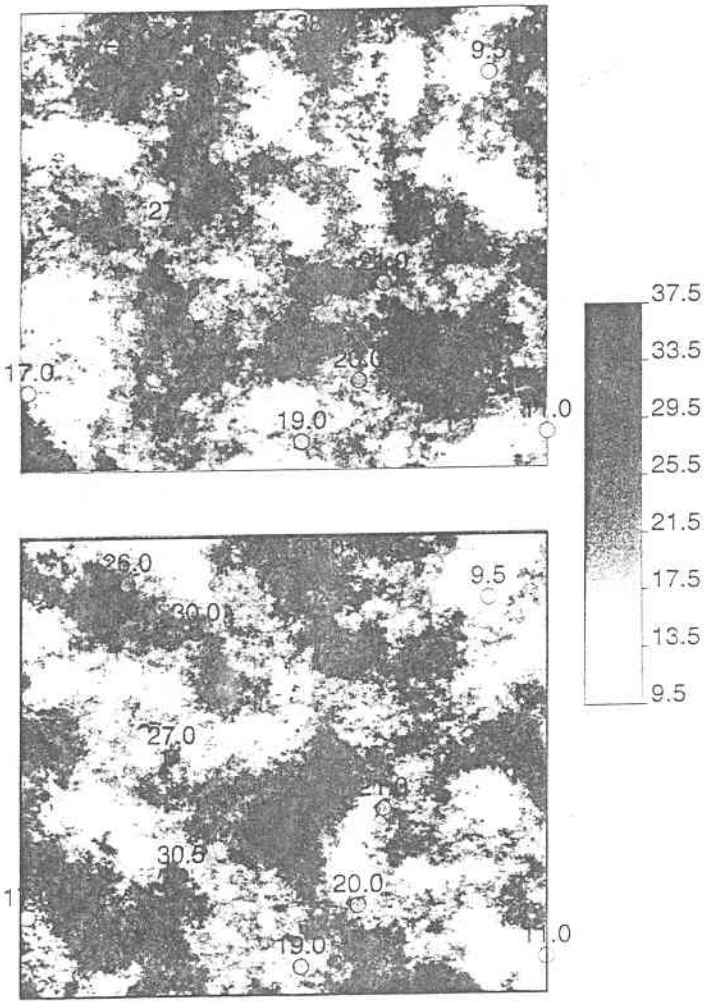


Figure 5.2: Two SGS realizations created for illustration purposes. Note how the data are reproduced in both cases. There is significant uncertainty in the realizations since the range of correlation is less than the average data spacing.

5.3 Direct Sequential Simulation

The theoretical foundation for sequential Gaussian simulation (SGS) was outlined in the previous section. The procedure works remarkably well. The data do not have to have a Gaussian histogram, because we can transform it to be Gaussian. Back transformation ensures that the simulated values are in real units.

The histogram of any particular SGS realization does not match the input declustered histogram exactly. The back transformation in SGS would only impose the histogram exactly if the Gaussian or normal values were exactly normal with a mean of 0, variance of 1, and correct shape. There are statistical or ergodic fluctuations between realizations. These variations are an important part of uncertainty; we will expect variability in the sample statistics over any study area of finite size. It is wrong to transform the results of SGS to impose the histogram exactly. Chapter 10 discusses how to handle the variability between multiple realizations. Let's return to the Gaussian transformation.

Gaussianity (of the input data and residual distribution shape) is required in SGS so that the simulated values have the correct global histogram. Notwithstanding this key advantage of the Gaussian distribution, we are tempted to consider non-Gaussian distributions in the sequential path. The spatial features of the resulting realizations could be altered by considering some non-Gaussian distribution shape in sequential simulation; this is called *direct sequential simulation (DSS)*. DSS suffers from two unavoidable consequences that undermine its tempting flexibility (1) the resulting histogram is not controlled, and (2) the multivariate spatial features are essentially similar to SGS.

The histogram of a DSS realization is some hybrid between the original data histogram, the residual histograms chosen, and a Gaussian distribution. The original data always play a role and become more important as there are more data within the range of correlation. The residual histogram influences the final simulated values. The kriged mean and variance are at the heart of DSS. The kriged mean is a weighted average of the data and previously simulated values. The ubiquitous central limit theorem tells us that these average values tend to a Gaussian or normal distribution. Thus, the histogram of a DSS realization does *not* honor that of the original data.[4] Moreover, it does not honor the specified residual histogram or the Gaussian histogram.

Some form of transformation (see Section 2.4) must be used to convert the DSS simulated values back to the right units. This back transformation introduces a significant limitation: the variability or uncertainty due to ergodic fluctuations is removed. Uncertainty in the resulting histogram is significant because it is an important first-order aspect of uncertainty.

[4]SGS honors the input standard Gaussian Y histogram and, therefore, the input Z histogram after back transformation

At present, DSS is primarily of research interest. The use of DSS may become important in the future as these limitations are addressed. The indicator formalism, however, is an alternative to Gaussian simulation that has seen wide use over the last 25 years.

5.4 Indicator Formalism

The key idea behind the indicator formalism is to code all of the data in a common format, that is, as *probability* values [3, 106, 144, 151, 153, 156, 240]. The two main advantages of this approach are (1) simplified data integration because of the common probability coding, and (2) greater flexibility to account for different continuity of extreme values. The comparative performance of indicator methods has been studied extensively by Goovaerts [107, 109, 108].

The indicator approach for continuous data variables requires significant additional effort versus Gaussian techniques. The indicator formalism applied to categorical data has seen wide application in facies modeling, particularly for carbonate reservoirs and high net-to-gross siliciclastic reservoirs. Regardless of the variable type, the indicator approach leads directly to a distribution (histogram) that is a model of uncertainty at each unsampled location estimated. Nevertheless, the implementation details for continuous data are quite different from categorical variables; the two approaches are discussed separately.

Continuous Data

The aim of the indicator formalism for continuous variables is to directly estimate the distribution of uncertainty $F_Z(\mathbf{u})$ at unsampled location $\mathbf{u}$. The cumulative distribution function is estimated at a series of threshold values: $z_k, k = 1, \ldots, K$. As an example, Figure 5.3 shows $K = 5$ probability values at five threshold values that provide a distribution of uncertainty. The probability values are evaluated by first coding the data as indicator or probability values. The indicator coding at location $\mathbf{u}_\alpha$ is written:

$$\begin{aligned} i(\mathbf{u}_\alpha; z_k) &= Prob\{Z(\mathbf{u}_\alpha) \leq z_k\} \\ &= \begin{cases} 1, \text{ if } z(\mathbf{u}_\alpha) \leq z_k \\ 0, \text{ otherwise} \end{cases} \end{aligned} \tag{5.18}$$

The expected value of this indicator variable is the stationary prior probability to be less than the threshold, that is, $F_Z(z_k)$. As in Equation 5.3, above, we consider residual data:

$$Y(\mathbf{u}_\alpha; z_k) = i(\mathbf{u}_\alpha; z_k) - F(z_k), \;\; \alpha = 1, \ldots, n, \; k = 1, \ldots, K$$

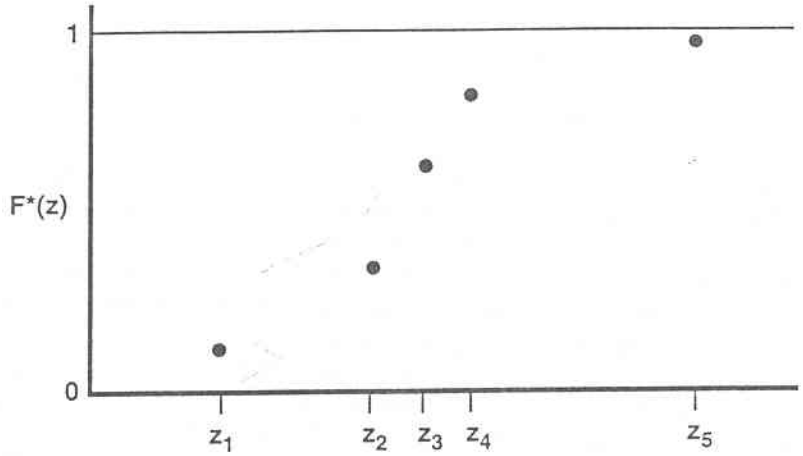


Figure 5.3: Schematic illustration of probability distribution $F(z)$ at a series of five threshold values, $z_k, k = 1, \ldots, 5$.

As in Equation 5.7, a linear estimate of this residual is considered:

$$Y^*(\mathbf{u}; z_k) = \sum_{\alpha=1}^{n} \lambda_\alpha(z_k) \cdot Y(\mathbf{u}_\alpha; z_k) \tag{5.19}$$

where $Y^*(\mathbf{u}; z_k)$ is the estimated residual from the prior probability at threshold z_k, $F(\mathbf{u}; z_k)$, $\lambda_\alpha(z_k), \alpha = 1, \ldots, n$ are the weights, and $Y(\mathbf{u}_\alpha; z_k), \alpha = 1, \ldots, n$ are the residual data values.

The indicator kriging derived cumulative distribution function at an unsampled location at threshold z_k is calculated as:

$$F_{IK}(\mathbf{u}; z_k) = \sum_{\alpha=1}^{n} \lambda_\alpha(z_k) \left[i(\mathbf{u}_\alpha; z_k) - F(z_k) \right] + F(z_k) \tag{5.20}$$

In presence of no data ($n = 0$) the indicator kriging (IK) estimate of the distribution of uncertainty is simply the prior mean at that threshold, that is, $F(z_k)$.

The IK process is repeated for all K threshold values $z_k, k = 1, \ldots, K$, which discretize the interval of variability of the continuous attribute Z. The distribution of uncertainty, built from assembling the K indicator kriging estimates, represents a probabilistic model for the uncertainty about the unsampled value $z(\mathbf{u})$. This indicator kriging procedure requires a variogram measure of correlation corresponding to each threshold $z_k, k = 1, \ldots, K$ so that the weights $\lambda_\alpha(z_k), \alpha = 1, \ldots, n; k = 1, \ldots, K$ can be determined.

The correct selection of the threshold values z_k for indicator kriging is essential: too many threshold values and the inference and computation becomes needlessly tedious and expensive; too few, and the details of the distribution are lost. Between 5 to 11 thresholds are normally chosen. The

thresholds are rarely above the 0.9 quantile or below the 0.1 quantile since it becomes difficult to reliably infer the corresponding indicator variograms. The thresholds are often chosen to be equally spaced quantiles, for example, the nine deciles are often chosen. Implementation details will be covered in more detail in Chapter 8.

The indicator residual data in the IK expression (5.19) originate from hard data $z(\mathbf{u}_\alpha)$ that are deemed perfectly known; thus, the indicator data $i(\mathbf{u}_\alpha; z)$ are hard in the sense that they are valued either 0 or 1 and are available at any cutoff value z. There are many applications where some of the z-information is due to soft secondary data such as seismic data. This soft data may enter indicator kriging in a variety of ways.

The first approach for soft data is to use them for locally variable prior probability values $F(\mathbf{u}; z_k)$. The soft data would be calibrated to hard data by considering co-located hard and soft data. The hard data are cross plotted against the soft data and "conditional" distributions of the original data variable for classes of the soft data variable are extracted. The $F(\mathbf{u}; z_k)$ values are taken from the conditional distributions considering the soft data variable at each location $\mathbf{u}$. This is illustrated by looking ahead to the bottom of Figure 5.5.

A second approach is to consider the $F(\mathbf{u}; z_k)$ values as a covariate in a cokriging context rather than as a prior mean. Any licit model of coregionalization can be used including the linear model of coregionalization or a collocated model. The Markov-Bayes model [279] is a model of coregionalization specially formulated for soft and hard indicator data.

A third approach for soft data is to consider them as inequality constraints or soft indicator data. At location $\mathbf{u}_\alpha$ inequality data tell us that z-value is between a lower limit $z_{low}(\mathbf{u}_\alpha)$ and upper limit $z_{upper}(\mathbf{u}_\alpha)$, that is,

$$i_{inequality}(\mathbf{u}_\alpha; z_k) = \begin{cases} 0, \text{ if } z_k < z_{low}(\mathbf{u}_\alpha) \\ \text{missing or undefined, if } z_{low}(\mathbf{u}_\alpha) \geq z_k \leq z_{upper}(\mathbf{u}_\alpha) \\ 1, \text{ if } z_k > z_{upper}(\mathbf{u}_\alpha) \end{cases} \tag{5.21}$$

The inequality constraint data can be used for thresholds that are defined; otherwise, there is no data for thresholds that are "missing or undefined." Figure 5.5 illustrates the indicator representation of hard data, inequality constraints, and soft data.

Indicator Variograms

One of the primary advantages of the indicator formalism is the ability to specify different spatial continuity at each threshold. This amounts to specifying problem-specific continuity for the low, median, and high values. Figure 5.4 presents indicator variograms at seven thresholds. These variograms were calculated from a 2-D exhaustive image of laminated sandstone. Note the changing zonal anisotropy and nugget effect for the different thresholds. The calculated variogram values and the models should be standardized to

unit variance by dividing by $p(1-p)$ where $p = F(z_k)$ is the global mean of the indicator variable or proportion of 1s. The parameters (relative nugget effect, ranges, and so on) of the indicator variograms are related since they arise from a common continuous-variable origin.

Order Relations

The series of indicator-derived probability values $F_{IK}(\mathbf{u}; z_k), k = 1, \ldots, K$ must meet the requirements of a cumulative distribution function, that is, a non- decreasing function between zero and one. This is not guaranteed by indicator kriging; therefore, an a-posteriori correction is applied. The corrections required are usually small. Large deviations imply a problem with the data or inconsistency with the indicator variograms. The correction method documented in GSLIB [62] is simple and effective.

Categorical Data

The aim of the indicator formalism for categorical variables is to directly estimate the distribution of uncertainty in the categorical facies variable. The probability distribution consists of estimated probabilities for each category: $p^*(k), k = 1, \ldots, K$. As an example, Figure 5.6 shows $K = 5$ probability values for five facies. The probability values are estimated by first coding the data as indicator or probability values:

$$\begin{aligned} i(\mathbf{u}_\alpha; z_k) &= Prob\{\text{facies } k \text{ being present}\} \\ &= \begin{cases} 1, \text{ if facies } k \text{ is present at } \mathbf{u}_\alpha \\ 0, \text{ otherwise} \end{cases} \end{aligned} \tag{5.22}$$

The expected value of this indicator variable is the stationary prior probability of facies k, that is, $p(k)$. As above, we consider residual data:

$$Y(\mathbf{u}_\alpha; z_k) = i(\mathbf{u}_\alpha; k) - p(k), \quad \alpha = 1, \ldots, n, \; k = 1, \ldots, K$$

Kriging of these residual data is used to derive the probability of each facies $k = 1, \ldots, K$ at an unsampled location. Once again, a variogram measure of correlation is required for each facies $k = 1, \ldots, K$. The result of indicator kriging is a model of uncertainty at location $\mathbf{u}$:

$$p_{IK}(\mathbf{u}; k) = \sum_{\alpha=1}^{n} \lambda_\alpha(k) \cdot [i(\mathbf{u}_\alpha; k) - p(k)] + p(k), \quad k = 1, \ldots, K \tag{5.23}$$

Seismic and other secondary data sources may be coded as soft probabilities for cokriging; details of this are covered later in Chapter 8.

Order Relations

The estimated probabilities $p_{IK}(k), k = 1, \ldots, K$ must meet the requirements of a probability distribution, that is, being non-negative and sum to one. As

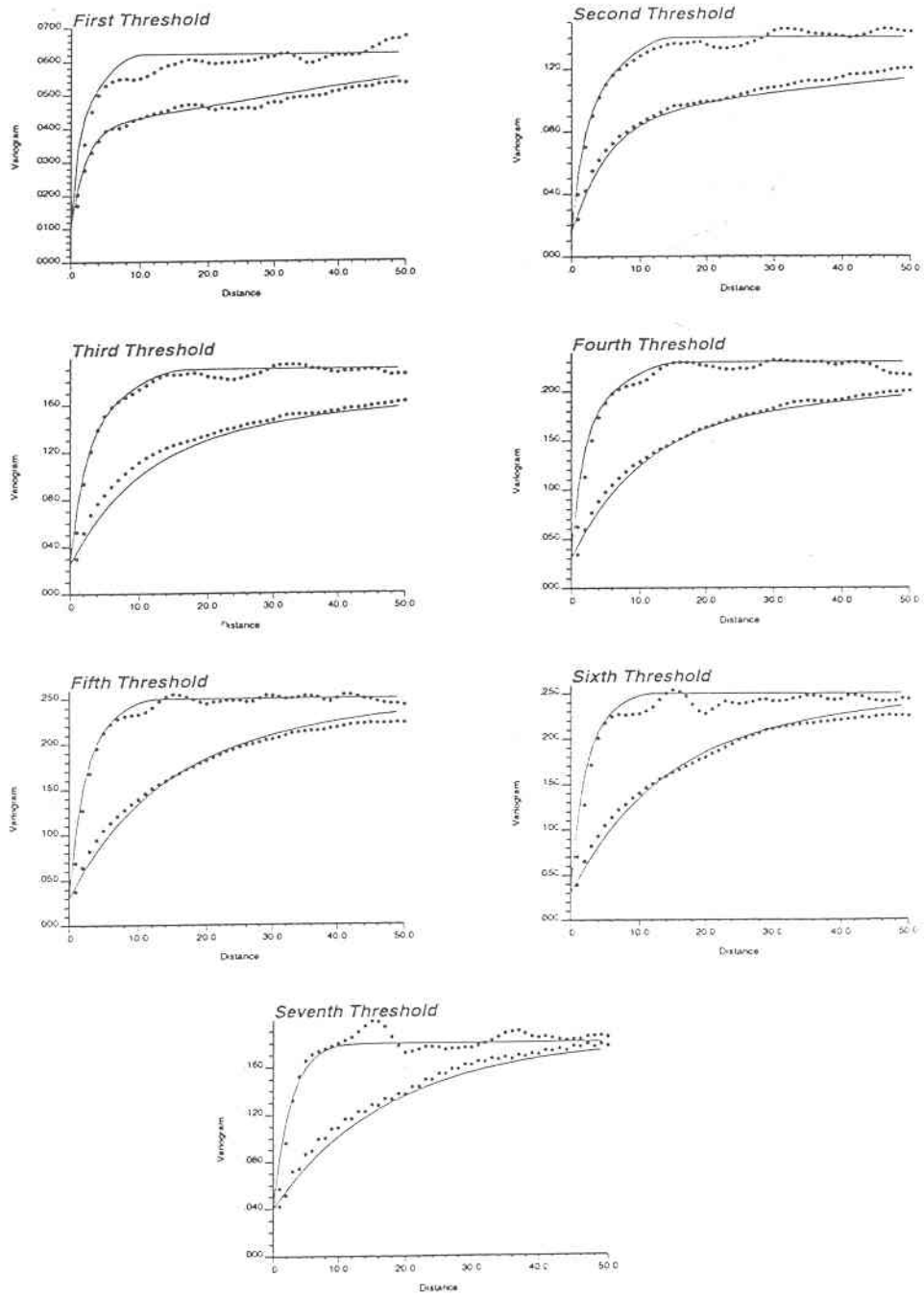


Figure 5.4: Indicator variograms for seven thresholds. Note the changing zonal anisotropy and nugget effect for the different thresholds. Also note how the variogram is different at low and high thresholds.

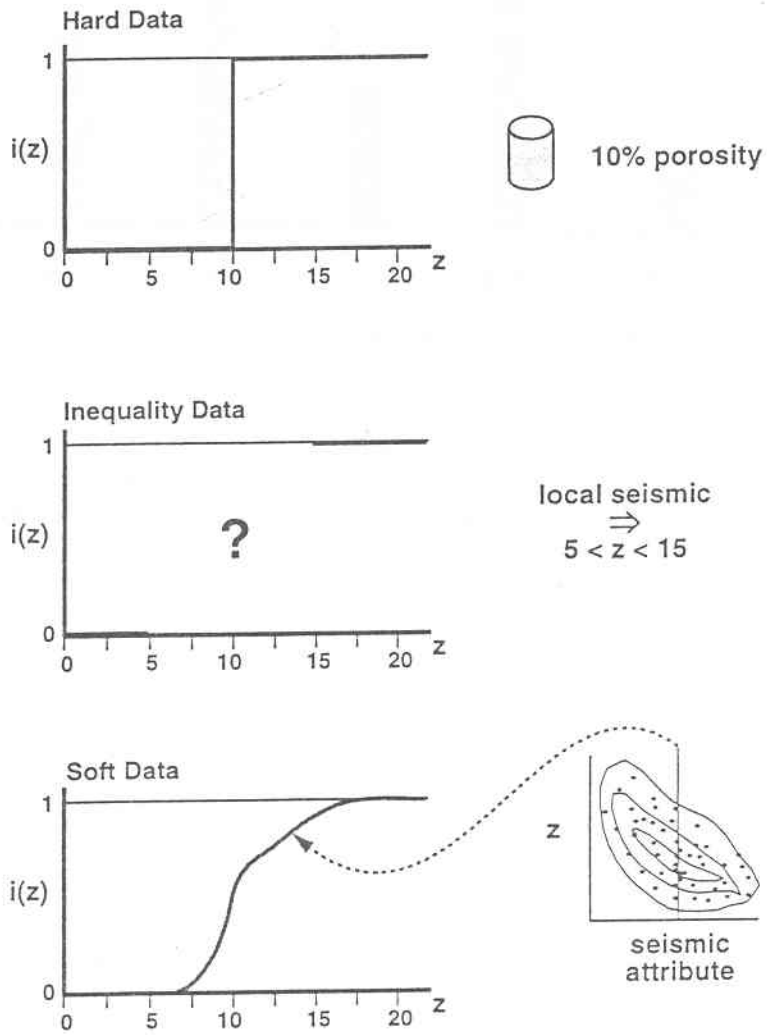


Figure 5.5: Schematic illustration of (top) hard data where the indicator transform i is zero or one, (middle) inequality data where the indicator data are zero below the lower limit and one above the upper limit, and (bottom) a soft data derived distribution.

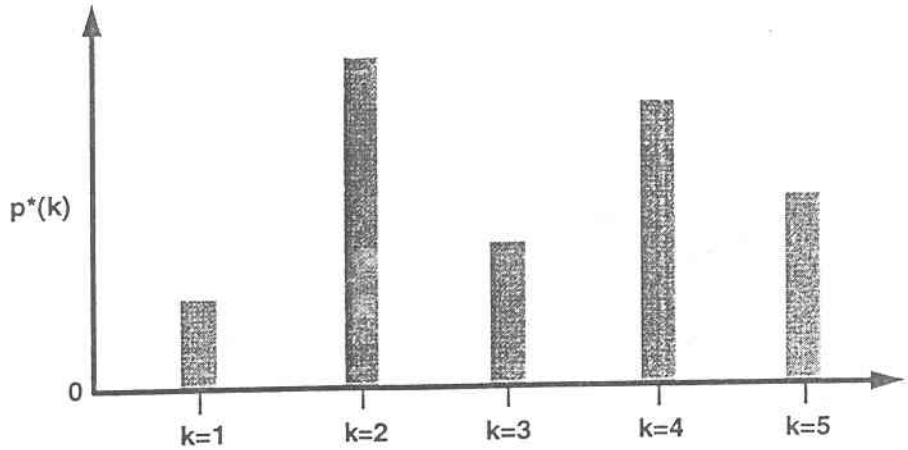


Figure 5.6: Schematic illustration of probability distribution $p^*(k), k = 1, \ldots, 5$ for five facies types, $k, k = 1, \ldots, 5$.

with continuous variables, these order relation requirements are not guaranteed by indicator kriging; therefore, an a-posteriori correction is applied. The probabilities are set to zero if they are negative and then reset according to:

$$p^*(k) = \frac{p_{IK}(k)}{\sum_{k=1}^{K} p_{IK}(k)}, \quad k = 1, \ldots, K$$

which ensures they sum to one. Large deviations from licit probabilities imply a problem with the data or the variograms.

Sequential Indicator Simulation

The concept of sequential simulation, described in context of Gaussian (SGS) and direct (DSS) distributions, can be extended to the indicator-based model of uncertainty. The grid nodes are visited sequentially in a random path. At each grid node:

1. Search for nearby data and previously simulated values.
2. Perform indicator kriging to build a distribution of uncertainty (correcting order relation problems if necessary).
3. Draw a simulated value from the distribution of uncertainty.

Multiple realizations are then generated by repeating the entire procedure with a different random number seed. Implementation details are included in Chapters 6 and 8.

5.5 P-Field Methods

Sequential simulation proceeds by (1) constructing a conditional distribution model that accounts for all original data and previously simulated grid nodes and (2) drawing with Monte Carlo simulation from the conditional distributions. Multiple realizations are generated by repeating the procedure. The CPU cost of such methods is mostly due to the cost of constructing each local conditional distribution over again for each new realization.

In sequential simulation, correlation between simulated values is achieved in the construction of the conditional distributions. R. M. Srivastava proposed a different approach [93, 230]. The two key aspects of his proposal were (1) to construct the conditional distributions using *only* the original data so that it can be done only once instead of repeatedly for each realization, and (2) draw from the conditional distributions using correlated probability values instead of the random numbers used in traditional Monte Carlo simulation.

- Spatial correlation between the probability values imparts spatial correlation in the resulting simulated values.
- The original data are honored since the conditional distributions at data locations have zero variance and the original data values are retrieved regardless of the simulated probability value.

There is a significant CPU advantage to p-field simulation when multiple realizations must be constructed.

Figures 5.7 and 5.8 presents a 1-D schematic illustration of the procedure. The construction of the conditional distributions of uncertainty at each location, using the original data alone, is the first step, see Figure 5.7. Three correlated probability fields and the corresponding simulated values are shown on Figure 5.8.

The p-field simulation methodology requires unconditional realizations of probability values. There are a number of numerical methods that are suited to fast simulation of uncorrelated spatial fields such as FFT, fractal, moving average, and turning bands methods [36, 114, 128, 129, 130]. These will not be recalled here; however, we note that the SGS method is an inefficient way to generate unconditional probability fields for p-field simulation because it is sequential.

In spite of the apparent advantages of the p-field simulation methodology it is not recommended. R. M. Srivastava and R. Froidevaux pointed out two artifacts that were explored more completely in the Math Geology paper of Pyrcz and Deutsch. The artifacts are (1) the conditioning data appear as local minima and maxima, and (2) the spatial continuity is not reproduced next to the conditioning data; the values are too smooth or too well correlated near the conditioning data.

The first artifact is caused by the separation of conditioning and the Monte Carlo simulation. The local distributions collapse to a step function at the local hard data locations. This causes the simulated values to "pinch" toward

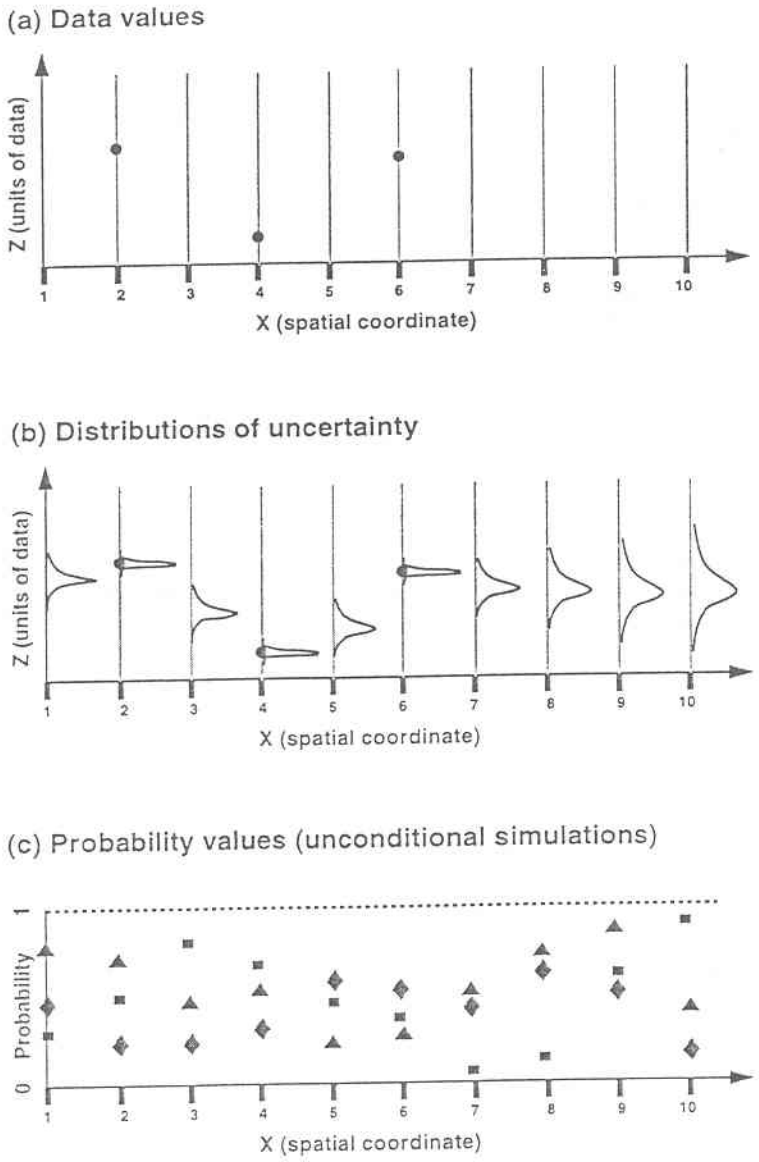


Figure 5.7: The first step in p-field simulation, that is, the construction of the conditional distribution of uncertainty at each location using the original data: (a) 1-D example conditioning data, (b) distributions of uncertainty, and (c) probability values from unconditional simulation. Note how the variance of the conditional distributions in (b) is lower near the data values.

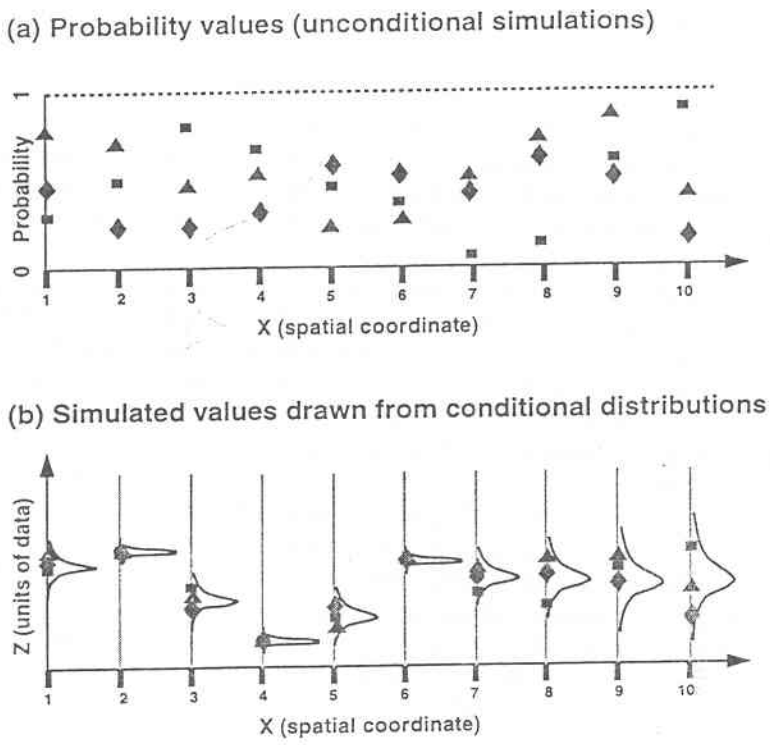


Figure 5.8: The second step in p-field simulation is the generation of correlated probability fields and the simultaneous drawing from the conditional distributions of uncertainty: (a) three sets of correlated probability values, and (b) distributions of simulated values drawn from the distributions of uncertainty shown in Figure 5.7.

the hard data values in the vicinity of the hard data. The artifact becomes more significant as the probabilities depart from 0.5 in the vicinity of the data location. Probability field values greater than 0.5 in the vicinity of the data cause a local minimum. Probability values less than 0.5 cause a local maximum. The only instance in which this artifact will not occur is when the non conditional correlated probability value is exactly 0.5 at the data location, which is unlikely.

The second artifact is that there is increased continuity next to conditioning data. Although expected from theory, this bias causes practical problems. P-field simulation, sometimes referred to as the cloud transform, should be avoided because of these two artifacts. It should be mentioned that these artifacts do not appear significant when dealing with categorical variables, where the method is equivalent to the truncated Gaussian method, see Chapter 6.

5.6 Accounting for Trends

Locally Varying Directions of Continuity

Reservoir facies and petrophysical properties may exhibit changing directions of continuity, that is, the principal direction of continuity may depend on location. There is no theoretical challenge to account for such locally varying directions of continuity; however, the problem (as usual) is the data to evaluate the local variations.

Figure 5.9 shows an example where the underlying detailed pixel plot shows locally varying directions of anisotropy. This data is the Walker Lake data from [138]. A number of authors have presented approaches to handle locally varying directions of continuity. The central idea is to modify the kriging equations locally. Deutsch and Lewis [64] setup the local kriging equations using the local direction of continuity. Xu [271] implemented a similar procedure in the context of mimicking the facies distribution in a sinuous fluvial channel setting. Soares implemented a procedure where only the right-hand side to the kriging equations were altered; the left-hand side, which accounts for data redundancy, used a global measure of anisotropy.

Programs to implement kriging and simulation with variable anisotropy for Gaussian and indicator sequential simulation are straightforward. The idea is to locally adjust the direction of anisotropy within each search neighborhood. The problem of calculating distances in presence of locally varying anisotropy and ensuring positive definiteness of the resulting kriging matrices is much simplified by retaining a single direction of anisotropy for each kriging. That direction of anisotropy is then changed from location to location.

Figure 5.10 shows an example of sequential indicator simulation with two local directions of anisotropy; the top half is at an azimuth angle of 60°and the bottom half is at an angle of 120°. Any complex grid of local directions could be used; however, the reproduction of the input direction grid depends on the kriging search radius and the conditioning data. Complex small scale features will not be well reproduced because of the key assumption that the anisotropy does not change within a kriging neighborhood.

The problem of calculating a reliable variogram in presence of locally varying directions of continuity is significant. Conventional software does not permit using such information. The common alternative is to calculate the variogram in directions or subareas where the anisotropy follows a nearly constant direction. The actual anisotropy will be more than observed because of mixing the directions.

Locally Varying Mean

A *trend* is defined as a general tendency or pattern that is a deterministic or predictable aspect of the spatial distribution of reservoir facies or properties. Different authors have different definitions.

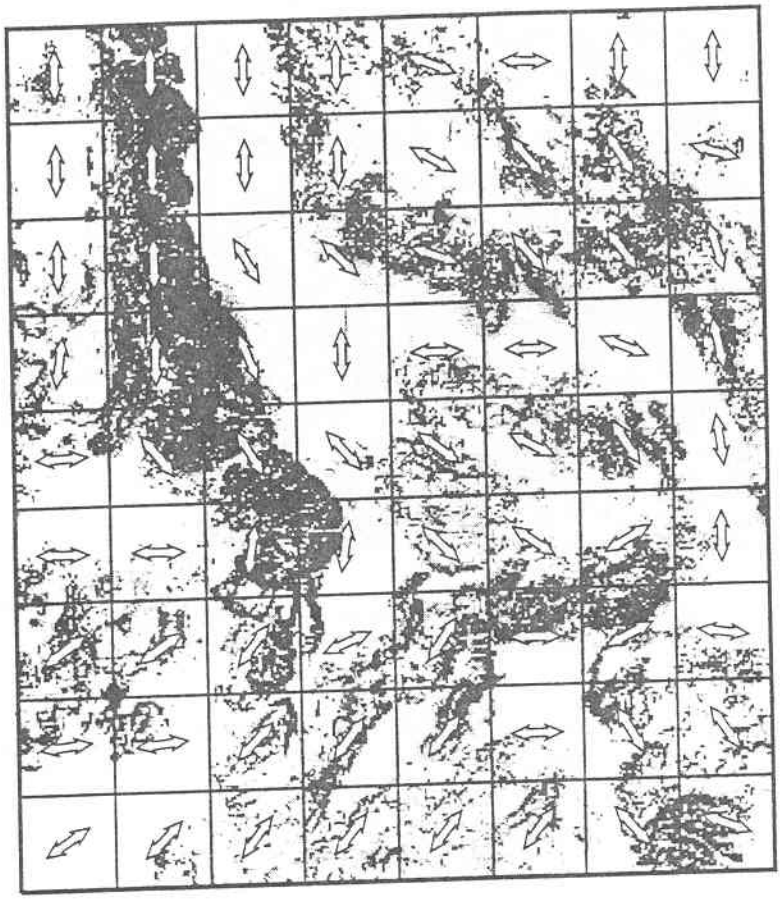


Figure 5.9: Specification of direction vectors within a large-scale grid network superimposed on the fine-scale distribution of properties. The fine-scale distribution is the Walker Lake data from [138].

Virtually all natural phenomena exhibit trends. For example, the vertical profile of porosity and permeability may "fine upward" within each successive strata simply because coarse particles settle first within a depositional event. Moreover, there may be other large-scale trends related to long term changes in the sediment source, space for deposition, or climate changes during deposition. The geological details of why there are systematic large-scale trends are less important than the fact that there are such trends and that they often affect reservoir performance predictions.

Any discussion on trends is subjective. All geostatisticians struggle with the separation of deterministic features, like trends, and stochastic correlated features, which are left to simulation. This separation is truly a decision of the modeler since reservoirs are deterministic; a "trend" and "residual" are meaningless if the reservoir is known exactly. Trends are necessarily

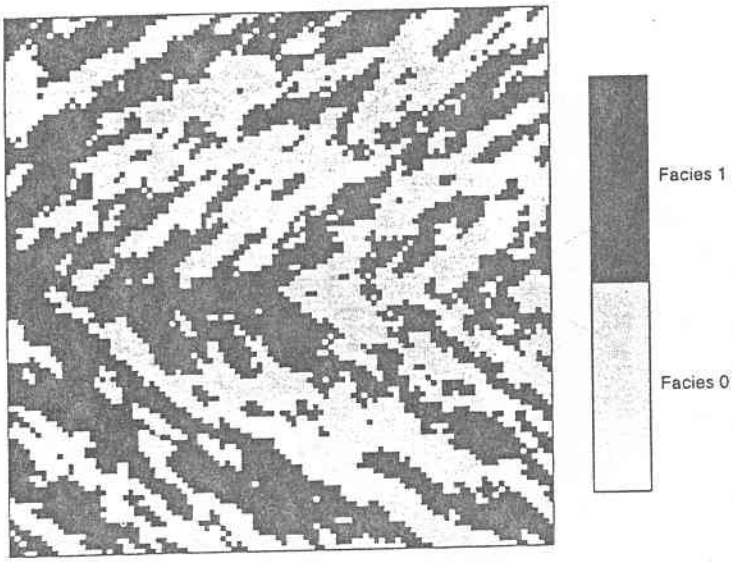


Figure 5.10: Example of sequential indicator simulation with two local directions of anisotropy; the top half is 90°different from the bottom half.

dependent on the available data. This is illustrated on Figure 5.11 where the trend appears different at (1) different scales, and (2) with different data. Most of the variation appears stochastic in presence of sparse data; as more data becomes available, the more refined the trend model and less of the intrinsic variation is left to geostatistical modeling.

Certain trends may be inferred from general geological knowledge. For example, the porosity within individual sedimentary beds in a deepwater depositional system is known to fine upwards. This is due to well-understood physical principles and does not depend on an abundance of data. In some situations the presence of a trend is observed through the data. There is an apparent contradiction with many well data. Abundant data permit the trend to be identified, but it is less important since the data impose their trend features as local conditioning data on the resulting model; therefore, explicitly modeling a trend may be considered less important. In all realistic reservoir scenarios (versus some mining or environmental studies) there are rarely enough well data to enforce a trend throughout the study area. Trends are only reproduced in the immediate vicinity of the wells. Some variant of the following procedures must be considered in the realistic case of widely spaced well data.

An implicit assumption of all geostatistical algorithms such as kriging is that the variable being considered is "stationary." This assumption entails that spatial statistics do not depend on location, for example, the mean and variogram are constant and relevant to the entire study area. Special efforts

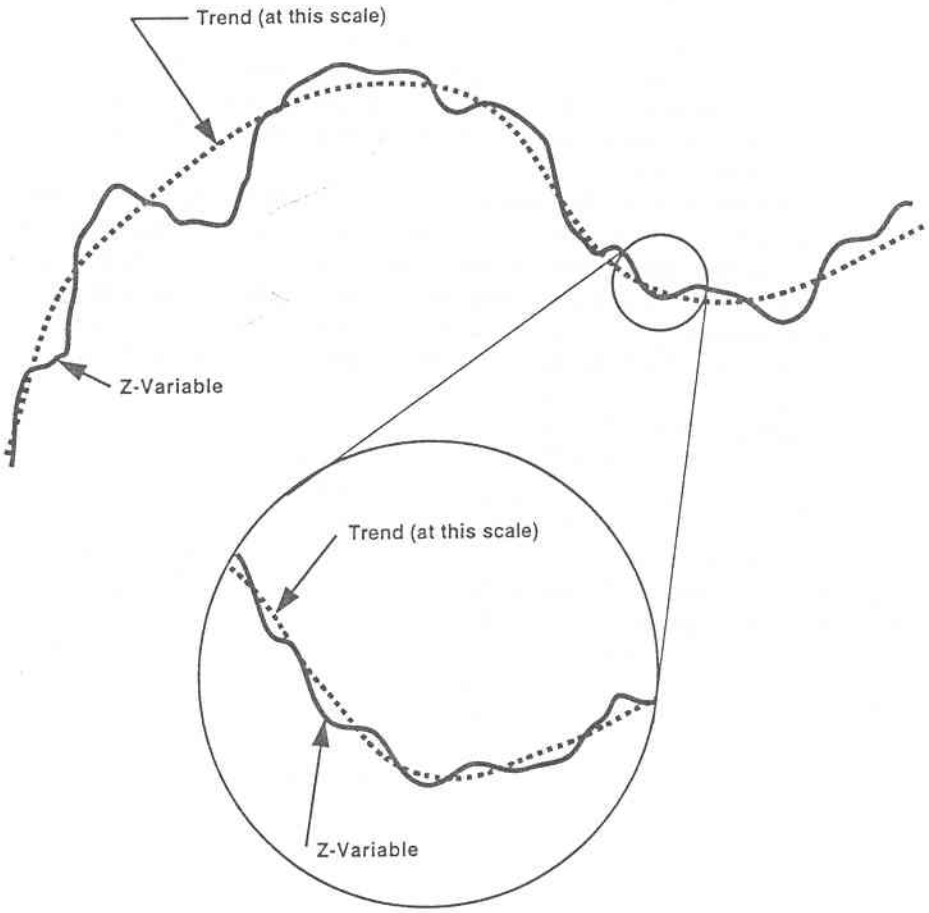


Figure 5.11: Schematic illustration of how trends appear different at different scales. A decision must be made about how much of the variability to put in the trend and how much to consider stochastic variations.

must be made to account for trends or so-called non-stationarities when they are known to exist from our geological understanding of the data.

Constraints have been added to kriging in an attempt to capture trends. Kriging with a trend model (KT) permits a trend of polynomial shape to be fitted within each local kriging neighborhood. Kriging with an external drift permits an independent variable to define the shape of the local trend surface. Experience has shown that these forms of kriging work when there are adequate data within each local neighborhood to define the trend. For practical reservoir modeling, where data are often sparse, these methods do not satisfactorily account for the trend.

Building a Deterministic Trend Model

Automatic fitting of a trend within the kriging formalism only works in presence of many data. A far better approach to define a trend for reservoir modeling purposes is to model it by hand contouring, guided machine contouring, or fitting simple polynomial models. The linear trend fitted in Figure 4.12 in Chapter 4 is an example.

In practice, we consider the vertical trend first and then map areal trends. Trend analysis is preceded by determination of appropriate stratigraphic coordinates. The first step in vertical trend analysis is visual inspection of the well data. Most trends will be known geologically before any geostatistical modeling is considered. A vertical trend plot for continuous variables or vertical proportion curve for facies should be constructed to look for and quantify vertical trends. Figure 5.12 shows an example of each.

There are a number of comments to be made regarding the continuous trend curve on the left side of Figure 5.12. A vertical trend does not need to be considered in modeling when all points fall close to the same average or there is considerable overlap in the simplified "box plots." A trend should be considered when there are distinct features on this plot. These points can be used directly for the trend, that is, the trend or mean is considered constant within each stratigraphic interval equal to the average of the data in the interval. It takes more time to fit a polynomial model and check that the fit is reasonable; however, it may be worthwhile to avoid artifact discontinuities at the boundaries of the chosen intervals.

Similar comments can be made about the facies proportion curve on the right side of Figure 5.12. Different techniques are used to account for trends in continuous and categorical variables, see below.

It is worth emphasizing the importance of trend modeling one more time. Consider the sketch in Figure 5.13. If the wells are relatively far apart, that is, distances $d1$and $d2$ are large with respect to the horizontal range of correlation; then, locations A, B, and C will have the same average over many realizations. This would be wrong; the average at A should be more than C even when there are no nearby well data. Fitting the vertical trend and working with residuals from enforce this trend.

Areal trends are usually modeled separately from vertical trends. It is hard to visualize and fit a 3-D trend model all at once; it is just easier to fit the vertical trends and then areal trends. Of course, the areal and vertical trends must be reconciled into a consistent 3-D trend model prior to any geostatistical modeling, see below.

For facies indicator variables, the proportion of each facies is calculated over the vertical extent of the stratigraphic layer. For continuous variables, the average of the variable (within each facies) is calculated over the vertical extent of the stratigraphic layer. These values are displayed on a 2-D map and contoured by hand or with the computer. Kriging could be used. Figure 5.14 gives an illustration of two maps. The trend on the left should be used. The trend on the right should not. We look for trends or variations at scales

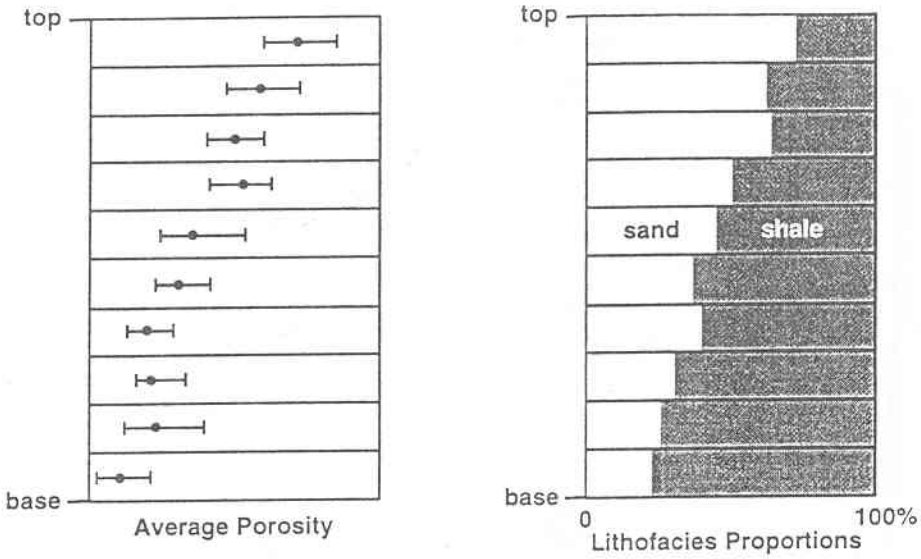


Figure 5.12: Example of vertical trends within a stratigraphic layer. The layer is partitioned into equal stratigraphic intervals (10 in this case), the average of the continuous property is calculated or the proportion of different facies. The black dots are the average of all porosity measurements falling in that stratigraphic interval. The attached lines are the 80% (or some other arbitrary) probability interval of the values falling in the interval.

larger than the data spacing. It would be nice to have an unambiguous way of deciding whether or not to model a trend; however, that is not the case. Do *not* model the trend when in doubt.

The areal trend is modeled by a gridded map at the same grid spacing as the geostatistical facies and property models to be built later. The areal trend map $m(x, y)$ and vertical trend curve $m(z)$ must now be merged to a 3-D trend model.

There is no unique way to merge a trend map with a trend curve. Ideally, additional geological knowledge should be brought to bear to determine the best approach. A simple and practical approach is to scale the vertical curve to the correct areal average:

$$m(x, y, z) = m(z) \cdot \frac{m(x, y)}{\overline{m}} \tag{5.24}$$

where $m(x, y, z)$ is the needed 3-D trend values, $m(z)$ is the vertical trend for the z-location, $m(x, y)$ is the areal trend value for this x, y location, and $\overline{m}$ is the global average of the variable being considered. This approach amounts to scale the vertical trend curve to the correct areal average. The resulting curves for high and low values of $m(x, y)$ should be checked for non-physical

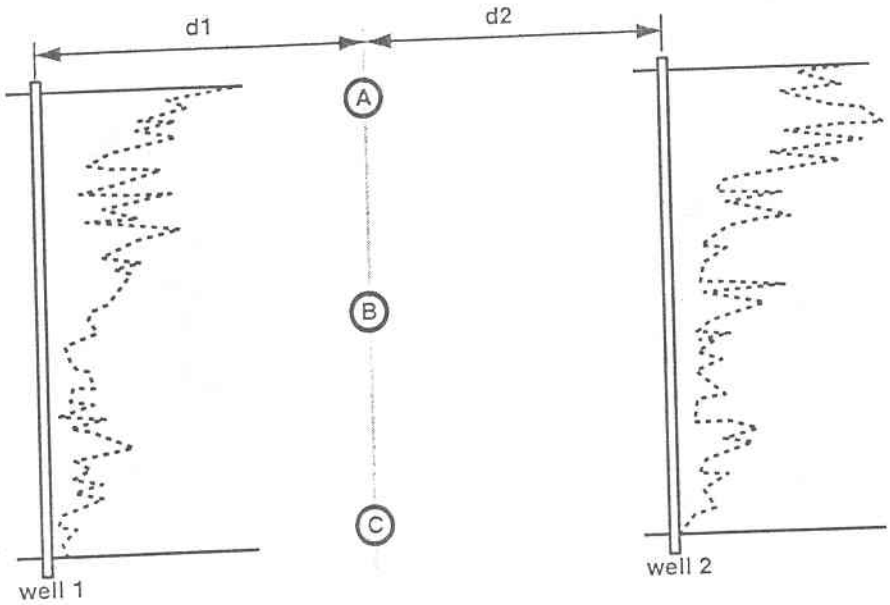


Figure 5.13: Illustration of two wells displaying an obvious vertical trend and a location between them. The expected average or mean values at the three locations A, B, and C should be different; however, they will be the same if $d1$ and $d2$ are beyond the range of correlation from the wells and the trend is not modeled explicitly.

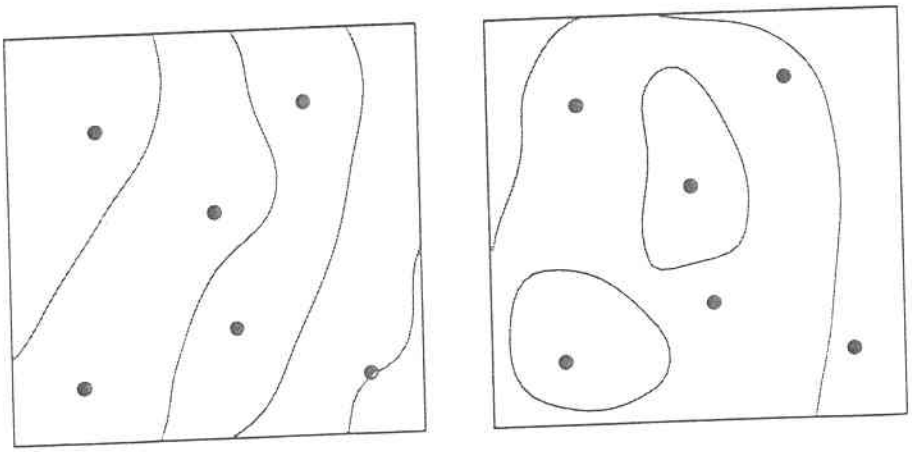


Figure 5.14: Illustration of two maps for areal trends. The example on the left shows a trend at a scale much larger than the data spacing. The example on the right is not so clear. We may choose not to model areal trends in this case (right side).

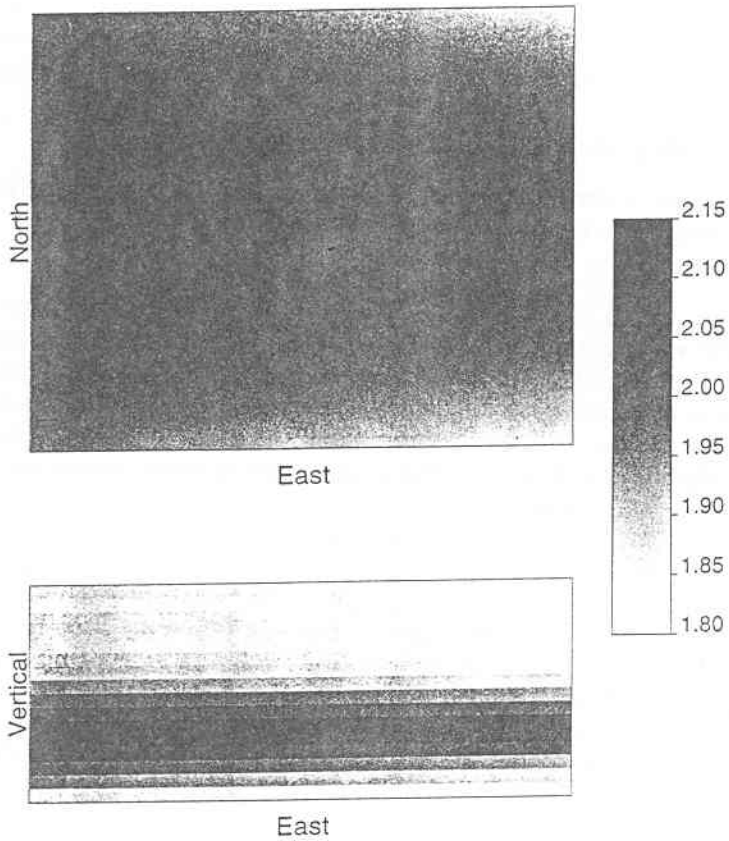


Figure 5.15: Horizontal and vertical cross section through a 3-D trend cube constructed for the log permeability within a reservoir layer (stratigraphic vertical coordinate). Note the high-permeability layer and the decrease in permeability to the east of the reservoir.

results. One possible problem is that high values are too high and low values are too low. Upper and lower limits to $m(x, y, z)$ can be imposed.

An alternative to equation (5.24) is to average the trend values, that is, $m(x, y, z) = (m(z) + m(x, y))/2.$, but this often smoothes the trends so much that neither can be reproduced in subsequent modeling.

Figure 5.15 illustrates a trend model of the logarithm of permeability (from a real reservoir study). The vertical trend is very pronounced. Figure 5.15 shows two cross sections through a 3-D trend model for the entire reservoir layer.

These procedures lead to a field-wide deterministic trend model for all x, y, z locations. For continuous variables this is a local mean or average value. For categorical variables this is the expected probability of each facies. We must now consider how to use the trend models. There are two

evident approaches for continuous variables (1) decomposition into a mean and residual, and (2) treating the trend as a secondary variable. These two approaches are discussed first and then categorical variables are discussed.

Decomposition into Mean and Residual

The idea is to decompose the original continuous Z variable (porosity or permeability) into a locally varying mean and a residual:

$$Z(\mathbf{u}) = m(\mathbf{u}) + r(\mathbf{u})$$

The locally varying mean $m(\mathbf{u})$ or $m(x, y, z)$ is defined at all locations over the entire reservoir. The residual $r(\mathbf{u})$ is only defined at data locations. The residual may be estimated at all locations and added to the mean for a numerical model; however, it is more appropriate to simulate multiple realizations of the residual and add them to the mean to create multiple realizations of the underlying continuous variable.

This approach is not without problems since the arbitrary separation of the data into trend and residual can easily lead to artifacts. One artifact is non-physical Z values, that is, negative porosities resulting from large negative residuals added to low mean values or too-high values resulting from large positive residuals and high mean values. Most artifacts are a consequence of modeling the residuals independently from the prior mean.

It is good practice to calculate the variance of the z-values, the m-values, and the r-values. By definition of the variance and covariance (see Chapter 2) we know that:

$$\sigma_Z^2 = \sigma_M^2 + \sigma_R^2 + C_{R-M}(0)$$

where σ_Z^2, σ_M^2 and σ_R^2 are the variances of the z, m and r values and $C_{R-M}(0)$ is the covariance between m and r. $C_{R-M}(0)$ can be negative or positive. If it is relatively close to zero then there should be few artifacts of considering them independently. The importance of the trend can be determined by the ratio σ_M^2/σ_Z^2, which is the fraction of variability explained by the trend or local mean values.

If the relation between R and M is significant, that is, $C_{R-M}(0)/\sigma_Z^2 > 0.15$, then artifacts are likely to be significant and the remnant correlation between the mean and residual must be accounted for.

The simple kriging formalism presented at the beginning of this Chapter implicitly assumes that a locally variable mean $m(\mathbf{u})$ is available at all locations.

Secondary Variable

The central idea here is to treat the locally varying mean $m(\mathbf{u})$ as a secondary data variable $Z_2(\mathbf{u})$ to be used in cokriging or cosimulation of the residual $r(\mathbf{u})$ or the original data variable $z(\mathbf{u})$. Considering the Z variable as the

primary variable has the advantage of direct enforcement of the Z histogram and variogram.

The local mean or secondary variable $Z_2(\mathbf{u})$ is available at all locations by mapping, see above. A conventional cokriging formalism would require a linear model of coregionalization (LMC) between the primary variable and the $Z_2(\mathbf{u})$ variable. The variogram of $Z_2(\mathbf{u})$ is likely to be very smooth due to the origin of this variable. The variogram of Z and R will not be as smooth. The LMC will require a mixture of nested variogram structures: some with great continuity (Gaussian) and some with less (spherical).

Since the secondary variable (the local mean) is available at all locations, a collocated alternative could be considered. As described above, this requires only the correlation coefficient between the primary and secondary variables. The smoothness of the secondary variable, however, may cause trouble in collocated cokriging.

Decomposition into mean and residual is the best approach unless significant problems arise in reproducing the histogram of the original Z values with simulation.

Categorical Variables

Locally variable facies proportions can be used with the indicator transformation of categorical variables presented above. Trend models in average facies proportions are geologically reasonable and expected. There are both vertical and areal trends. These trend models are constructed in the same way as for continuous variables.

Only one trend model is needed when there are two facies types; the proportion or probability of the second facies is one minus the proportion of the first. When there are three or more ($K > 2$) facies, the sum of the proportions should equal one. A restandardization step may be required:

$$p_k^* = \frac{p_k}{\sum_{k'=1}^{K} p_{k'}}$$

The way that the trend models enter subsequent modeling depends on the type of facies modeling adopted. The next chapter, Chapter 6, discusses cell-based modeling. Chapter 7 discusses object-based modeling. The specific approach to use these proportion values will be discussed there.

5.7 Work Flow

Figure 5.16 illustrates a work flow for kriging to build a map. The kriging type chosen determines whether a the mean must be explicitly mapped. The results are appropriate for visualizing trends in data and, perhaps, for smoothly varying surfaces.

Figure 5.17 illustrates the work flow for sequential Gaussian simulation (SGS) of a stationary random function. A declustered representative distribution and variogram model are the key inputs. There are a number of other details such as search, but those are essentially the same as required for kriging.

Figure 5.18 illustrates the work flow for indicator kriging for mapping local uncertainty. A very similar procedure could be used for indicator simulation of categorical or continuous variables.

Figure 5.19 illustrates the work flow to build a trend model for porosity, permeability, or facies modeling. For convenience, a 1-D vertical trend and 2-D areal trend are constructed first and then combined into a 3-D trend model.

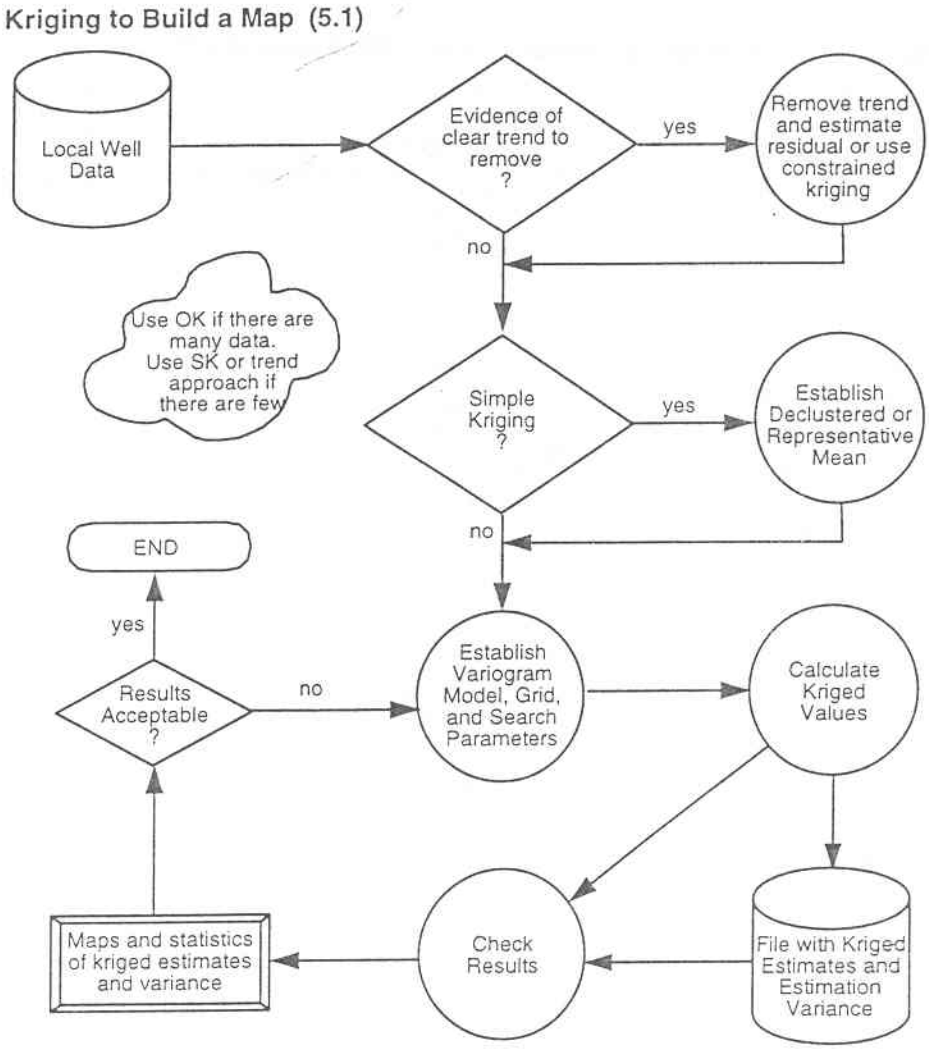


Figure 5.16: Work Flow 5.1: kriging to build a map.

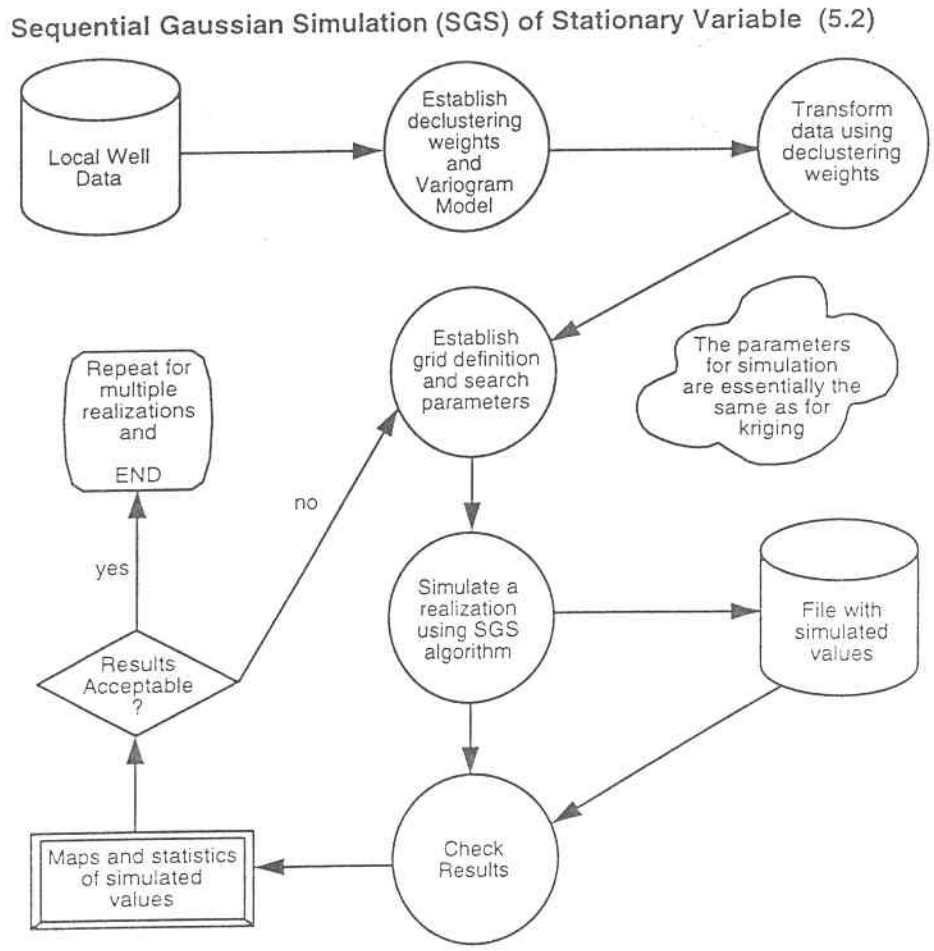


Figure 5.17: Work Flow 5.2: sequential Gaussian simulation (SGS) of a stationary random function.

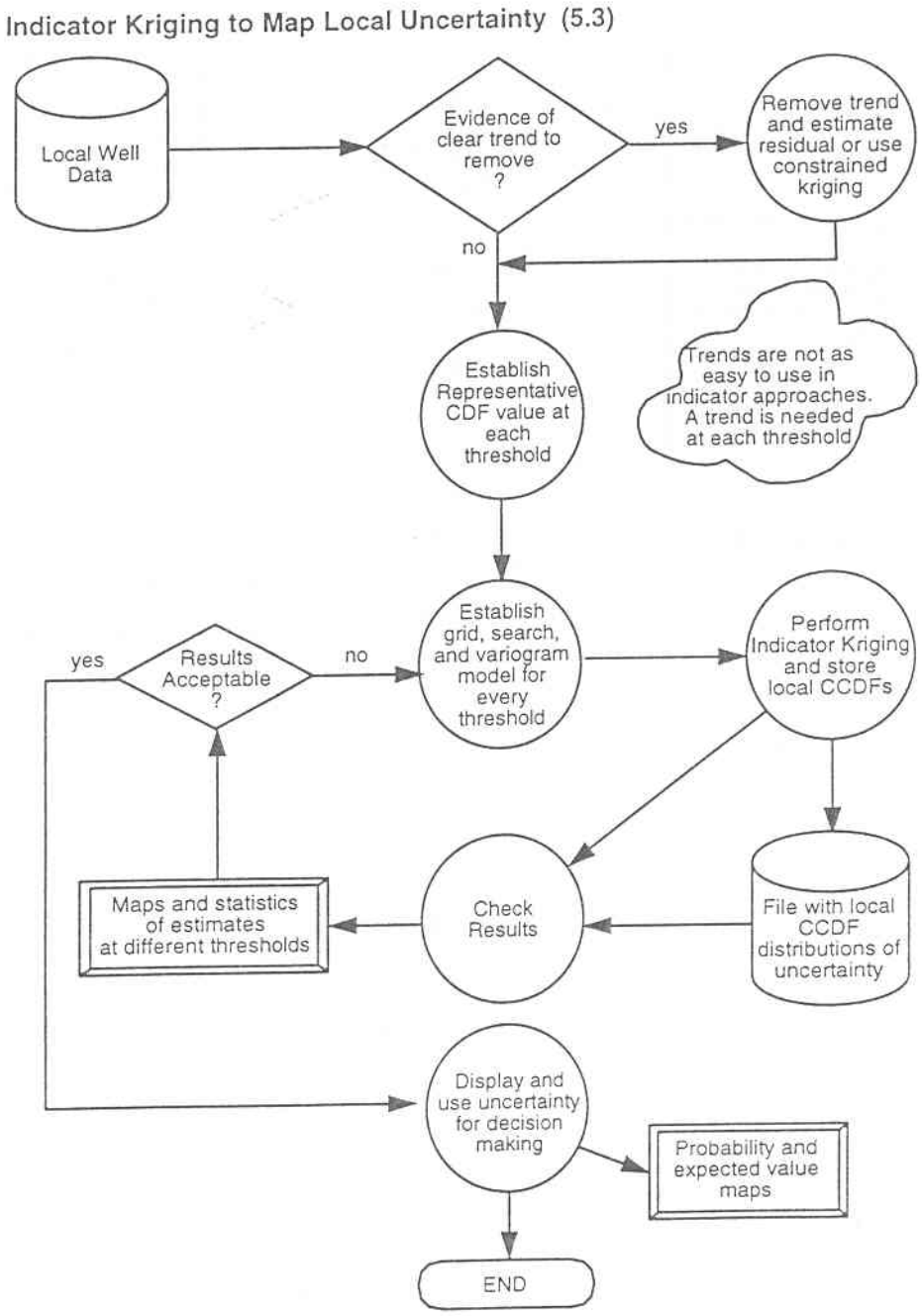


Figure 5.18: Work Flow 5.3: indicator kriging for mapping local uncertainty.

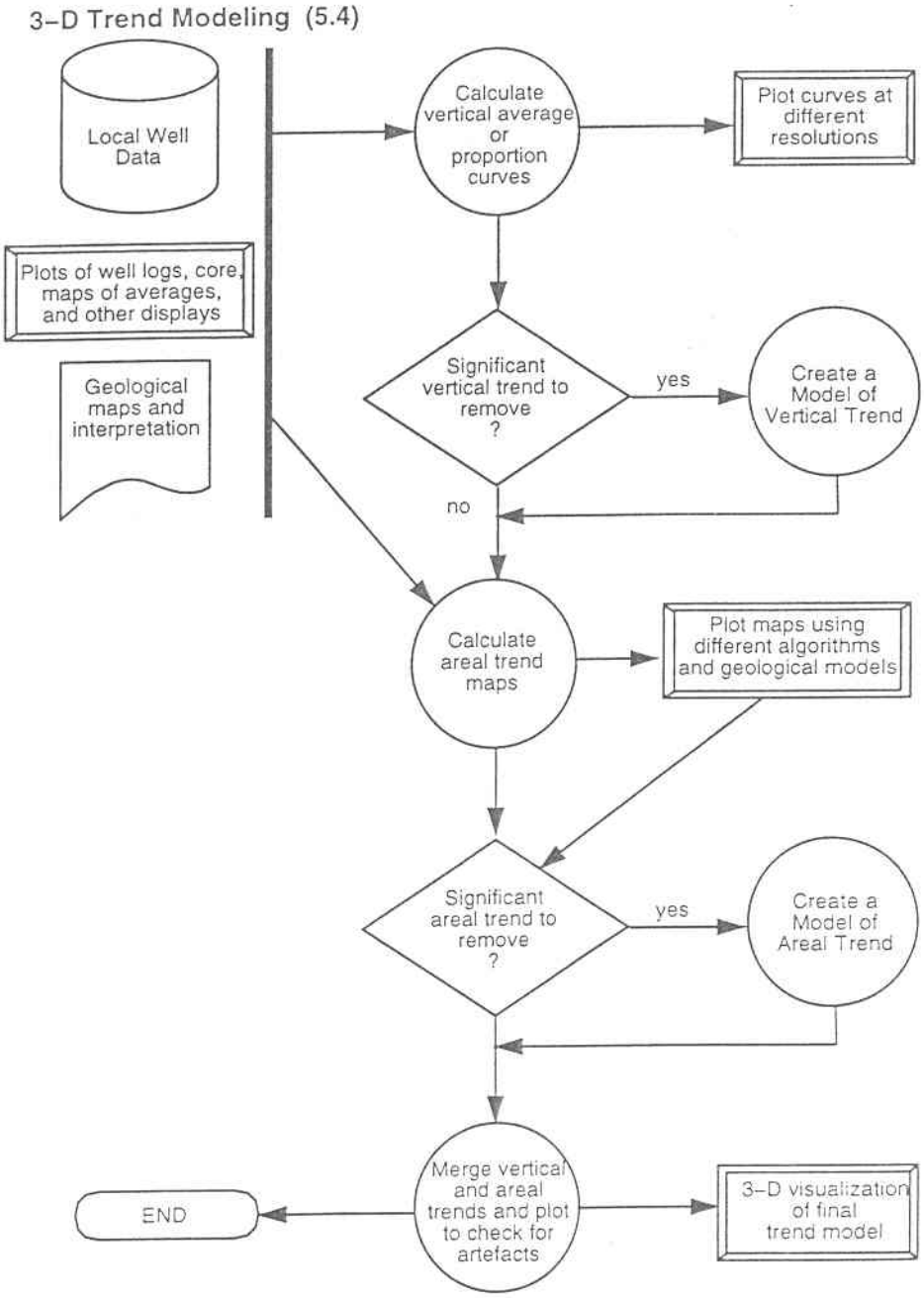


Figure 5.19: Work Flow 5.4: build a trend model for porosity, permeability, or facies modeling.

Chapter 6

Cell-Based Facies Modeling

Facies are often important in reservoir modeling because the petrophysical properties of interest are highly correlated with facies type. Knowledge of facies constrains the range of variability in porosity and permeability. Moreover, saturation functions depend on the facies even when the distributions of porosity and permeability do not. This chapter presents cell-based approaches to modeling facies.

Section 6.1 gives some guidance on choosing the appropriate facies modeling method. Object-based techniques are presented in detail in Chapter 7. They are introduced here in the context of choosing the most appropriate method.

Section 6.2 presents sequential indicator simulation (SIS), which is a widely used cell-based facies modeling technique. The background of SIS was discussed in Section 5.4. Implementation details including the integration of seismic data are discussed more thoroughly here.

Section 6.3 presents the truncated Gaussian simulation method. A categorical facies realization is created by truncating a continuous Gaussian simulation. The implementation details of determining the facies ordering, the continuous conditioning data, the right variogram, and integration of seismic data are presented.

Cell-based facies models often show unrealistic short-scale variations that can be removed with image processing techniques; Section 6.4 presents a simple and yet powerful algorithm for image cleaning.

6.1 Choosing the Appropriate Method

Facies are distinguished by different grain size or different diagenetic alteration. There are clear distinctions between sandstone and shale and, often,

between limestone and dolomite. Our concern is not how the facies types are defined, but how to construct realistic 3-D distributions of the facies that may be used in subsequent reservoir decision making.

The first question is whether or not to bother with facies before porosity and permeability modeling. Some considerations:

1. The facies must have a significant control on the porosity, permeability, or saturation functions; otherwise, modeling the 3-D distribution of facies will be of little benefit since uncertainty will not be reduced and the resulting models will have no more predictive power. There is unavoidable ambiguity in the definition of *significant control*: the histograms of porosity (permeability) within the different facies should have significantly different mean, variance, and shape, for example, by more than 30%. The saturation function measurements from different facies should not overlap.

2. The facies must be identifiable at the well locations, that is, from well log data, as well as core data. Careful examination of core data permits identification of many facies; however, there would be large uncertainty in the 3-D modeling of those detailed facies unless they can also be determined from the more abundant well log data. This consideration is easy to check; one must resist the temptation of taking geological detail available from visual examination of core data.

3. An additional constraint on the choice of facies is that they must have straightforward spatial variation patterns. The distribution of facies should be at least as easy to model as the direct prediction of porosity and permeability.

The choice of facies is a difficult case-dependent problem for which there may be no clear solution. Simplicity is preferred. Two facies, a "net" reservoir rock type and a "non-net" rock type, are often sufficient. At times there are two different "net" rock types of different quality, which have clearly different spatial variation patterns. Rarely should three or more be considered. At times there are two different "non-net" rock types, such as a depositional shale and a diagenetic-controlled cemented rock type, which have different spatial variation patterns. It is extremely rare to have to consider three or more "non-net" rock types. Thus, there are up to a maximum of six different facies types. Once the facies are defined, relevant data must be assembled and a 3-D modeling technique selected

The alternatives are (1) cell-based geostatistical modeling, (2) object-based stochastic modeling, or (3) deterministic mapping. Deterministic mapping is always preferred when there is sufficient evidence of the facies distribution to remove any doubt of the 3-D distribution. In many cases, there is evidence of geologic trends, which should be included in stochastic facies modeling.

The five main steps for stochastic facies modeling: (1) relevant well data together with areal and vertical geologic trend information must be assembled within the stratigraphic layer being modeled, (2) seismic attributes are often related to facies in an approximative way; if available, the seismic data must be tested to see if they calibrate to facies probabilities, (3) declustered, spatially representative, global proportions of facies must be determined for any facies modeling, (4) spatial statistics such as variograms or size distributions must be assembled as input to stochastic modeling, and finally, (5) 3-D facies models must be constructed and validated. Each of these steps will be discussed with the different facies modeling techniques.

Cell-based techniques are commonly applied to create facies models prior to porosity and permeability modeling [74, 75, 169, 193, 271]. The popularity of cell- based techniques is understandable: (1) local data are reproduced by construction, (2) the required statistical controls (variograms) may be inferred from limited well data, (3) soft seismic data and large-scale geological trends are handled straightforwardly, and (4) the results appear realistic for geological settings where there are no clear geologic facies geometries, that is, when the facies are diagenetically controlled or where the original depositional facies have complex variation patterns. Of course, when the facies appear to follow clear geometric patterns, such as sand-filled abandoned channels or lithified dunes, object-based facies algorithms (Chapter 7) should be considered.

Systems Tract Changes

Before presenting the details of cell-based facies modeling, we must consider major changes in the geological depositional environment. In sequence stratigraphic parlance, a *systems tract* is a particular form or nature of sedimentation. Many reservoirs fall into a single systems tract, for example, fluvial, deltaic, shoreface systems, or deepwater depositional systems. Nevertheless, the geological systems tract may change across the areal extent of the model. For example, fluvial and related terrestrial facies may give way to deltaic and other marine facies. Such large-scale changes in facies must be accounted for in facies modeling, regardless of whether cell- or object-based modeling is used.

Object-based modeling may be customized to handle such trends, that is, the objects are constructed to progressively change across the reservoir. Fluvial channels are related to their shoreface or marine counterparts; turbidite channels are related to their deeper water sheet and lobe counterparts. When such customization is not straightforward, areal and vertical trends may be considered. Such trends apply to cell-based facies modeling as well.

Another approach to account for large-scale changes in the depositional system is to deterministically divide the area into large-scale regions or domains. Then, the facies are modeled separately within each region. The regions are modeled sequentially, say from most proximal to most distal, with

data conditioning overlapping from one region to the next. This reproduces gradational changes in the facies at the borders between the regions. Separate modeling of different areas could cause artificial discontinuities at the boundaries.

6.2 Sequential Indicator Simulation

The building blocks of sequential indicator simulation (SIS) have been discussed. Section 2.5 discussed the calculation of representative global proportions. Section 5.6 discussed construction of locally varying proportion models. Chapter 4 covered all the details of variogram calculation. Section 5.4 was entirely devoted to the indicator formalism. There remain some implementation details, particularly with respect to seismic data, that should be discussed.

SIS is widely used for diagenetically controlled facies because the results have high variability and yet correct anisotropy and variogram measures of spatial correlation. We consider K mutually exclusive facies categories $s_k, k = 1, \ldots, K$. The indicator transform at a particular location $\mathbf{u}_j$ for a particular facies s_k is the probability of facies s_k prevailing at that location: 1 if it is present, and 0 otherwise.

There are K indicator transforms at each data location. Sequential indicator simulation consists of visiting each grid node in a random order. At each grid node a facies code will be assigned by the following steps:

1. Find nearby data and previously simulated grid nodes,
2. Construct the conditional distribution by kriging, that is, calculate the probability of each facies being present at the current location, $p_k^*, k = 1, \ldots, K$, and
3. Draw a simulated facies from the set of probabilities.

This entire procedure is repeated with different random number seeds to generate multiple realizations. A random path is considered to avoid any artifacts that can result from a combination of a regular path and a restricted search, which is needed to limit the kriging matrix size. Original data and previously simulated values are both used so that the realization has correct spatial structure between original data, between original data and simulated values, and between pairs of simulated values.

Figure 6.1 gives an illustration of sequential indicator simulation from a major Arabian reservoir [22]. This is an areal stratigraphic slice through a model with two facies: dolomite and limestone. The greater continuity in the Northeast direction is expected geologically and was revealed by the directional indicator variograms.

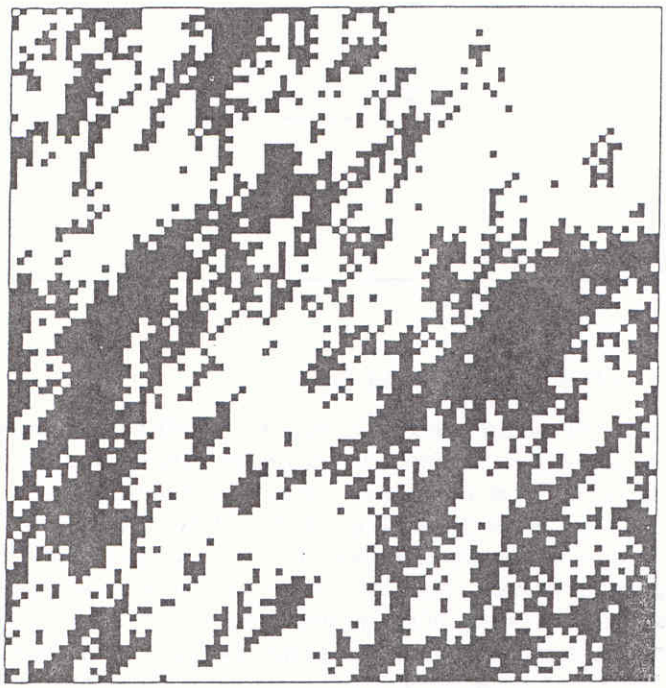


Figure 6.1: An illustration of sequential indicator simulation from a major Arabian reservoir (SPE 29869).

Seismic Data in SIS

Seismic data is often of great value in constraining facies models; it is areally extensive over the reservoir and can be sensitive to facies variations. Of course, facies vary over scales smaller than the seismic measurement volume and there is variability of acoustic properties within the same facies. A first step will be to calibrate the seismic data to facies proportions.

There are typically few well data and an abundance of seismic data. The horizontal indicator variograms for SIS are poorly defined from the well data alone. A solution consists of using the seismic data to help define the 3-D variogram models.

At the heart of indicator simulation is the use of kriging to determine the conditional probability of each facies type. In the presence of seismic data, some form of cokriging or block kriging must be considered to account for the seismic data. The third step will be to choose the form of cokriging and assemble the required input parameters.

Table 6.1: An illustration of the calibration of facies proportions to seismic acoustic impedance (ai).

ai-class	$p(k=1,ai)$	$p(k=2,ai)$	$p(k=3,ai)$
$-\infty$ - ai_1	-	-	-
ai_1 - ai_2	-	-	-
ai_2 - ai_3	-	-	-
...	-	-	-
ai_9 - ∞	-	-	-

Calibration of Seismic Data to Facies

Cokriging does not require the seismic data to be explicitly calibrated to facies proportions; the original seismic units could be retained and the calibration enters through the cross covariances between the hard indicator data and soft seismic data. Notwithstanding this flexibility of cokriging, calibrated probability values are preferred because (1) the calibrated probability values are in units we understand, and (2) cross variograms or covariances are often difficult to infer in presence of sparse well data.

Seismic data is called acoustic impedance (AI) in the following. There is no need to work with acoustic impedance; the same procedures could be used for any particular seismic attribute.

The range of AI-variability is divided into a series of classes. Too few classes and important relationships are lost. Too many classes and there are insufficient data in each class to reliably infer its relation to facies. In general, 10 classes based on the deciles of the AI histogram is sufficient.

Figure 6.2 shows a histogram of a seismic attribute at well locations. The upper plot is the cumulative histogram with the deciles specified by the horizontal gray lines. The 10 classes of seismic are identified by the 9 vertical gray lines that go from the cumulative histogram through the lower histogram. The classes are numbered from 1 through 10 in the center. Determining these classes is essentially automatic.

The calibration procedure consists of determining the probability of each facies for each class of seismic, that is,

$$p(k|ai_j), \quad k = 1, \ldots, K, \quad j = 1, \ldots, N_{ai}$$

where $p(k|ai_j)$ is the probability of facies type k for the jth seismic class ai_j. These probabilities are evaluated at the well locations. For each ai datum, at a well location, there are corresponding actual proportions of each facies. The result of this calibration exercise is a table of prior probabilities. Table 6.1 shows an empty calibration table for $K = 3$ facies types and 10 ai classes.

Of particular concern in constructing such a calibration table is the vertical resolution of the seismic data. The seismic data (ai-values) may be

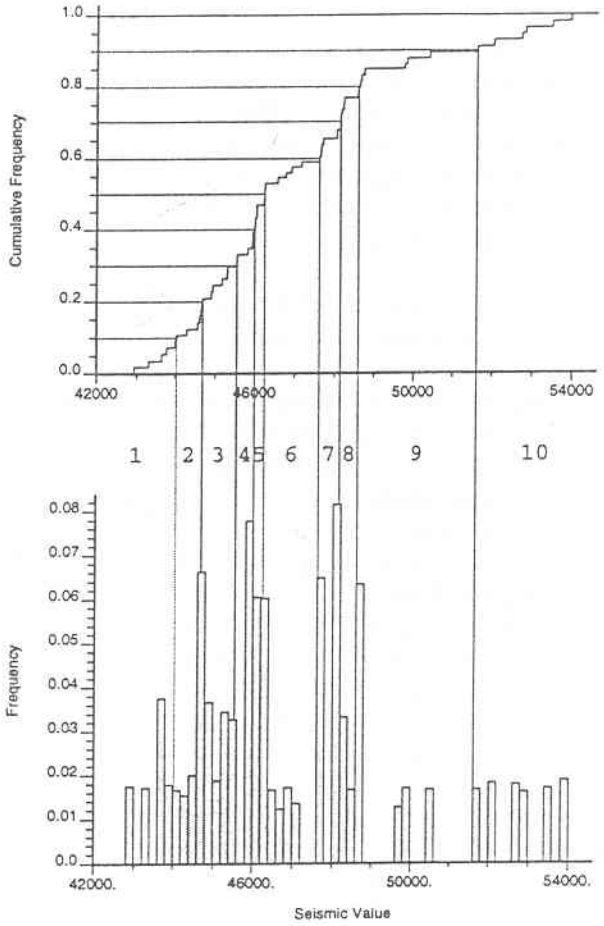


Figure 6.2: Cumulative histogram and histogram of seismic attribute with 10 classes defined by decile thresholds; there is the same number of data in each class.

recorded at a small sample rate (2 ms or less); however, the "real" vertical resolution may be much coarser. It is necessary to consider a vertical window of realistic size to calculate the prior probabilities. Otherwise, the calibration will not be representative of the true seismic resolution and short-scale noise may mask the information value of the seismic data.

The seismic data carries *information* when the calibration probabilities $p(k|ai_j)$ depart from the global probabilities of each facies, p_k. Ideally, the calibration probabilities will be close to 1 and 0; however, that is never the case due to small-scale variations of facies and the variability of acoustic properties within each facies.

The subject of indicator variogram inference will be discussed before describing the use these prior probabilities in cokriging.

Indicator Variogram Inference

Regardless of which variant of indicator simulation is used, we must infer a set of K indicator variograms that describe the spatial correlation structure of each facies type. The main challenge in variogram inference is the horizontal direction; there are often too few wells to calculate a reliable horizontal variogram. The vertical resolution, however, usually permits a reliable vertical variogram to be calculated. The seismic data provides an excellent source of information to guide the selection of horizontal parameters for the indicator variograms.

Each facies indicator variogram will be calculated in turn. The first step is to calculate the vertical indicator variogram $\gamma_k(\mathbf{h})$ from the well data and standardize it to unit sill by the variance $p_k(1-p_k)$. Secondly, we calculate the horizontal variogram from the corresponding seismic proportions, that is, the $p(k|ai_j)$ values. The seismic proportion variograms depend on the facies type $k = 1, \ldots, K$; however, there is unavoidable overlap since the proportions come from the same underlying acoustic rock properties. The seismic variogram is also standardized to unit sill. Finally, the vertical and horizontal variograms are fit with the shape of the vertical indicator variogram: only the length scales or ranges and any zonal anisotropy is taken from the seismic proportion variogram; see Chapter 4 for more detail on variogram modeling.

There are a number of assumptions in this hybrid approach to determine indicator variograms. The most important assumption is that the seismic proportions variogram provides a reasonable approximation to the horizontal indicator i range. This is reasonable if the seismic is highly correlated to the facies proportions and if the vertical averaging of the seismic is not too pronounced. One must consider slightly reducing the horizontal range by the horizontal support of the seismic data since the seismic variogram range is often longer than the indicator horizontal range (see Section 4.4).

Locally Varying Mean

Rather than cokriging, the calibrated seismic data probabilities can be used as locally varying means for kriging, that is, the probability of facies k is estimated by kriging as:

$$\hat{i}(\mathbf{u};k) = \sum_{\alpha=1}^{n} \lambda_\alpha \cdot i(\mathbf{u}_\alpha;k) + \left[1 - \sum_{\alpha=1}^{n} \lambda_\alpha\right] \cdot p(k|ai(\mathbf{u}))$$

where $\hat{i}(\mathbf{u};k), k = 1, \ldots, K$ are the estimated local probabilities to be used for simulation, n is the number of local data, $\lambda_\alpha, \alpha = 1, \ldots, n$ are the weights. $i(\mathbf{u}_\alpha;k)$ are the local indicator data, and $p(k|ai(\mathbf{u}))$ is the seismic-derived probability of facies k at location $\mathbf{u}$ being considered. When there are few local data, the sum of the weight to the data is small and the seismic- derived probabilities receive a large weight. When the weight to local data is high the seismic-derived probabilities are given low weight.

The set of probabilities $\hat{i}(\mathbf{u};k), k = 1, \ldots, K$ must be corrected for order relations, see Chapter 5; otherwise, the SIS procedure remains unchanged. This procedure is simple and quite effective. There is no explicit variogram control on the influence of the seismic data. Some form of cokriging would be required for such control.

The Bayesian Updating Approach

The simplest form of cokriging is collocated cokriging or the *Bayesian Updating Approach* [69, 72, 73]. This method is gaining in popularity because of its simplicity and ease with which seismic is accounted for.

At each location along the random path (recall procedure described in Section 5.4), indicator kriging is used to estimate the conditional probability of each facies, $i^*(\mathbf{u};k), k = 1, \ldots, K$, from hard i data alone, then Bayesian updating modifies or updates the probabilities as follows:

$$i^{**}(\mathbf{u};k) = i^*(\mathbf{u};k) \cdot \frac{p(k|ai(\mathbf{u}))}{p_k} \cdot C \quad k = 1, \ldots, K \tag{6.1}$$

where $i^{**}(\mathbf{u};k), k = 1, \ldots, K$ are the updated probabilities for simulation, $p(k|ai(\mathbf{u}))$ is the seismic-derived probability of facies k at location $\mathbf{u}$ being considered, p_k is the overall proportion of facies k, and C is a normalization constant to ensure that the sum of the final probabilities is 1.0. The factor $p(k|ai(\mathbf{u}))/p_k$ operates to increase or decrease the probability depending on the difference of the calibrated facies proportion from the global proportion. No change is considered if $p(k|ai(\mathbf{u})) = p_k$, which is the case where the seismic value $ai(\mathbf{u})$ contains no new information beyond the global proportion. An additional step is required to ensure exactitude, that is, exact reproduction of facies observations at well locations.

The simplicity and utility of this approach is appealing. There are two implicit assumptions behind Bayesian updating that may be important: (1) the collocated seismic data screens nearby seismic data, and (2) the scale of the seismic is implicitly assumed to be the same as the geological cell size. A relatively simple check on the simulated realizations will judge whether these assumptions are causing artifacts. The calibration table (just like Table 6.1) can be checked using the seismic data and the SIS realization; a close match to the probabilities calculated from the original data indicates the approach is working well. Problems would be revealed by high probabilities (large $p(k|ai)$'s) becoming exaggerated to even higher values in the final model; low values would be even lower.

Markov-Bayes Soft Indicator Kriging

Once the seismic data has been calibrated to probability units, indicator kriging can be used to integrate that information into improved conditional probability values for SIS [3, 145, 280].

The Bayesian approach described above is one approach to update the results of indicator kriging using well data alone. Different sources of information such as well data and seismic data are available. With enough data one could directly calculate and model the matrix of covariance and cross covariance functions (one cross covariance for each facies k). An alternative to this difficult exercise is provided by the Markov-Bayes model [280]. For simplicity, we write this model in terms of covariance. The three covariances needed are $C_I(\mathbf{h};k)$: the covariance of the *hard* facies indicator data for facies type k, which is calculated from the well data or a combination of well and seismic data (see Chapter 4 and above), $C_{IS}(\mathbf{h};k)$: the cross covariance between the hard facies indicator data and the seismic probability values (the $p(k|ai(\mathbf{u}))$ values), and $C_S(\mathbf{h};k)$: the covariance of the seismic probability values. The latter two covariances are given by the following model:

$$\begin{aligned} C_{IS}(\mathbf{u};k) &= B_k C_I(\mathbf{h};k),\ \forall\, \mathbf{h} \\ C_S(\mathbf{u};k) &= B_k^2 C_I(\mathbf{h};k),\ \forall\, \mathbf{h} > 0 \\ &= |B_k| C_I(\mathbf{h};k),\ \mathbf{h} = 0 \end{aligned} \tag{6.2}$$

The coefficients B_k are obtained by simple manipulation of the calibrated seismic probabilities:

$$B_k = E\{P(k|ai(\mathbf{u})) \mid I(\mathbf{u};k) = 1\} - E\{P(k|ai(\mathbf{u})) \mid I(\mathbf{u};k) = 0\} \quad \in\ [-1,+1] \tag{6.3}$$

$E\{\cdot\}$ is the expected value operator or simply the arithmetic average. The term $E\{P(k|ai(\mathbf{u})) \mid I(\mathbf{u};k) = 1\}$ is close to 1 if the seismic data is good, that is, the seismic-predicted probability of the facies being present is very high if the facies is present. The term $E\{P(k|ai(\mathbf{u})) \mid I(\mathbf{u};k) = 0\}$ is close to 0 if the seismic data is good, that is, the seismic-predicted probability of the facies being present is very low if the facies is *not* present.

The K parameters, $B_k, k = 1,\ldots,K$, measure how well the soft seismic probabilities distinguish the different facies. The best case is when $B_k \approx 1$, and the worst case is when $B_k = 0$. When $B_k = 1$, the seismic-probability data $p(k|ai(\mathbf{u}))$ are treated as hard indicator data. Conversely, when $B_k = 0$, the seismic data are ignored, that is, their kriging weights will be zero.

Keeping the closest seismic-datum amounts to the collocated cokriging option [274]. This model and alternatives are covered in Chapter 4.

Block Cokriging

The versions of cokriging presented above are all approximations that try to avoid (1) calculation of cross variograms or covariances, (2) fitting a full model of coregionalization, and (3) handling the large- scale nature of seismic data. We could do the job right. All variograms and cross variograms could be calculated (see Chapter 3). These variograms could be downscaled to point scale assuming the seismic data provides information on a linear average of facies indicator data. A model of coregionalization could be fit (see Chapter

3). Finally, a block cokriging could be performed in SIS that would account for the actual scale of the seismic data *and* the "softness" of the calibration from acoustic impedance to facies proportions. The cokriging formalism is classical.

Most geostatistical modeling software consider the different data types to be at the same spatial scale. Nevertheless, once a small scale model of coregionalization has been established, linear spatial averages of the variogram can be easily used in the cokriging equations. Implementation of block cokriging in a conventional SIS program (such as `sisim` in GSLIB [62]) is straightforward.

Comments on Implementation

SIS is applicable to heterogeneous facies that have been diagenetically altered or that have no clear geometric shapes associated to discontinuities in petrophysical properties. The steps for SIS have been described above and are summarized in the work flows at the end of this chapter. Some comments:

- There is a need for representative global proportions calculated via declustering or reliable local proportions determined through trend modeling.
- An indicator variogram model is needed for each facies type. In fact, there are $K-1$ degrees of freedom in the required indicator variograms. Only one variogram is needed for a binary system since the continuity of one facies completely specifies that of the other. We calculate and fit all K variograms in presence of $K \geq 3$; inconsistency between the K variograms will be revealed in order relations problems.
- The kriging type used in the sequential simulation framework is adapted to handle seismic data.

The lengthy discussion on the type of cokriging to use for seismic data could be confusing. The methods increase in complexity and data requirements. The simplest is best, provided it does not contradict available data.

Nested Indicators

An increasingly common implementation of SIS is to apply it in a nested or hierarchical manner. For example, consider a first SIS to separate net and non-net facies. Then, a second SIS could be applied within the net facies to distinguish different types of net facies. This procedure allows the "nested" relationships between the facies to be effectively captured. The truncated Gaussian simulation also considers transitions between facies.

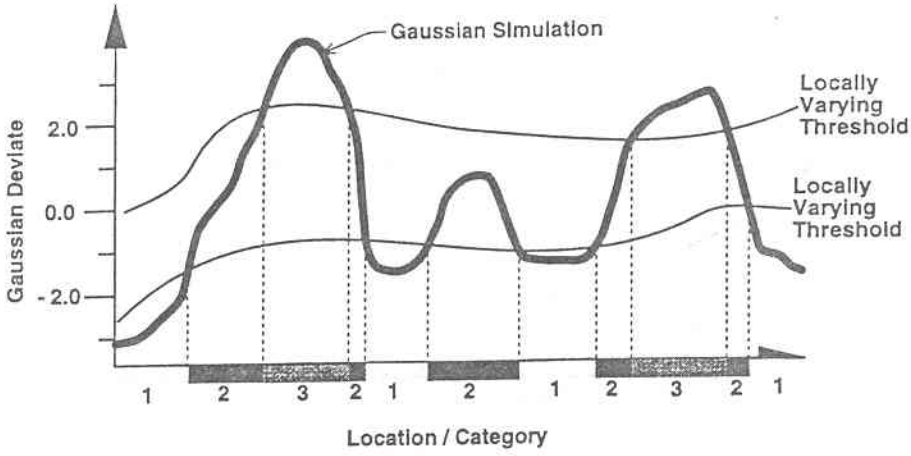


Figure 6.3: Schematic illustration of how truncated Gaussian simulation works: a continuous Gaussian variable is truncated at a series of thresholds to create a categorical variable realization.

6.3 Truncated Gaussian Simulation

The key idea Truncated Gaussian simulation [25, 187, 272] is to generate realizations of a continuous Gaussian variable and then truncate them at a series of thresholds to create a categorical facies realizations.

A 1-D schematic example is shown on Figure 6.3. Although this example is only 1-D, it illustrates many of the concepts behind truncated Gaussian simulation. The categorical simulation, shown along the lower axis, is derived from a continuous Gaussian simulation shown by the black curve. The thresholds for truncating the Gaussian variable to facies need not be constant; we see a greater proportion of facies 1 toward the right side where the first threshold is higher.

Ordering of Facies

An important feature of truncated Gaussian simulation is the ordering of the resultant facies models. The facies codes are generated from an underlying continuous variable. In the three-facies example of Figure 6.4 we would most often see facies 2 between 1 and 3. Only rarely would code 1 be next to code 3; when the continuous variable changes very quickly from low to high values.

This ordering is a significant advantage of truncated Gaussian simulation when dealing with facies that *are* ordered. For example, the facies may be defined by the amount of shale: 1 = shale, 2 = interbedded sand and shale, and 3 = clean sand. Truncated Gaussian simulation may work well for this case. SIS will work better when there is no clear ordering. The correct

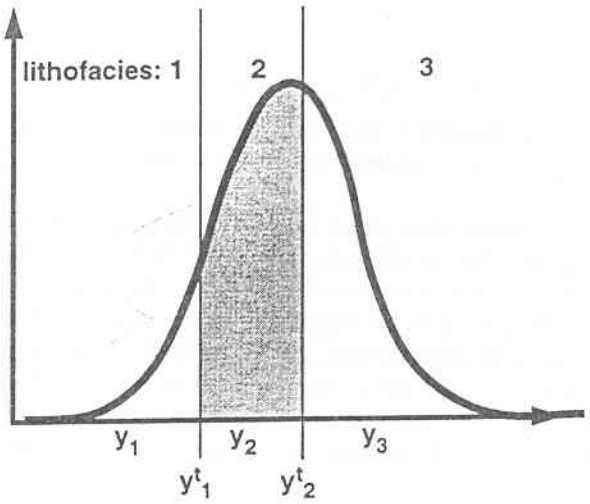


Figure 6.4: Standard normal distribution with three facies codes. There are two thresholds (y_1^t and y_2^t) and three continuous y values for each facies (y_1, y_2, and y_3). The proportion of each facies is the area under the standard normal curve.

ordering for the facies is usually evident from the geological context of the problem.

Proportions, Thresholds, and Conditioning Data

The facies proportions are taken as constant or they have been modeled as locally variable with the trend modeling techniques of Chapter 5. In any case, assume that the proportion of each ordered facies is known at each location $\mathbf{u}$ in the layer A, that is, $p_k(\mathbf{u}), k = 1, \ldots, K, \mathbf{u} \in A$. These proportions can be turned into cumulative proportions:

$$cp_k(\mathbf{u}) = \sum_{j=1}^{k} p_j(\mathbf{u}), \quad k = 1, \ldots, K, \ \forall \mathbf{u} \in A \tag{6.4}$$

where $cp_0 = 0$ by definition and $cp_K = 1.0$ since the proportions sum to 1.0 at all locations. The $K-1$ thresholds for transforming the continuous Gaussian variable to facies are given by:

$$y_k^t(\mathbf{u}) = G^{-1}\left(cp_k(\mathbf{u})\right), \quad k = 1, \ldots, K-1, \ \forall \mathbf{u} \in A \tag{6.5}$$

where $y_k^t(\mathbf{u}), k = 1, \ldots, K-1, \forall \mathbf{u} \in A$ are the thresholds for the truncated Gaussian simulation, $y_0^t = -\infty$, $y_K^t = +\infty$, $G^{-1}(\cdot)$ is the inverse cumulative distribution function for the standard normal distribution,[1] and $cp_k(\mathbf{u}), k =$

[1] Although there are no closed-form analytical expressions for this function, there are excellent polynomial approximations.

$1, \ldots, K-1$ are the cumulative probabilities for location $\mathbf{u}$, see (6.4). Given a normal deviate or variable, these thresholds may be used to assign a facies code:

$$\text{facies at } \mathbf{u} = k \text{ if } y^t_{k-1}(\mathbf{u}) < y(\mathbf{u}) \leq y^t_k(\mathbf{u}) \tag{6.6}$$

The locally variable thresholds (6.5) introduce the trend information. The Gaussian values $y(\mathbf{u})$ are considered stationary, that is, $E\{Y(\mathbf{u})\} = 0, \forall \mathbf{u} \in A$.

The categorical facies data must be transformed into continuous Gaussian conditioning data for conditional simulation of the stationary Gaussian variable, Y. There are two considerations. First, the local proportions and corresponding thresholds must be considered to ensure that the correct facies are obtained on back transformation. Second, the facies are categorical and some decision must be made about the "spikes" caused by categorical data. Handling locally varying proportions and thresholds is straightforward: a local transformation is used. Despiking the categorical data requires additional decisions.

Random despiking (used in the `nscore` program of GSLIB) would introduce unreasonable short-scale randomness. A more sophisticated despiking (see Chapter 2) could be used, but with little advantage. Decisions regarding the parameters for despiking may introduce other artifacts. The simplest solution is to set the normal score transform of each facies to the center of the class in the normal distribution, that is, just leave the spikes unchanged:[2]

$$y(\mathbf{u}) = G^{-1}\left(\frac{cp_{k-1}(\mathbf{u}) + cp_k(\mathbf{u})}{2}\right) \tag{6.7}$$

where $y(\mathbf{u})$ is the normal score transform at location $\mathbf{u}$, k is the facies code at location $\mathbf{u}$, and the cp_ks are the cumulative proportions as in equation (6.4). All facies measurements are transformed and used in conditional Gaussian simulation (for example `sgsim` from GSLIB).

Variogram

Truncated Gaussian simulation requires a single variogram for the y variable. This is a convenience in the sense that K indicator variograms need not be calculated and modeled in a consistent manner. This is also a significant disadvantage since there is no ability to control the different patterns of variability for different facies. In general, it is not a good idea to directly calculate and fit the y-variogram with the y-data defined above in (6.7) due to the spikes. In the case of two facies (a single threshold y^t_1), the normal score covariance is directly linked, through a one-to-one relation, to the facies indicator covariance [158]. We could exploit this analytical relation, but it only exists for the case of two facies.

[2] The reduced variance (less than 1.0) of the input y data may cause an artifact in the y-realizations, but this will not be seen in the categorical variable realization.

One way around this limitation would be to consider a series of Gaussian RFs $Y_k(\mathbf{u}), k = 1, \ldots, K$, each used to simulate (after truncation) a nested set of only two facies; that is, simulate each set nested into a previously simulated set. This would be a more complex version of the nested indicator approach mentioned in the previous Section 6.2. This approach is not developed further since the facies rarely have such an arbitrary pattern of "nested" variability.

Another approximation consists of inverting the single covariance model $C_Y(\mathbf{h})$ from some average of the K indicator variograms. It is this single average indicator variogram that would be reproduced through simulation, not any particular facies indicator variogram model.

Recent efforts to extend the truncated Gaussian method for simulation of multiple ($K > 2$), non-nested, facies call for multiple Gaussian RFs. Contrary to the case $K = 2$, simulation of multiple, non-nested, facies involves a highly non-unique inversion. Iterative procedures based on trial and error selection of the input normal score covariances and cross covariances between the multiple Gaussian RFs have been considered, yet with limited success [174].

The recommended approach is to take the most important indicator variogram and "invert" it numerically to the corresponding normal score variogram. This can be done numerically with code available with the paper by Kyriakidis, Deutsch, and Grant, 1998. The normal score variogram has a similar range and anisotropy; the shape is smoother, that is, more parabolic or "Gaussian" at short distances. Figure 6.5 illustrates the relationship between the indicator and Gaussian variogram for a series of different facies proportions.

Figure 6.6 illustrates the consequence of using the wrong variogram for truncated Gaussian simulation. The Gaussian realization, truncated indicator realization, and indicator variograms on the left side correspond to using the standardized indicator variogram for the normal score (Gaussian) realization. Note that the resulting indicator variogram has a too-high nugget effect. The realization on the right side corresponds to using the theoretically correct normal score variogram.

Comments on Implementation

Truncated Gaussian simulation is applicable to heterogeneous facies that have been diagenetically altered, that have no clear geometric shapes, and that are ordered in some predictable manner. The steps for truncated Gaussian simulation have been described above and are summarized in the work flows at the end of this chapter. Some comments:

- As with SIS, there is a need for representative global proportions calculated via declustering or reliable local proportions determined through trend modeling. These proportions determine the locally varying thresholds and conditioning data.

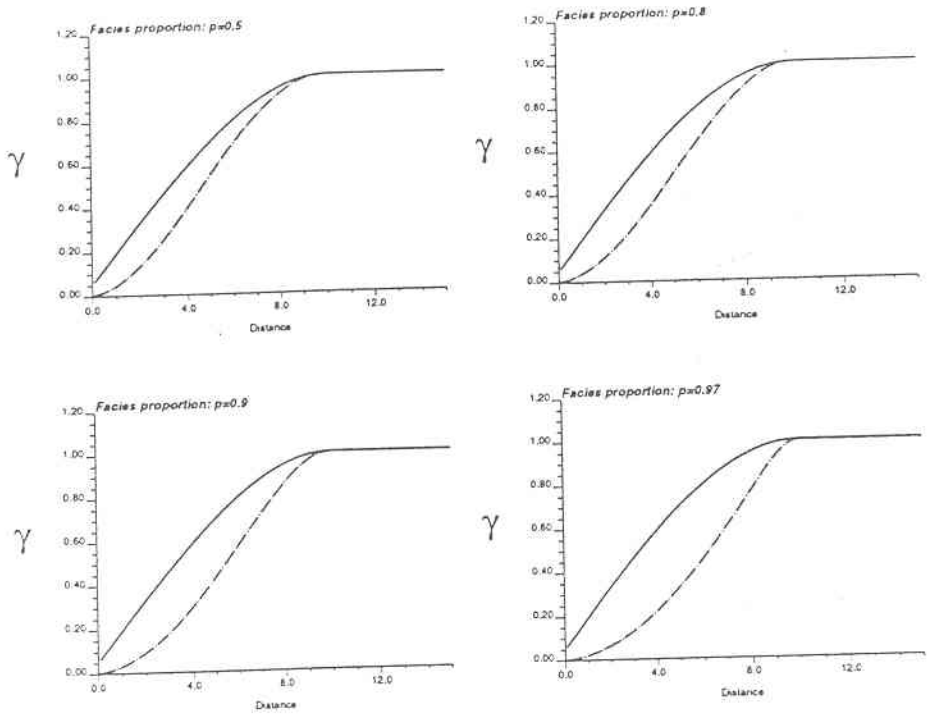


Figure 6.5: Corresponding standardized indicator variograms shown by the solid lines and normal score semivariograms shown by the dashed lines for different univariate proportions.

- An indicator variogram model is needed for the most important facies type. This variogram is converted to a variogram for simulation of the Gaussian variable; in practice, the shape is changed to a Gaussian shape.
- Realizations of conditional Gaussian simulation are truncated by the locally varying thresholds to create facies realizations.

Truncated Gaussian simulation requires essentially the same professional involvement and CPU effort as SIS. The consideration to choose one or the other is based on the ordering or nesting of the facies. Truncated Gaussian simulation should be used when there is a clear ordering and no significant difference in anisotropy of the various facies geometries.

Figure 6.7 shows horizontal slices through a SIS realization and a truncated Gaussian realization. These realizations are models of a reservoir offshore Africa. There is no particular ordering or proximity of the facies in the SIS realization. The facies are clearly ordered in the truncated Gaussian realization. The variograms used to create the SIS realizations are shown

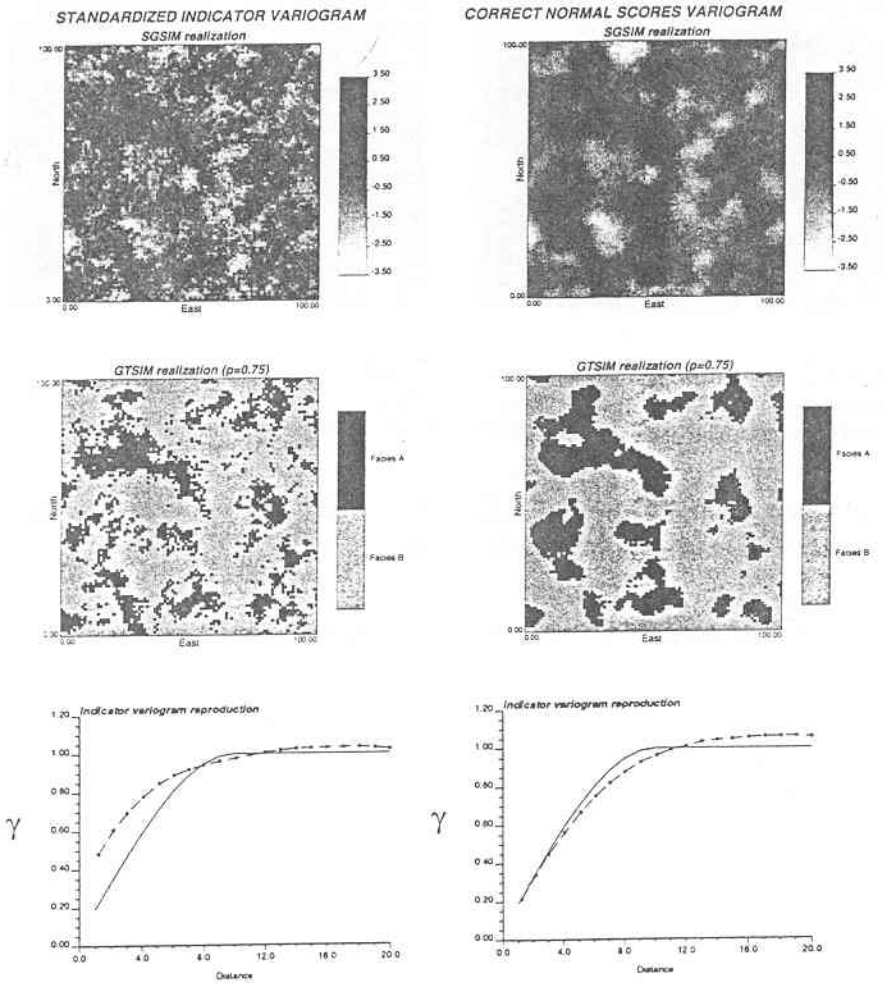


Figure 6.6: An illustration of the consequence of using the wrong variogram for truncated Gaussian simulation. The Gaussian realization, truncated indicator realization, and indicator variograms on the left side correspond to using the standardized indicator variogram for the normal score (Gaussian) realization. Note that the resulting indicator variogram has a too-high nugget effect. The realization on the right side corresponds to using the theoretically correct normal score variogram.

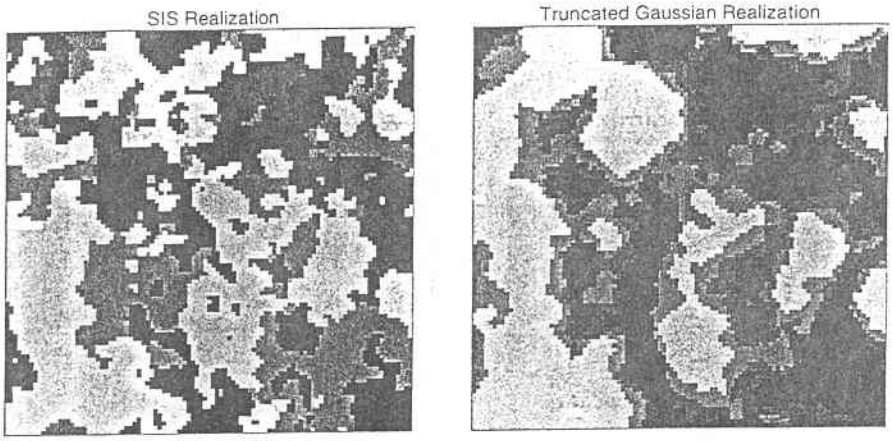


Figure 6.7: Horizontal slices through a SIS realization and a truncated Gaussian realization. The same essential variogram structure is present in both realizations; however, note the ordering of the facies in the truncated Gaussian realization.

in Figure 6.8. The 3-D variogram used to create the truncated Gaussian realization on the right of Figure 6.7 is shown on Figure 6.9.

6.4 Cleaning Cell-Based Facies Realizations

One concern with cell-based facies realizations is the presence of unavoidable short-scale variations, which appear geologically unrealistic. In some cases, such variations affect flow simulation and predicted reserves; a more justifiable reason to consider realization cleaning algorithms. A second concern is that the facies proportions often depart from their target input proportions. In particular, facies types with relatively small proportions (5 to 10%) may be poorly matched. In indicator simulation, the main source of this discrepancy is the order relations correction (the estimated probabilities are corrected to be non-negative and sum to 1.0). There is no evident alternative to the commonly used order relations correction algorithms; post-processing the realizations to honor target proportions is a convenient and attractive solution. An unavoidable consequence of this post-processing is a reduction in the space of uncertainty. The uncertainty in the global proportions could be established ahead of time by, for example, a bootstrap and then used in the cleaning algorithm described below.

The general problem of image cleaning has been tackled by a number of workers in the area of image analysis and statistics [14, 24, 100, 112]. One proposal to clean facies realizations, closely related to some of these image analysis methods, is based on the concepts of dilation and erosion [219]. This approach is well suited to cleaning binary (only two facies) images;

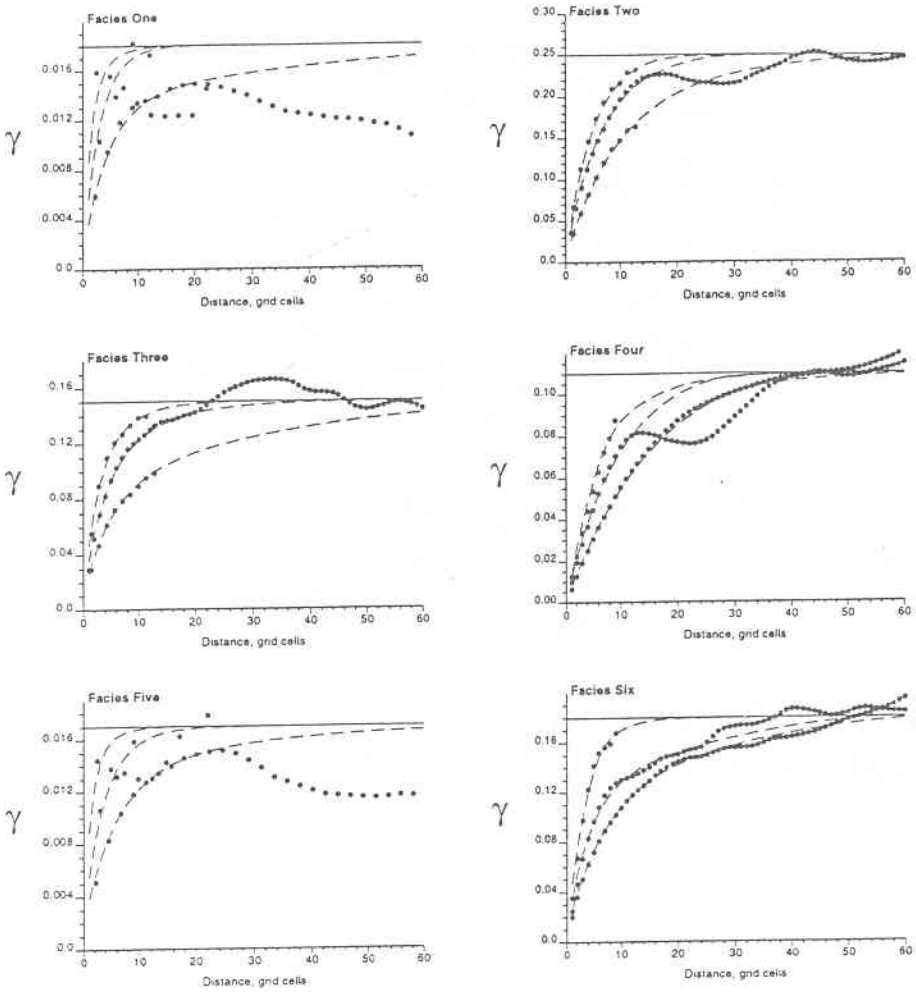


Figure 6.8: The six indicator variograms (the three directions for each variogram are shown by the three dashed lines) used to create SIS realization of Figure 6.7.

however, there is no general extension to more than two facies types. One could consider a nested application two facies at a time.

Iterative or simulated annealing-type algorithms can be designed to clean facies realizations. These methods can be very powerful; however, there are a number of practical problems (1) they tend to be CPU intensive, (2) it is difficult to determine the appropriate values for a number of tuning parameters, and (3) often, a training image is required.

The *maximum a-posteriori selection* or MAPS technique, amounts to replacing the facies type at each location $\mathbf{u}$ by the most probable facies type based on a local neighborhood. The probability of each facies type, in the

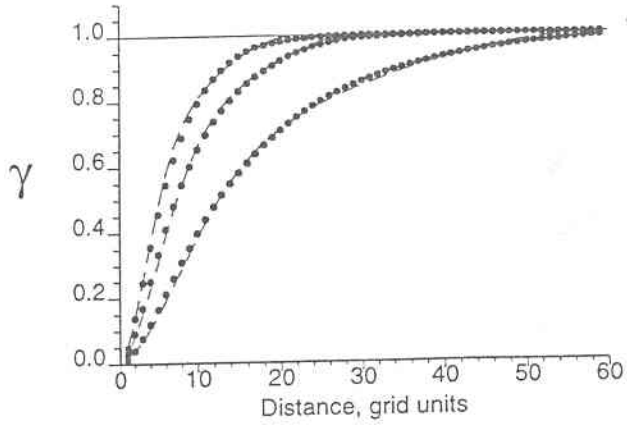


Figure 6.9: Variogram used for the Gaussian simulation needed for truncated Gaussian simulation; see right side of Figure 6.7.

local neighborhood, is based on (1) closeness to the location $\mathbf{u}$, (2) whether or not the value is a conditioning datum, and (3) mismatch from the target proportion.

Consider an indicator realization $i_k^{(0)}(\mathbf{u})$, $k = 1, \ldots, K$, $\mathbf{u} \in A$ of N locations taking one of K facies types, $s_k, k = 1, \ldots, K$. There is no requirement for the facies $s_k, k = 1, \ldots, K$ to be in any particular order. The proportions of each facies type in the realization are $p_k^{(0)} = Prob\{I_k^{(0)} = 1\} \in [0, \ldots, 1], k = 1, \ldots, K$ with $\sum_k p_k^{(0)} = 1$. The target proportions of each facies type are $p_k, \in [0, \ldots, 1], k = 1, \ldots, K$ with $\sum_k p_k = 1$.

Consider the following steps to clean the realization $i_k^{(0)}(\mathbf{u})$, $k = 1, \ldots, K$, $\mathbf{u} \in A$ and to bring the proportions $p_k^{(0)}, k = 1, \ldots, K$ closer to the to target probabilities $(p_k, k = 1, \ldots, K)$. At each of the N locations $\mathbf{u} \in A$, calculate a local probability $q_k(\mathbf{u})$ based on a weighted combination of surrounding indicator values:

$$q_k(\mathbf{u}) = \frac{1}{S} \sum_{\mathbf{u}' \in W(\mathbf{u})} w(\mathbf{u}') \cdot c(\mathbf{u}') \cdot g_k \cdot i_k^{(0)}(\mathbf{u}'), \quad k = 1, \ldots, K \tag{6.8}$$

where S is a standardization constant to enforce $\sum_k q_k(\mathbf{u}) = 1.0$, $W(\mathbf{u})$ is a template of points centered at location $\mathbf{u}$, and $w(\mathbf{u}')$, $c(\mathbf{u}')$, and $g_k, k = 1, \ldots, K$ are weights to account for closeness to $\mathbf{u}$, conditioning data, and mismatch from the global proportions, more precisely:

- $w(\mathbf{u}')$ = weights that define a neighborhood or template to achieve realization cleaning. The definition of the local neighborhood and the weights $w(\mathbf{u}')$ within the neighborhood control the extent of realization cleaning. The appearance of the results is used to determine the optimal neighborhood (see examples below).

- $c(\mathbf{u}')$ = weight to ensure reproduction of conditioning data. $c(\mathbf{u}') = 1.0$ at all non-data locations, and equal to $C; C \geq 10$ at conditioning data locations, $\mathbf{u}' = \mathbf{u}_\alpha, \alpha = 1, \ldots, n$. This discontinuous function ensures that the conditioning data are given priority.

- g_k = weight to ensure that the facies proportions in the cleaned image are closer to the target global proportions: the probability of a particular facies q_k will be decreased if the original realization has a too-high proportion of that facies and will be increased if the original realization has a too-low proportion, specifically,

$$g_k = \frac{p_k}{p_k^{(0)}}, \quad k = 1, \ldots, K$$

  where, $p_k, k = 1, \ldots, K$ are the target proportions and $p_k^{(0)}, k = 1, \ldots, K$ are the proportions in the initial realization. Further scaling factors ($f_k, k = 1, \ldots, K$ multiplying the g_k values) could be determined by trial-and-error to force a closer match.

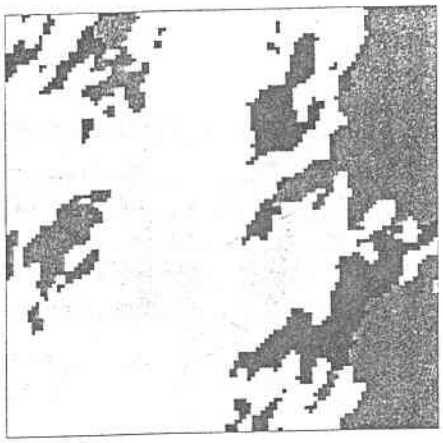
The size of the local neighborhood and the nature of the weights $w(\mathbf{u}')$ have a significant impact on the "cleanliness" of the results. In general, the image will appear cleaner when a larger window is considered. Figure 6.10 illustrates how an SIS realization is cleaned with three different templates W and weights $w(\mathbf{u}')$. Anisotropy can be introduced to the template W and weights $w(\mathbf{u}')$ to avoid "smoothing" of anisotropic features. Trial-and-error must be used to determine the weights for any specific case.

Conditioning data are enforced by the weights $c(\mathbf{u}')$ within the local neighborhood; $c(\mathbf{u}')$ is equal to 1.0 at all non-data locations, and equal to C, a larger number, at conditioning data locations, $\mathbf{u}' = \mathbf{u}_\alpha, \alpha = 1, \ldots, n$. Figure 6.11 shows two SISIM realizations before and after cleaning. The well data are reproduced in both cases. The facies observed at the vertical well are propagated continuously away from the well. The extent of the propagation depends on the magnitude of C and the size of the cleaning template W; a conditioning datum can affect only those nodes within the template W.

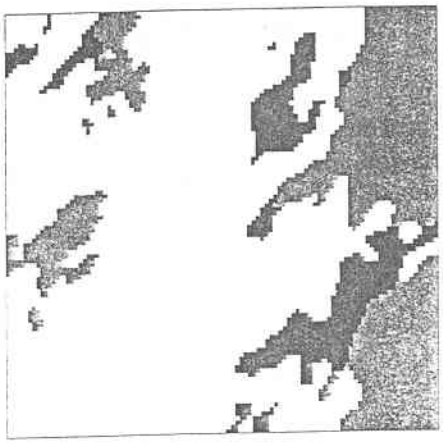
The weights $g_k, k = 1, \ldots, K$ lead the cleaned image toward the global proportions: the probability of a particular facies is decreased if the original realization has a too-high proportion of that facies and will be increased if the original realization has a too-low proportion. As expressed in (6.8), the proportions are not imposed, the probabilities are simply adjusted to be closer. In general, variability in facies proportions is an inherent aspect of uncertainty, and perfect reproduction of target proportions is not a critical goal.

The motivation for cleaning categorical facies realizations is to correct noisy facies realizations that are often geologically unrealistic and have a different flow character than clean images. A drawback of realization cleaning algorithms is that *real* short-scale variations in facies may be removed for the questionable sake of nice, clean, pretty pictures.

1	2	1
2	3	2
1	2	1



1	1	1	1	1
1	2	3	2	1
1	3	5	3	1
1	2	3	2	1
1	1	1	1	1



1	1	1	1	1	1	1
1	1	2	3	2	1	1
1	2	3	5	3	2	1
1	3	5	7	5	3	1
1	2	3	5	3	2	1
1	1	2	3	2	1	1
1	1	1	1	1	1	1

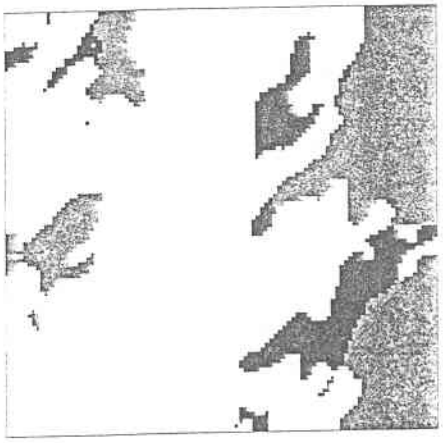


Figure 6.10: Three templates and the corresponding cleaned images starting from an SIS realization.

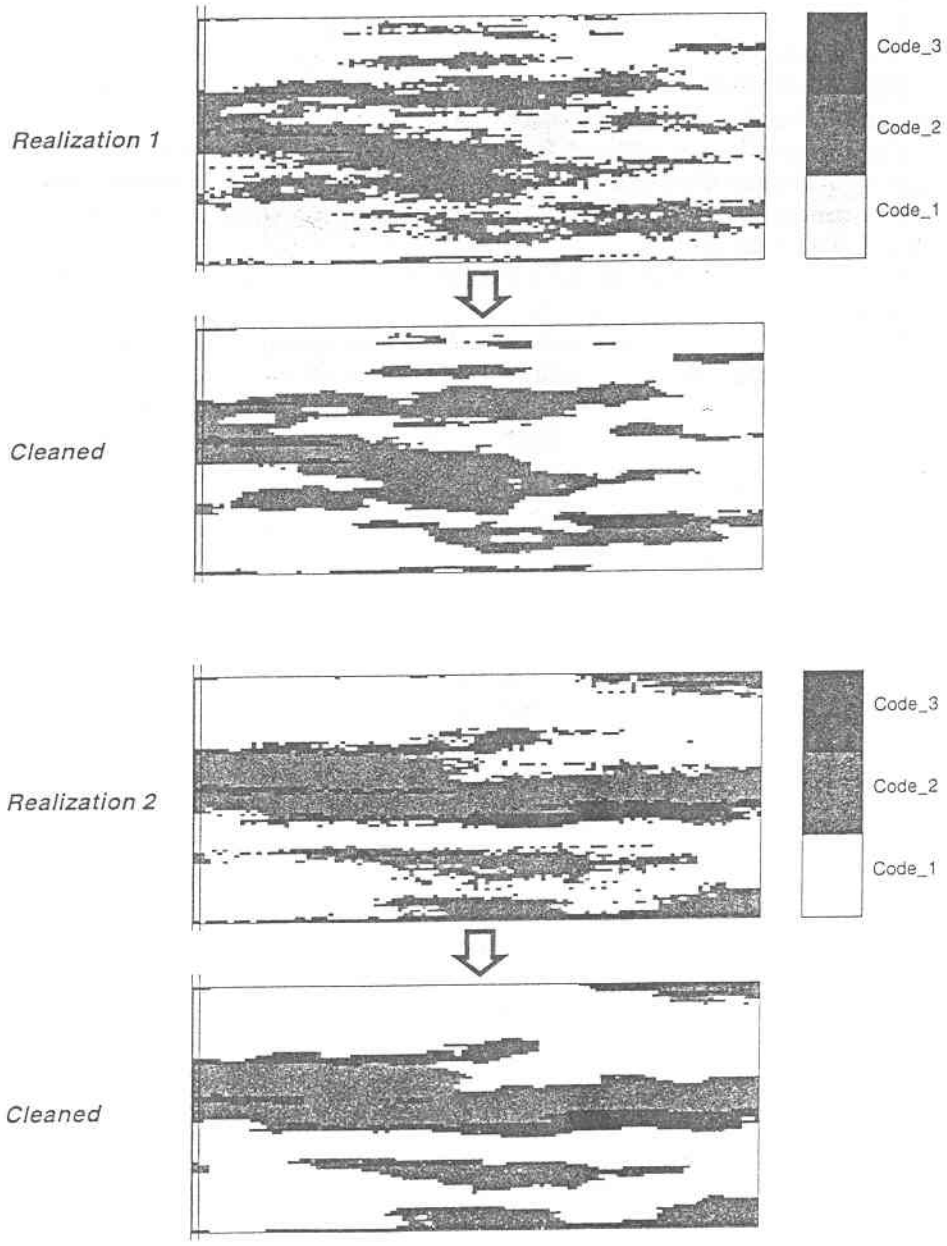


Figure 6.11: Two SISIM realizations before and after cleaning. Note the reproduction of the conditioning data at the vertical string of well data at the left.

6.5 Work Flow

Figure 6.12 illustrates a work flow for the calibration of facies proportions from seismic data. The results can be used with SIS or truncated Gaussian facies simulation.

Figure 6.13 illustrates the work flow for facies modeling using sequential indicator simulation (SIS) with seismic data.

Figure 6.14 illustrates the work flow for the pre-processing steps required prior to truncated Gaussian simulation. The locally varying thresholds and proportions must be established. The normal score transformation of the facies is passed to SGS.

Figure 6.15 illustrates the work flow for facies modeling using truncated Gaussian simulation. Stationary SGS is described in Chapter 5.

Figure 6.16 illustrates the work flow for cleaning cell-based facies realizations. The work flow is iterative to achieve the *correct* level of cleaning. Care should be taken to avoid cleaning real consequential short-scale facies variations.

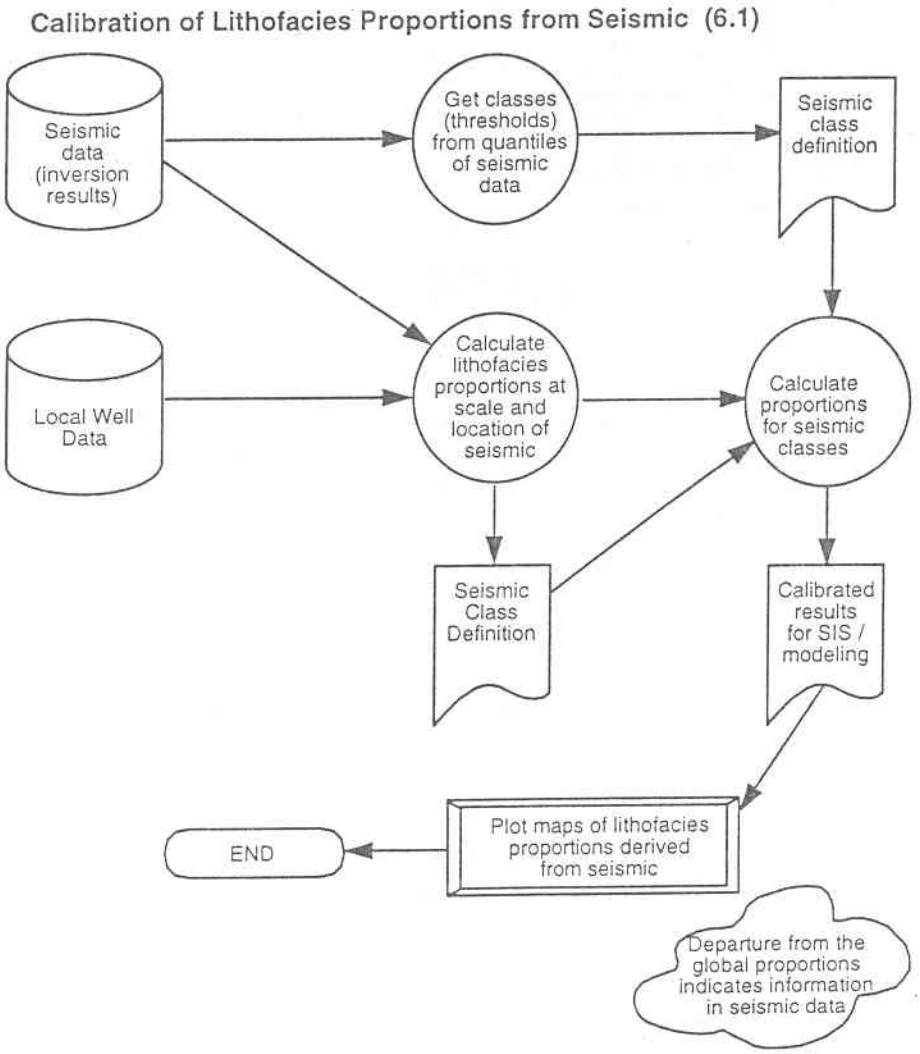


Figure 6.12: Work Flow 6.1: calibration of facies proportions from seismic data.

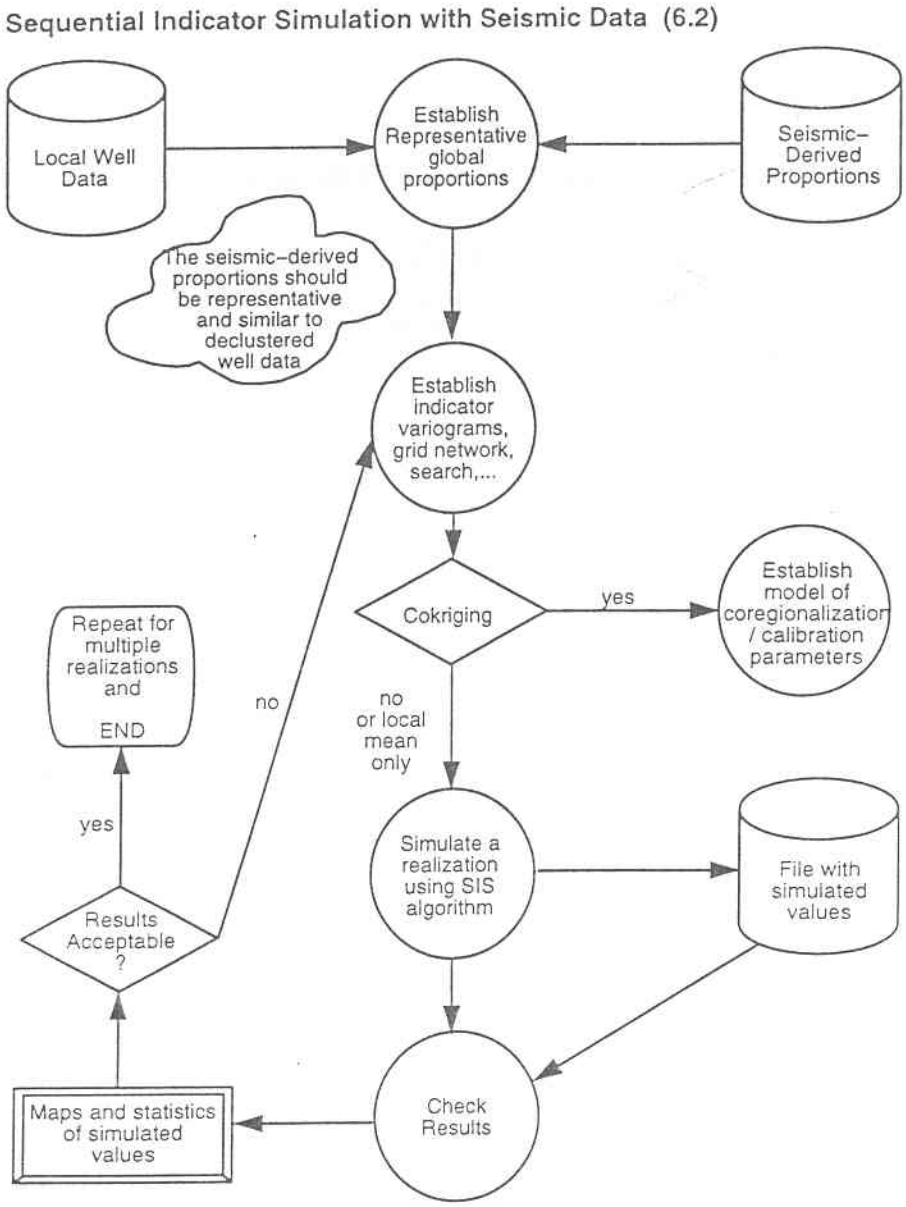


Figure 6.13: Work Flow 6.2: facies modeling using sequential indicator simulation (SIS) with seismic data.

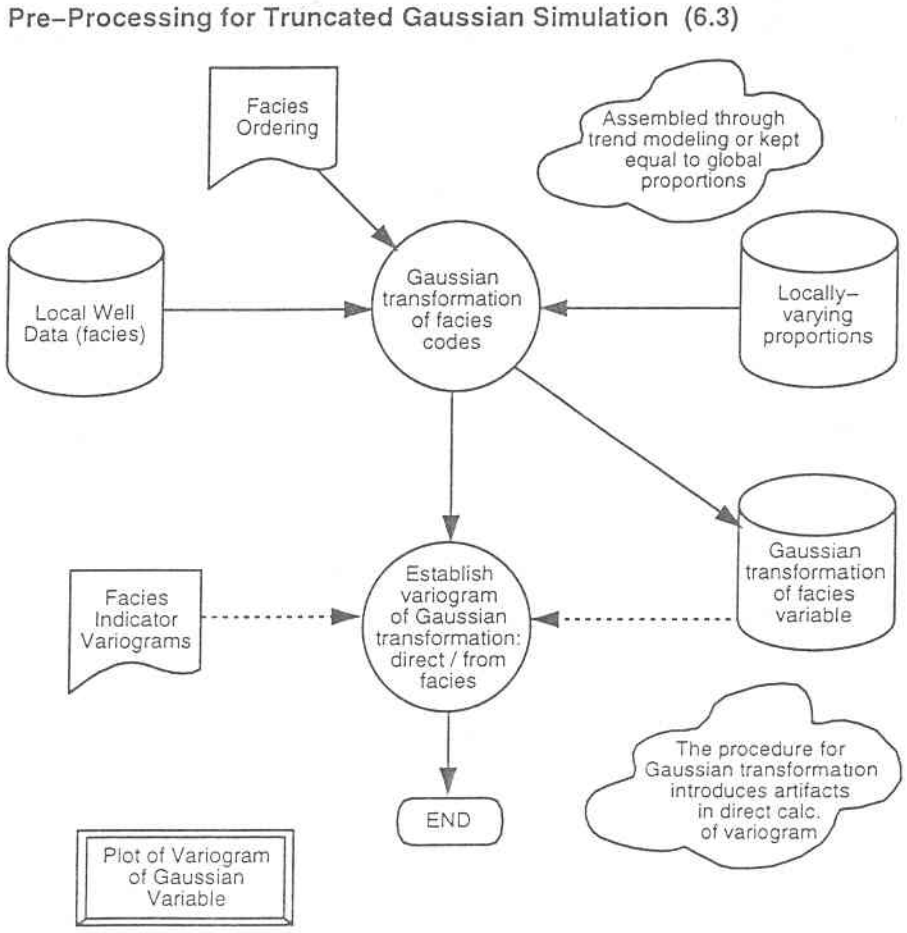


Figure 6.14: Work Flow 6.3: pre-processing steps required prior to truncated Gaussian simulation.

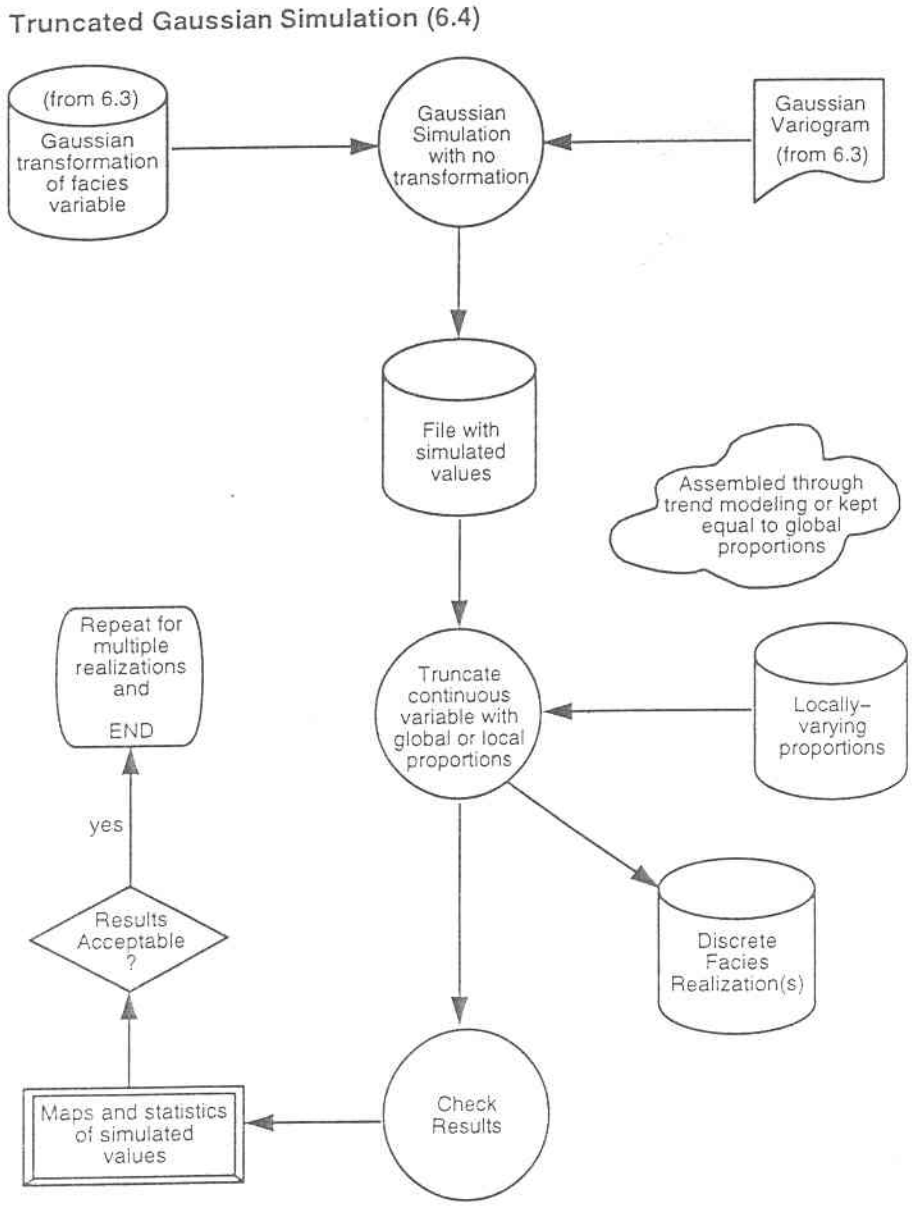


Figure 6.15: Work Flow 6.4: facies modeling using truncated Gaussian simulation.

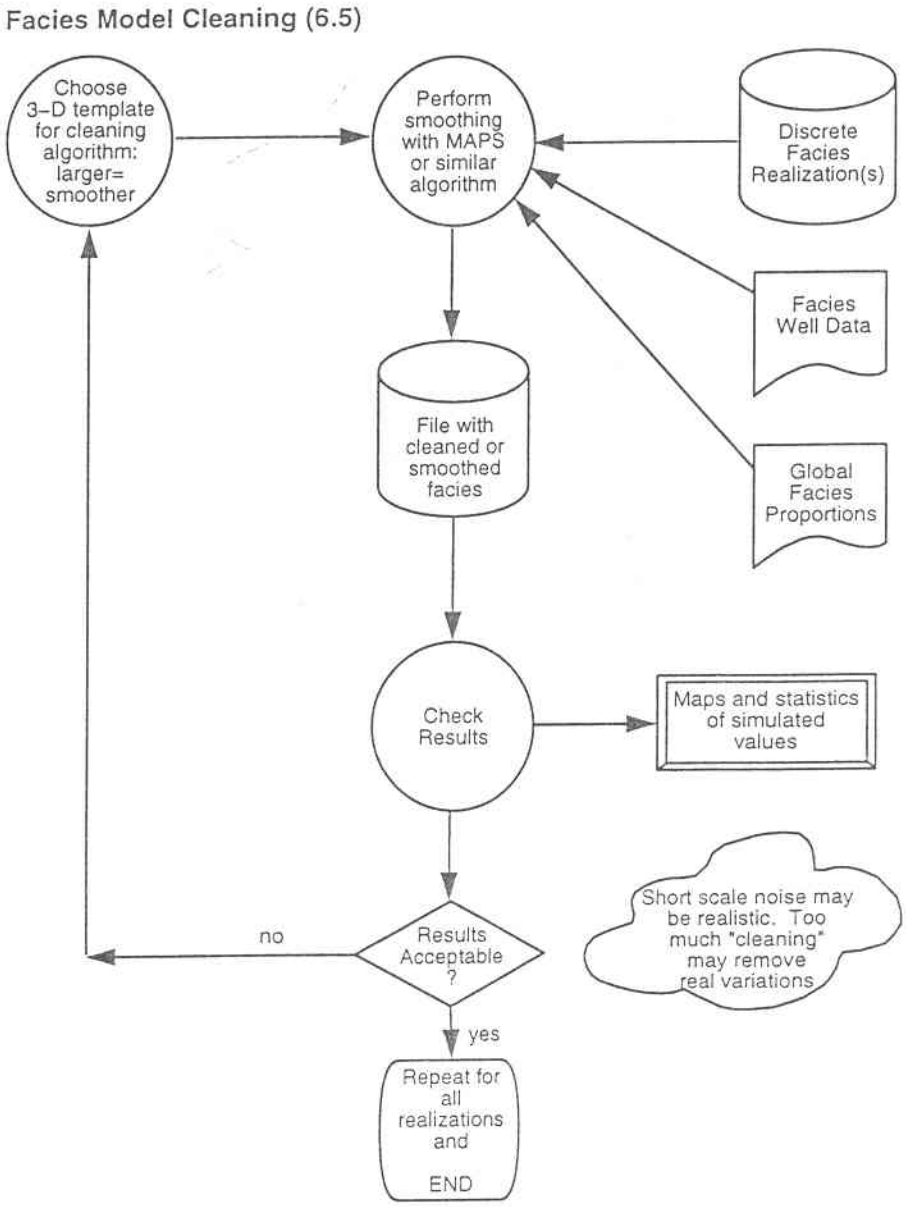


Figure 6.16: Work Flow 6.5: cleaning cell-based facies realizations.

Chapter 7

Object-Based Facies Modeling

Object-based facies models are visually attractive because the resulting facies realizations mimic idealized geometries interpreted in outcrops and modern analogues. Such models show facies belonging to clean geological shapes with realistic non-rectilinear continuity, which cannot be modeled with cell-based approaches. This chapter gives some details of modeling facies with object-based techniques.

Section 7.1 gives background on object-based techniques including an overview of some geological shapes that could be considered, the relevant data that object-based models must reproduce, and algorithms for object placement.

Section 7.2 presents the concepts of object-based modeling in the context of stochastic shales. Object-based modeling gained wide popularity in the context of modeling fluvial facies. Section 7.3 presents some details related to modeling abandoned sand-filled fluvial sand channels and related margin facies (levee and crevasse sands). Implementation details including the integration of well and seismic data are discussed.

Section 7.4 presents considerations for other depositional systems. In particular, the application to deepwater depositional systems. Finally, Section 7.5 presents work flow diagrams for some operations related to object-based facies modeling.

7.1 Background

From a geological perspective, it is convenient to view reservoirs from a chrono-stratigraphic perspective. The sedimentary architecture is considered in light of a hierarchical classification scheme. The reservoir facies are divided into sequences, parasequences, bed sets, beds, and so on. For example, het-

erogeneities in a fluvial setting are described by stratigraphic reservoir layers, channel complexes, channels, levees, and crevasses through additional smaller scale features. We consider modeling this genetic hierarchy of heterogeneities by successive coordinate transformations and geometric objects representing facies associations. Porosity and permeability models are then constructed at the appropriate scale using coordinate systems aligned with depositional continuity.

Significant effort has been focused on sedimentology and stratigraphy of reservoir systems. It is not possible to provide a reasonable overview of such studies here. Consider fluvial deposits as an example. The recent book by Miall [189] provides a well-illustrated description of fluvial sedimentary facies, basin analysis, and petroleum geology with more than 500 figures and 1000 references. The literature describing fluvial deposits is rich and varied. The history of quantitative computer models for fluvial systems is also extensive. Allen's early qualitative work in the 1960s and 1970s [7] led to quantitative computer simulations [8]. Leeder, at about the same time, was also building quantitative models [172]. Bridge published in this area with Leeder [30] and also published computer code [29] that was updated recently [179].

Although not specifically designed for fluvial facies, object-based models became popular in petroleum reservoir modeling in the mid-1980s due to the work of Haldorsen and others [117, 120, 236]. The importance of fluvial reservoirs in the Norwegian North Sea soon prompted the development of these object-based methods for fluvial facies [38, 45, 86, 113, 126, 203, 235]. The theory and implementation was refined over a number of years [101, 124, 134, 242, 250, 251, 253] with increasing practical application of these methods to Norwegian North Sea reservoirs [28, 252]. Such applications have set the standard for other oil producing regions of fluvial depositional setting. Other non-Norwegian oil companies also developed object-based modeling capability [6, 161].

Object-based models are now created routinely in reservoir characterization. The three key issues to be addressed in setting up an object-based model are the (1) geological shapes and their parameter distributions, (2) algorithm for object placement modification, and (3) relevant data to constrain the resulting realizations.

We must exercise a little caution in this chapter. This is *not* a rigorous presentation of geological concepts for different depositional systems. A pragmatic approach is considered. Pragmatic from a geological, flow modeling, and geostatistical perspective. The idea of modeling reservoirs by genetic forward modeling is attractive, but not considered here. Object-based models are *pseudo*-genetic in the sense that erosional rules (younger rocks on top) and other geological principles are used as much as possible.

There is the inevitable question of semantics and word choice. I choose to use "object-based" modeling. Of course, these object-based models end up as a collection of cells, but they are not called cell-based. The reference to object-based is in the method of creation and not the format of the result. Some

prefer "Boolean" models, which has a connotation of greater statistical rigor. Others prefer "marked point process" connoting a statistical point process for object centroids and then object properties such as size, orientation, and facies type are attributed to the point process.

Geological Shapes

There is no inherent limitation to the shapes that can be modeled with object-based techniques. The shapes can be specified by equations, a raster template, or a combination of the two. The geological shapes can be modeled hierarchically, that is, one object shape can be used at large scale and then different shapes can be used for internal small-scale geological shapes. Some evident shapes:

- Abandoned sand-filled fluvial channels within a matrix of floodplain shales and fine-grained sediments. The sinuous channel shapes are modeled by a 1-D centerline and a variable cross section along the centerline. Levee and crevasse objects can be attached to the channels. Shale plugs, cemented concretions, shale clasts, and other non-net facies can be positioned within the channels. Clustering of the channels into channel complexes or belts can be handled by large-scale objects or as part of the object placement algorithm.

- Lower energy meandering fluvial systems can be modeled as sand lenses within a non-net background. We sometimes consider modeling the entire channel (as above) and then assigning sand and shale within the channel in some realistic manner.

- Other channelized depositional systems including deepwater and estuarine systems are often modeled by adapting fluvial channel modeling techniques to system-specific considerations such as channel size, width to thickness ratios, and internal heterogeneities. The transition from channels to other facies types can also be handled by some rule-based scheme, that is, a channel may evolve into a more lobe-like geometry at some distance into the modeling domain (as energy is lost and the sediments disperse).

  Channel-type systems are discussed in detail in Section 7.2.

- Eolian dune shapes are obvious candidates for object-based modeling. Although the 3-D geometry of such object is not trivial to define analytically, we can make the necessary assumptions for practical modeling.

- Remnant shales may be modeled as disk or ellipsoid objects within a matrix of sand. This may be appropriate in high net-to-gross reservoirs. Although such shales may have a low proportion they will significantly affect the vertical permeability, hence, horizontal well production and coning.

- Deltaic reservoirs may be modeled as fan-shaped sand units within a matrix of poorer quality sediments. Historically, cell-based techniques have been used more in such systems.
- Shoreface environments have a systematic progression from proximal or near-shore facies through distal or offshore facies. The position of the shoreline could be modeled as a function of geological time (position in the stratigraphic layer) and the particular ordering of facies linked to this position. A hybrid cell- and object-based modeling scheme would be appropriate.
- Carbonate reservoirs may consist of predictable shapes. Less attention is paid to carbonates in this respect; however, there has been modeling of algal mounds and other clear shapes.

Another form of object-based modeling is "surface-based modeling" that is described briefly in Chapter 11. Stratigraphic surfaces are viewed as geological objects and modeled with a Boolean object-based formalism.

Relevant Data

A realistic facies model should reproduce all available data within the volumetric support and precision of each data source. The relevant data include local well data, seismic data, production data, and geological interpretations including deterministic objects, connections, and trends.

Local well data consist of identified facies codes at arbitrary locations in the 3-D reservoir layer. The facies may come from direct core observation or inference using a variety of well logs. Local well data take the form of intersections of facies types at arbitrary orientation and location.

Seismic and production data are most often imprecise large-scale information. There are different ways of representing such information. One way is to calculate the seismic-derived probability of each facies

$$p_k(\mathbf{u}_V), \;\; k = 1, \ldots, K, \;\; , \forall \mathbf{u}_V \in A$$

where $p_k(\mathbf{u})$ is the probability or proportion of facies k at location $\mathbf{u}_V$, there are $k = 1, \ldots, K$ facies codes, and the location $\mathbf{u}_V$ is defined everywhere in the reservoir A. The V subscript is the volume over which the facies proportions are defined. Note that a slightly different notation was considered for the cokriging formalism considered in Chapter 6.

Similar locally varying facies proportion information may be obtained from geological interpretation and mapping. Knowledge of the depositional system often allows specification of areas of greater and lesser proportions of different facies.

Seismic, production, and geological data may not be able to distinguish between some of the facies. For example, crevasse splay and levee sands may be indistuishable from the data. An important constraint in object-based modeling will be to reproduce proportions of "grouped" facies.

The location of some geological objects may be known exactly from seismic observation, interpretation of well test or production data, or from geological mapping. The positioning of such objects must be fixed and not left to an algorithm for stochastic placement. There are other types of information regarding object placement: (1) knowledge of the exact position of some objects, (2) knowledge that two (or more) intersections of the same facies *cannot* relate to the same object, or (3) a probabilistic measure of two (or more) intersections belonging to the same facies. These "connections" come from geological interpretation of the particular cored facies and from engineering data that relate pressure of fluid continuity between intersections in different wells. Many implementations of object-based modeling makes allowance for such interpreted "connections."

Lastly, the set of relevant data includes distributions of uncertainty related to the object sizes, shapes, orientations, and interactions. These parameters are specific to a particular problem, but are typically associated with some uncertainty. It is very difficult to establish reliable distributions for these parameters from sparse well data. In fact, it is hard to get reliable distributions of these parameters from well exposed outcrops. The inference of these distributions remains the single biggest limitation of object-based modeling.

Algorithms for Object Placement

The objects must be placed so that they appear realistic and honor the available data. This is not trivial. Simplistic algorithms can lead to artifacts near conditioning data. The algorithms for conditioning include analytical algorithms that force data reproduction, two-step positioning where conditioning data are reproduced first and then the remainder of the domain is filled, and iterative simulated annealing-type algorithms.

The basic algorithm underlying object-based facies modeling is the Boolean placement of objects. The objects may accumulate or build up from the stratigraphic base. Alternatively, objects may be embedded within a matrix facies according to some stochastic process and erosional rules (stratigraphically higher objects erode older lower objects) imposed afterwards. Unconditional simulation is straightforward; objects are placed randomly until the global proportions of the different facies are reproduced. Reproduction of dense well conditioning was considered difficult; however, many algorithms have been developed to address this challenge.

Direct conditioning algorithms modify the size, shape, or position of objects to honor local conditioning data by construction. The procedure may be implemented for a variety of geological shapes, but fluvial channels are the most common. Figure 7.1 gives an illustration of a 1-D process that models that centerline of a channel object. The departure from the main direction line is modeled with geostatistical (typically Gaussian) simulation. The departure from the centerline can be used for conditioning to local well data. Figure 7.2 gives two simplistic examples.

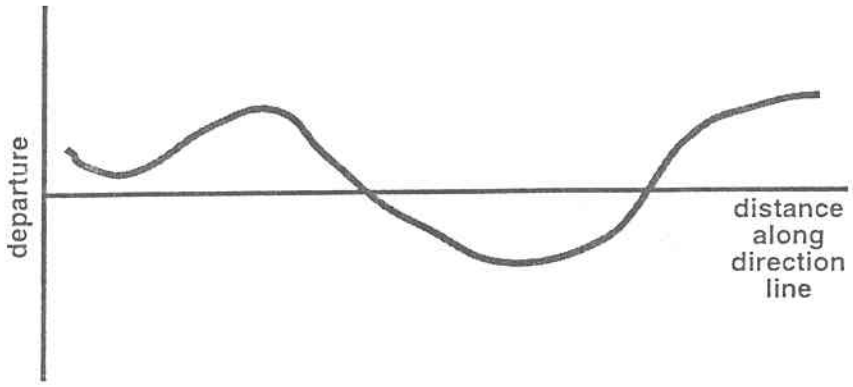


Figure 7.1: Illustration of 1-D variable that models that centerline of a channel object. The departure from the main direction line is modeled analytically or with geostatistical (typically Gaussian) simulation.

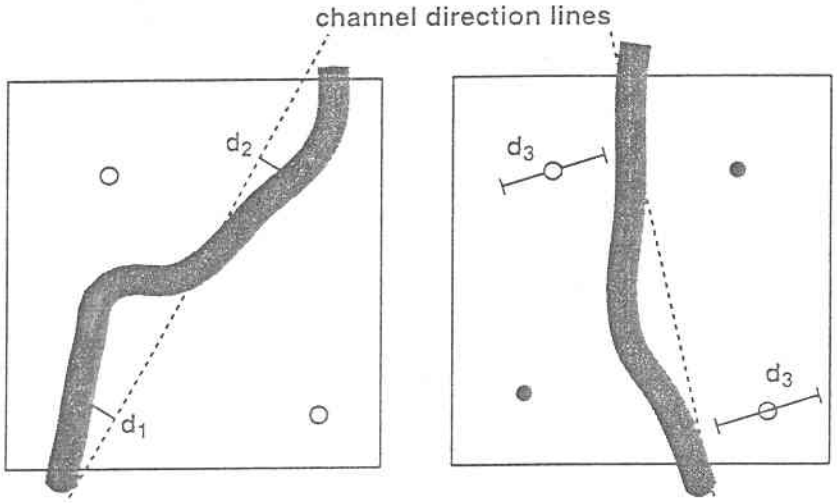


Figure 7.2: Two examples of direct conditioning of channel centerlines to honor local data in object-based modeling. In the left case, the distances d_1 and d_2 are used to force the channel to go through the two wells. In the right case, the channel is forced away from the wells where it does not exist.

Often, direct conditioning algorithms are applied in two steps. The first step is to add objects to "cover" local facies intersections, that is, match all conditioning data that are *not* background facies. The second step is to add additional objects, that do not violate known intersections of background facies, until the right proportion of each facies type is present.

A completely different approach to object placement is iterative. An initial set of objects are placed within the reservoir volume to match global proportions of the different facies. The placement may violate local well data and locally varying proportions. An objective function is calculated to measure the mismatch from the measured data. Then, the model is iteratively perturbed until the objective function is lowered sufficiently close to zero. Details of this will be discussed below.

7.2 Stochastic Shales

When the net-to-gross ratio is fairly high (greater than 80% or so) it is reasonable to model the remnant shales as objects within a matrix of sand. These remnant shales may not affect horizontal permeability, but could have a significant effect on vertical permeability, coning, cusping, and performance predictions of horizontal wells. Concretions or stylolites may also be modeled with the same procedure.

The required input to stochastic shale modeling is a geometric specification of the shale geometry, size and orientation parameters, well data, and perhaps information on locally varying shale proportions. The shape is often taken as ellipsoidal. The thickness distribution may be inferred from well locations. The length-to-thickness ratio must be established from geological analogue information or matching to some other data type such as well test or production data. Some orientation information may be available through knowledge of the depositional system. It is unlikely there is any information related to locally varying proportions if the fraction of shale is small (say, less than 20%).

Figure 7.3 shows three images of stochastic shales. Each cross section has the same percentage of shale (15%) and the same thickness distribution. The lateral extent of the ellipsoidal shales was changed systematically.

Conditional simulations are more interesting. Figure 7.4 shows three images that honor the succession of sand and shale at a well location. In this case, a two-step procedure was followed whereby shales were first added to honor the well intersections and then other shales were added (that do not violate the sand intersections at the well locations) to make up the proportion of 15% shales.

SIS models would look similar to these shale models, but with more noise. The object parameterization chosen for the shales is simplistic and linear. A more useful application of object-based techniques is for sinuous fluvial channels.

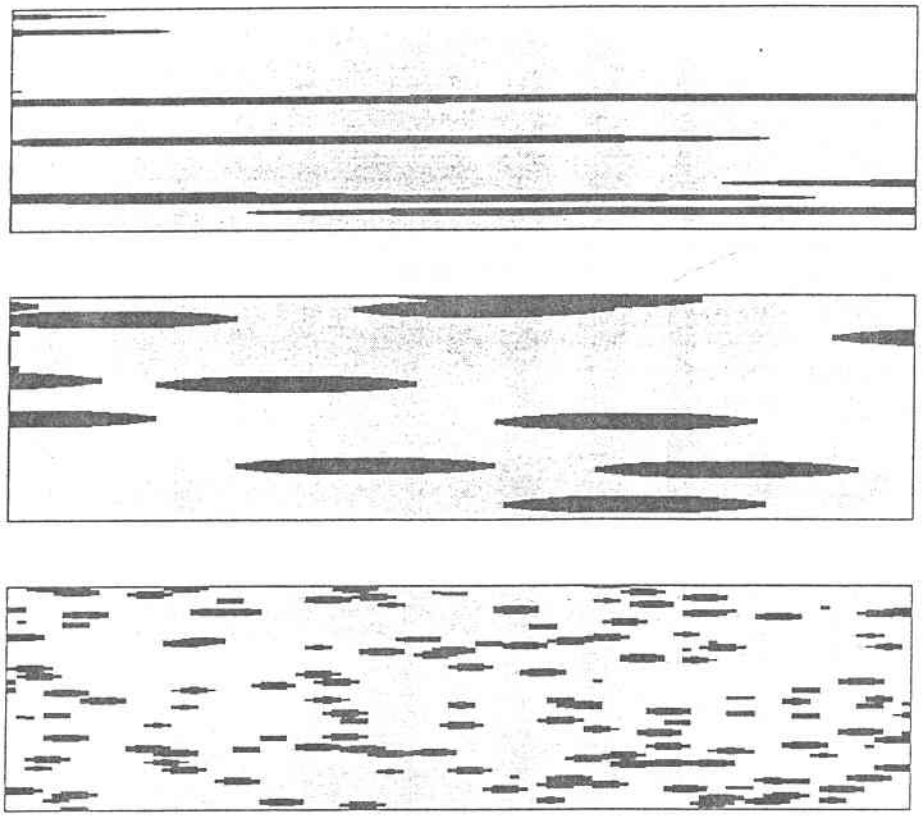


Figure 7.3: Illustration of three object-based stochastic shale models. The thickness distribution is the same for all three, but the lateral dimensions is clearly different.

7.3 Fluvial Modeling

Historically, object-based modeling has been devoted to abandoned sand-filled channel forms with a matrix of floodplain shales [66]. More complex fluvial sand geometries such as channel bar forms, laterally extensive sand regions due to lateral accretion, and more discontinuous point-bar sand bodies are considered using the same scheme with more sophisticated sedimentological models. Also, associated facies such as levee sands, crevasse sands, and cemented zones within the channel sands are modeled at the same time.

Figure 7.5 illustrates the hierarchical approach to modeling. The reservoir is first separated into a number of *major reservoir layers*. Figure 7.5 shows three major reservoir layers. Each of the three layers is modeled independently and then reassembled back into a complete reservoir model. Step (b) indicates a stratigraphic transformation transforming the layer from a vertically irregular volume into a regular 3-D volume. In step (c), a number of

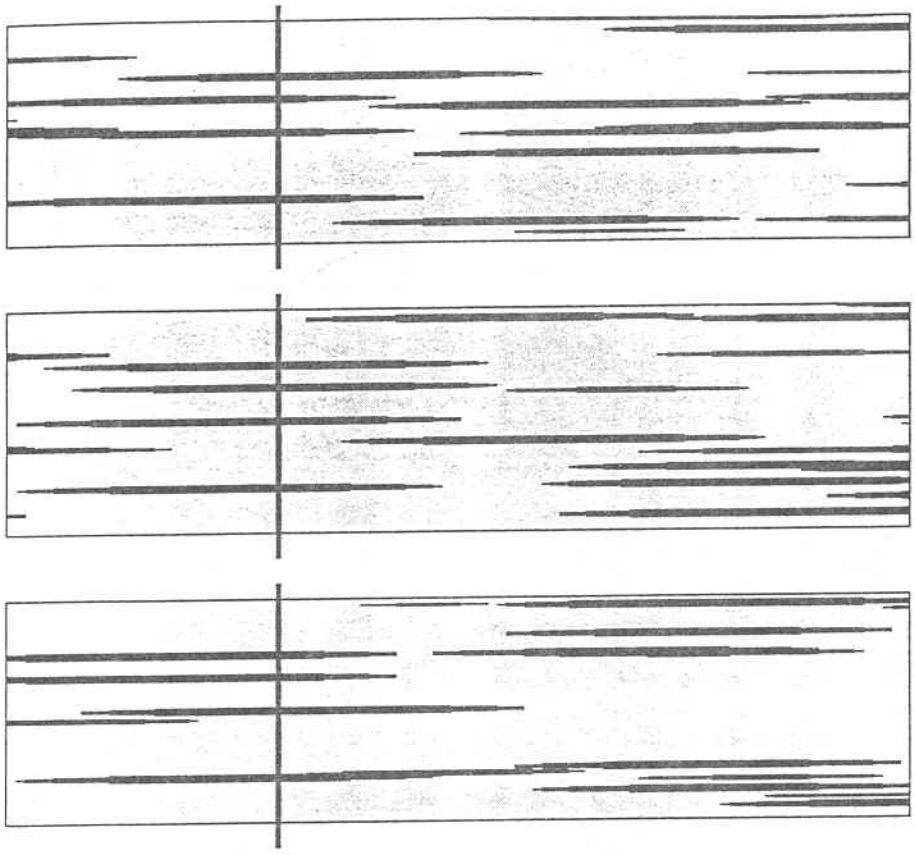


Figure 7.4: Three object-based stochastic shale models that reproduce the succession of sand and shale at a vertical well location.

channel complexes are stochastically simulated. Then each of these channel complexes is transformed from an irregular shape into a regular 3-D volume and a number of channels are stochastically simulated, as illustrated by steps (d) and (e). Each channel is in turn transformed from an irregular volume into a regular 3-D volume and filled by cell-based geostatistical algorithms with porosity and permeability, as illustrated in steps (f) and (g). In steps (h) and (i), the property-filled 3-D boxes representing channels are back transformed to their original geometry and position in the corresponding regular channel complex 3-D volume. When all channel complexes are filled with property-filled channels, they are back transformed to their original geometries and positions in the corresponding regular reservoir layer volume, as shown in steps (j) and (k). The last steps consist of back-transforming the reservoir layers and restoring them to their original reservoir positions, as sketched in steps (l) and (m).

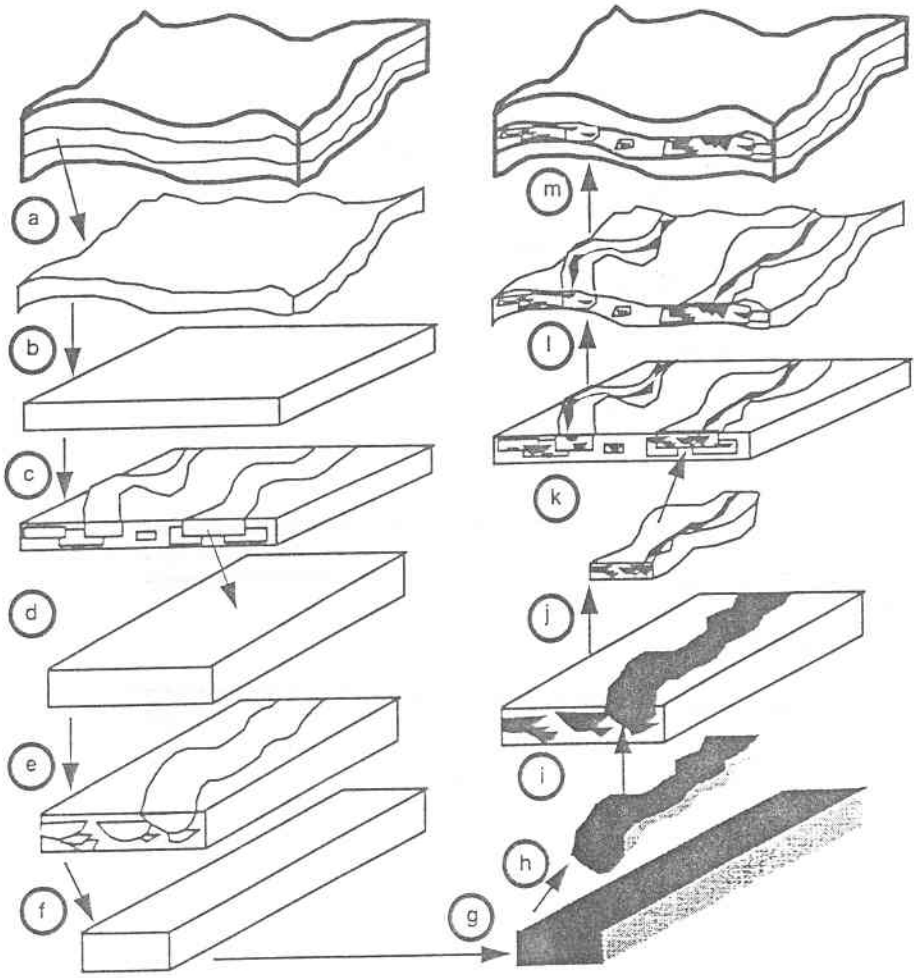


Figure 7.5: The conceptual approach to hierarchical object-based modeling: (a) each reservoir layer is extracted, (b) stratigraphic reservoir coordinates are calculated, (c) each channel complex is extracted from the layer, (d) a local stratigraphic channel complex coordinate system is established, (e) each channel is extracted from the channel complex, (f) a local stratigraphic channel coordinate system is established, (g) the petrophysical properties are modeled within the appropriate channel coordinate system, (h) the channel coordinates are restored, (i) the channel is placed back in the channel complex, (j) the channel complex coordinates are restored, (k) the channel complex is placed back in the layer, (l) the layer coordinates are restored, and (m) the layer is placed back in the reservoir.

Coordinate Transformations

One view of object-based modeling is that the coordinate system is adapted to the appropriate principal directions of continuity. The direction of continuity depends on the scale of observation and the specific geological feature being modeled. The different coordinate transformations have all been described in Chapter 3; here, they are applied successively and at different scales, for example:

- Each major stratigraphic *layer* or *sequence* is modeled independently. A layer-specific relative stratigraphic coordinate is defined from the existing top, existing base, restored top, and restored base.
- An areal translation and rotation is carried out to obtain an areal coordinate system aligned with the principal paleoslope direction, that is, the direction with the greatest continuity on average.
- Fluvial channels often cluster together in *channel complexes* or *channel belts*. The channel complexes may have large-scale undulations that are straightened. The channels within the channel complexes are also straightened for modeling of petrophysical properties within them.
- The "straightened" channel complexes and channels are not equally wide; a relative horizontal coordinate makes the boundaries parallel and simplifies subsequent petrophysical property modeling.

Each transform is reversible and the geological objects can be viewed and saved in any coordinate system.

Each channel is filled with porosity and permeability using the appropriate within-channel coordinate system. The vertical coordinate system could be taken parallel to the top channel surface or made proportional between the top and base surfaces. The base of each channel may have shale clasts or basal deposits that lower porosity and, more importantly, permeability. Furthermore, there may be a systematic trend in porosity from the base to the top of each channel. These features are handled easily, since the position of each channel base is known, by a trend in the porosity modeling program.

Fluvial Facies Objects

Figure 7.6 illustrates one possible conceptual model for fluvial facies. There are four facies types where the geometric specification of each is chosen to mimic shapes idealized from observation.

1. The first facies type is impermeable background floodplain shale, which is viewed as the matrix within which the reservoir quality or *sand* objects are embedded.
2. The second facies type is *channel sand* that fills sinuous abandoned channels. This facies is viewed as the best reservoir quality due to the

relatively high energy of deposition and consequent coarse grain size. There may be special features within the channel sands such as (1) heterogeneous channel fill, perhaps containing some fine-grained non-net material, (2) a channel lag deposit at the base, and (3) fining-upward trends within the channel fill.

3. The third facies type is *levee sand* formed along the channel margins. These sands are considered to be poorer quality than the channel fill.

4. The fourth and final facies type is *crevasse splay sand* formed during flooding when the levee is breached and sand is deposited away from the main channel. These sands are also considered to be poorer quality than the channel fill. As illustrated in Figure 7.6, crevasses often form where the channel curvature is high.

An "object" for fluvial modeling is a channel and all related levee and crevasse sands. More specifically, the object is a template of cells that would be coded as channel sand, levee sand, and crevasse sand. A combination of an analytical cross section (rectangle or half ellipse) with a 1-D function for the centerline also could be considered. The template provides significant CPU advantages. The connectivity of the resulting realizations is sensitive to the choice of an underlying grid size. The grid size must be chosen small enough to preserve the geological shapes represented by the templates.

The parameters used to define an abandoned sand-filled channel are illustrated in Figure 7.7 and Figure 7.8. The channels are defined by an orientation angle, the average departure from the channel direction, the "wavelength" or correlation length of that average departure, thickness, thickness undulation (and correlation length), width/thickness ratio, and width undulation (and correlation length). Each parameter may take a range of possible values according to a probability distribution.

Figure 7.9 shows a geometric form that could be adopted for the levee sand. The three distance parameters (A) lateral extent of the levee, (B) height above the channel datum elevation, and (C) depth below the channel datum are used to define the size. For simplicity, the geometric shape will remain fixed and only the size will vary. In general, the levee size parameters should depend on the size of the channel; a large channel has larger levees. The sizes of the left and right levee may be different.

A random walker procedure may be used to establish the crevasse geometry, see Figure 7.10. The location of a crevasse along the channel axis is chosen with probability increasing in proportion to the curvature. A number of random walkers are "released" from the location of the crevasse to establish its areal extent, see Figure 7.10. The four control parameters are (1) the *average* distance the crevasse sand reaches from the channel bank, (2) the *average* along-channel distance, (3) the irregularity of the crevasse sand or the number of random walkers used; more walkers leads to a smoother outline. The thickness of the crevasse sand decreases linearly from a maximum thickness next to the channel.

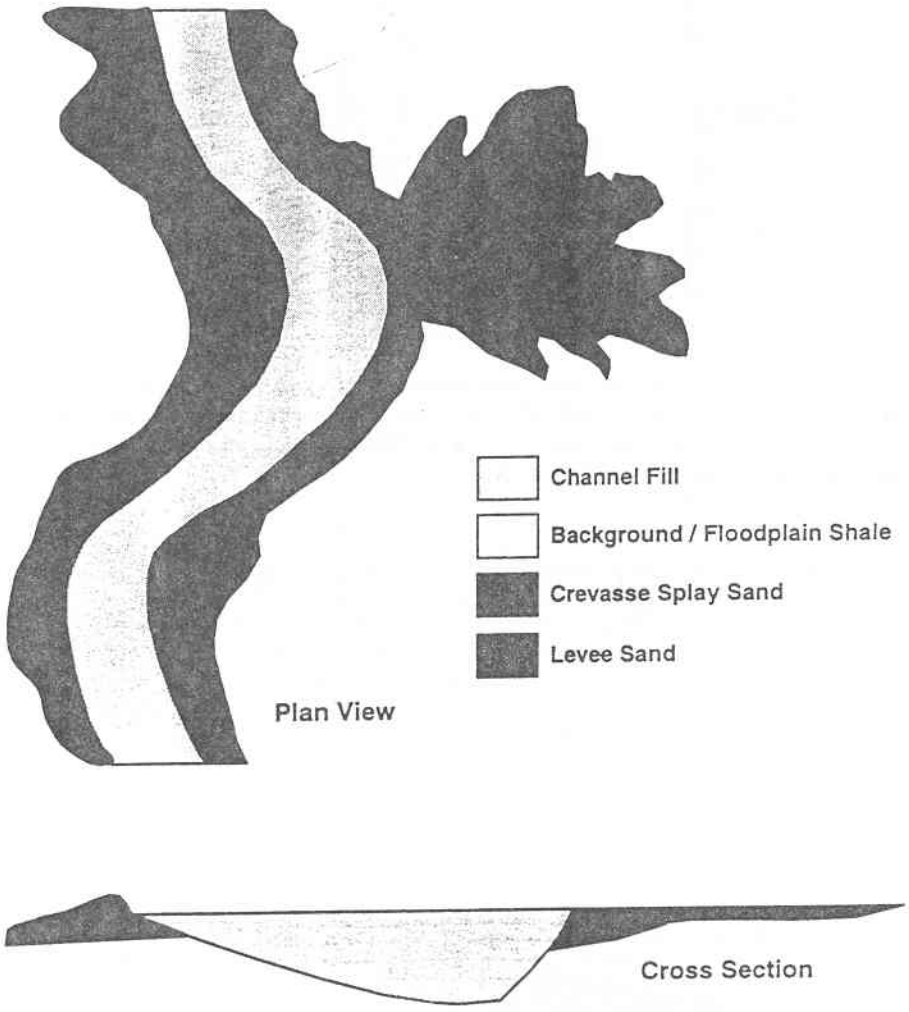


Figure 7.6: Plan and section view of conceptual model for fluvial facies: background of floodplain shales, sand-filled abandoned channel, levee border sands, and crevasse splay sands.

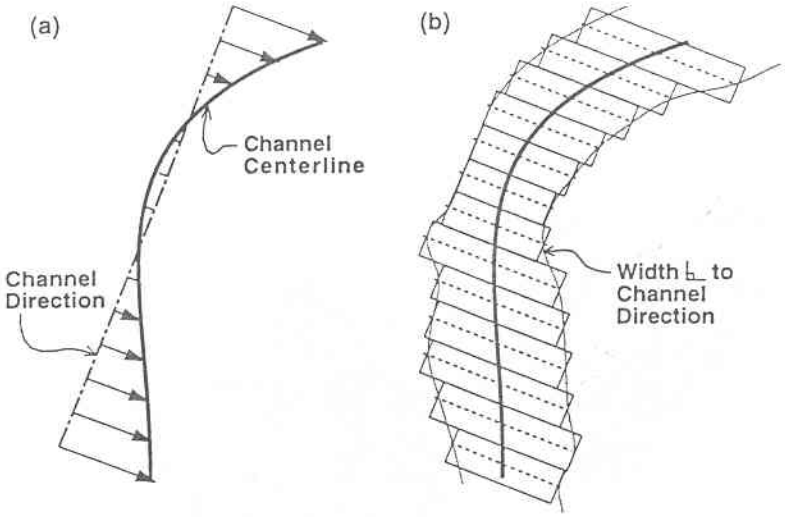


Figure 7.7: Areal view of some parameters used to define channel object: (a) angle for channel direction and deviation for actual channel centerline, and (b) variable channel width (and thickness) with "blocky" connection between channel cross section slices.

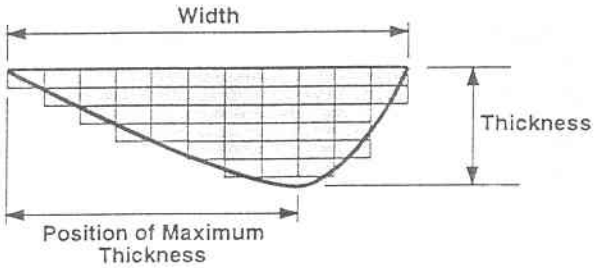


Figure 7.8: Cross section view of a channel object defined by width, thickness position of maximum thickness.

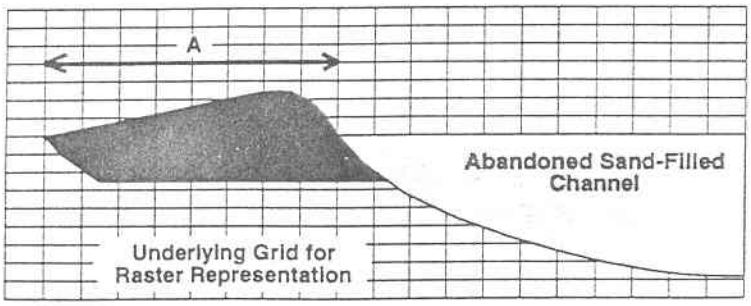


Figure 7.9: Cross section through an abandoned sand-filled channel and levee sand. Three distance parameters (A), (B), and (C) are used to define the size of the levee sand.

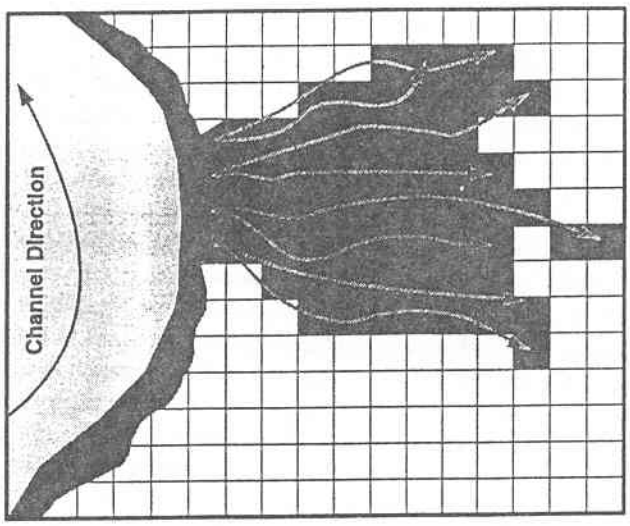


Figure 7.10: Areal view of channel and crevasse formed by breaching levee. A number of random walkers are released from breach to establish the crevasse geometry. The number, length, and lateral diffusivity control the geometry and size of the crevasse.

Geological Conditioning Data

Many of the object size parameters for object-based models are difficult to infer from available well and seismic data. Wells provide information on the thickness distributions and proportions of the different facies; however, it is often necessary to adapt measurements from analog outcrops, modern depositional systems, and similar densely drilled fields.

The proportion of each facies type (channel sand, floodplain shale, crevasse splay sand, and levee sand) may be specified by a vertical proportion curve, areal proportion maps, and reference global proportions. The reference global proportions are denoted as $P_g^k, k = 1, ..., K$, where K is the number of facies. A vertical proportion curve specifies the proportion of a facies k as a function of vertical elevation or time and is denoted as $P_v^k(z)$ where $z \in (0, 1]$. Areal proportion maps specify the facies proportion as a function of areal location (x, y) and are denoted as $P_a^k(x, y)$. These three types of proportions can be obtained through a combination of well and seismic data.

The vertical proportion curve and areal proportion map may have implicit global proportions that may be either inconsistent with each other or with the reference global proportions. The proportions can be scaled.

Two characteristics of the geological or geometrical parameters are that (1) each parameter may take a range of possible values according to a probability distribution, and (2) the range of values changes with the stratigraphic position/time z. Therefore, the parameters that define each channel are spec-

ified with a series of conditional distributions for a discrete set of z values between the stratigraphic base and the stratigraphic top.

A collection of closely spaced and genetically related channels is referred to as a channel complex. As described above, the orientation, shape, and size of a channel complex serve as a basic modeling unit. The following input distributions are required to specify the channel complex geometry, (1) angles α, (2) departure from channel complex direction and the correlation length of this departure along the axis of the channel complex, (3) thickness, (4) width-to-thickness ratio, and (5) net-to-gross ratio within the channel complex.

A greater number of parameters are required to specify the channel geometry because the thickness and width of each channel is not constant: (1) average thickness, (2) thickness undulation and correlation length of thickness undulation, (3) width-to-thickness ratio, (4) width undulation and correlation length of width undulation, and (5) channel base roughness. There are inevitable problems to reliably infer these parameters from available data.

The facies are known along each well. The data may come from any arbitrary number of wells. The well data in original coordinate space are translated to the appropriate coordinates at each stage of modeling.

Simulation Procedure

All of the geologic inputs described above must be specified prior to modeling. Many of these parameters will be difficult to infer from available data. Some parameters, such as width-to-thickness ratios, may be kept constant at some realistic value. Sensitivity studies can be considered on other parameters to judge the importance and to assess the visual acceptability of the resulting realizations.

The simulation procedure is sequential. First, the channel complex distribution is established. Second, the distribution of channels within each channel complex. Third, the porosity is assigned using appropriate channel coordinates. Finally, the permeability is assigned conditional to the facies and porosity. Conventional kriging-based algorithms can be used for porosity and permeability modeling.

Conditioning to the facies succession along wells and to facies proportion data should be accomplished directly if possible. An iterative procedure can be used to enforce the facies at wells and proportion curves and maps (see previous discussion on coordinate transformation).

7.4 Non-Fluvial Depositional Systems

Object-based facies modeling is applicable to many different depositional settings. The main limitation is coming up with a suitable parameterization for the geological objects. Deltaic or deepwater lobes are one object that could be defined. Parameters for a "simple" lobe geometry are illustrated in Figure 7.11. This set of parameters provides a balance between over

simplistic geometry (too few parameters) and flexibility with the associated difficult inference (too many parameters). The seven parameters illustrated in Figure 7.11:

1. *sw*= starting width; typically set to the final channel width. The user, however, could set this arbitrarily.
2. *ll*= lobe length; total length of the lobe from the channel to terminus.
3. *rm*= relative position of maximum width; the lobe reaches maximum width a distance of $rm \cdot ll$ from the start of the lobe.
4. *lwlr*= lobe width/length ratio; maximum width of the lobe is given by $lwlr \cdot ll$ at location specified by relative position rm.
5. *st*= start thickness; start thickness of the lobe next to channel (could be associated to the channel thickness at that point). This is the thickness at the lobe centerline.
6. *ft*= final thickness; thickness of the lobe at the terminus (at the lobe centerline) relative to starting thickness.
7. *ct*= cross section thickness correction; amount that the thickness reduces from the centerline to edge of lobe.

The parameterization could easily get more elaborate at the cost of additional parameters that must be inferred from limited observational data. One natural extension is the possibility to keep the base flat for some distance before tapering to zero thickness at the lobe terminus.

In plan view, the constraints on the geometry include the following: (1) the width is equal to starting width at transition point from channel - $y = w$ at $x = 0$, (2) the width is a maximum at relative position l - $y = W$ at $x = l$, (3) the width is zero at maximum lobe length L - $y = 0$ at $x = L$, (4) the tangent to the lobe shape has a zero slope at $x = l$, and (5) the tangent to the lobe shape has an infinite slope at $x = L$. The following equation satisfies these constraints:

$$y = \begin{cases} w + 4 \cdot (W - w) \cdot \left[\frac{x}{2l}\left(1 - \frac{x}{2l}\right)\right], & 0 \leq x \leq l \\ W \cdot \sqrt{1 - \left(\frac{x-l}{L-l}\right)^2}, & l \leq x \leq L \end{cases} \tag{7.1}$$

y is the distance from the centerline, x is the distance along the centerline, $w = sw/2.0$, $l = rm \cdot ll$, $L = ll$, and $W = (ll \cdot lwlr)/2$. The shape is based on the "$p(1-p)$" shape for the first part (closest to the connection with the channel) and an elliptical shape for the second part. The function and the first derivative are both continuous at all locations around the lobe outline.

Figure 7.12(a) shows a series of lobe shapes for different parameters. A combination of the width/length ratio and relative position parameter provides flexibility in the lobe shape. There are some natural extensions to this

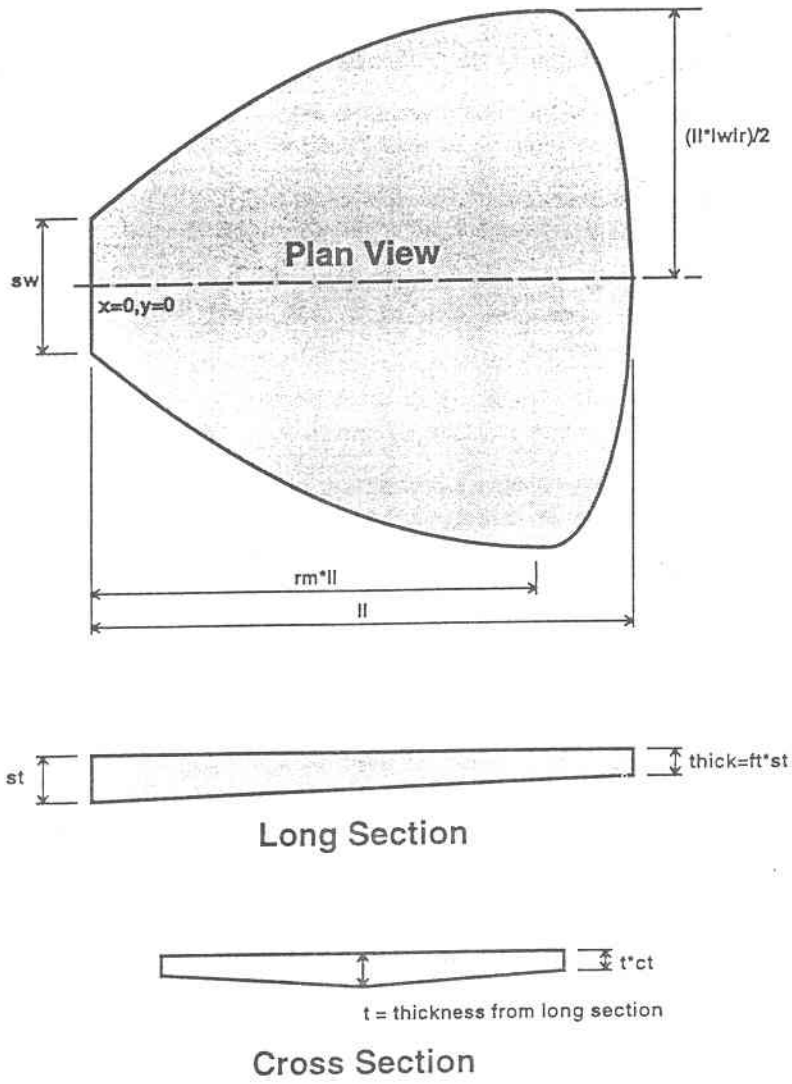


Figure 7.11: Parameters needed to describe the 3-D lobe geometry.

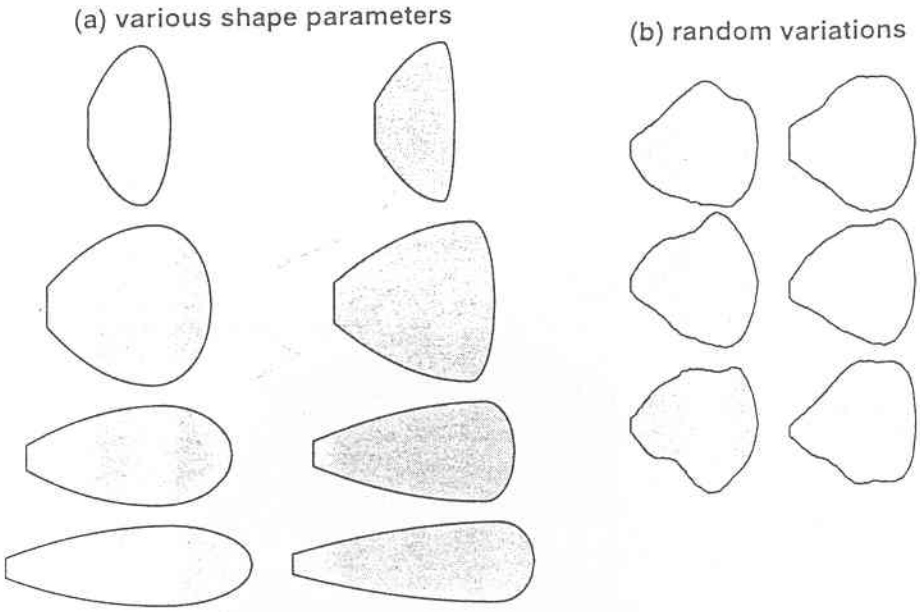


Figure 7.12: (a) example lobe geometries for various parameters. From top to bottom the width/length ratio decreases from 2.0:1 to 1:1 to 0.5:1 to 0.33:1. The relative position of the maximum width is 0.65 on the left and 0.85 on the right. (b) the basic geometry $W = 0.5, w = 0.01, L = 1.0$, and $l = 0.6$ with six different Gaussian simulations added.

areal shape including (1) addition of stochastic variations to the lobe shape for more realism (Figure 7.12(b) shows six examples) and (2) consideration of asymmetric lobe geometries, that is, the W, w, and l parameters could be different on the "top/bottom" of the lobe.

The lobes may be positioned at the end of channels. Figure 7.13 illustrates this relationship together with the channel and lobe template. The final configuration of channels and lobes in any model will ultimately be determined by the simulation algorithm and conditioning data. Areally varying facies proportions and well data could lead to lobes in preferred locations.

7.5 Work Flow

Following are the main workflows related to object-based facies modeling. Figure 7.14 illustrates the work flow for fluvial modeling. The main operation is the step to position the facies objects in a manner that reproduces all available data. Figure 7.15 shows more details of the important object-positioning work flow. In practice, there are a number of heuristic schemes to ensure that local conditioning data are reproduced quickly and exactly at the well locations.

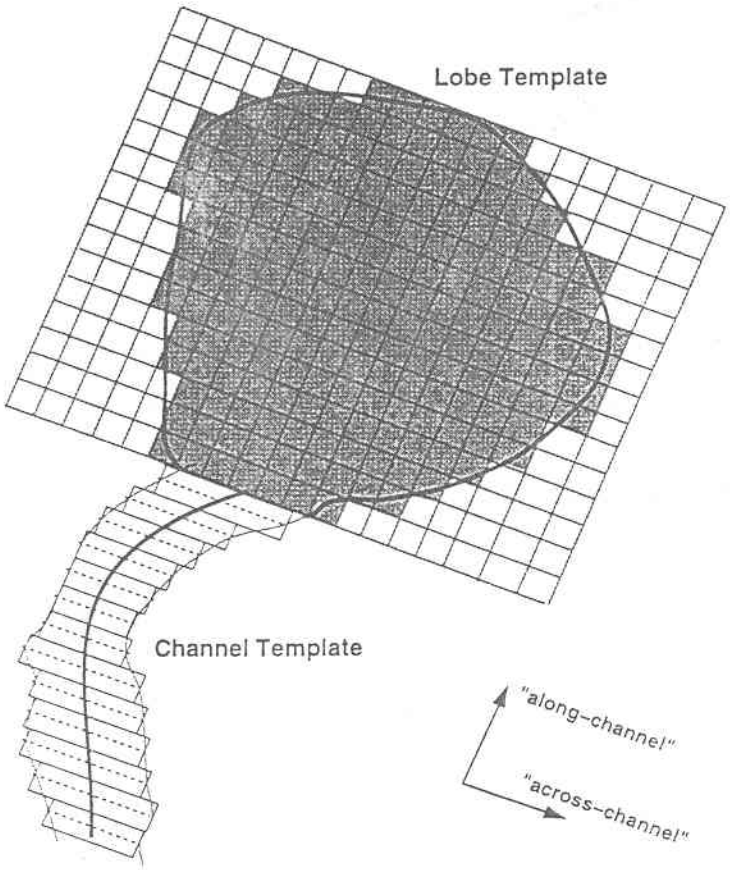


Figure 7.13: Channels and lobes are represented by templates. In the case of lobes, the gray-shaded area is kept in memory together with the depth under each cell.

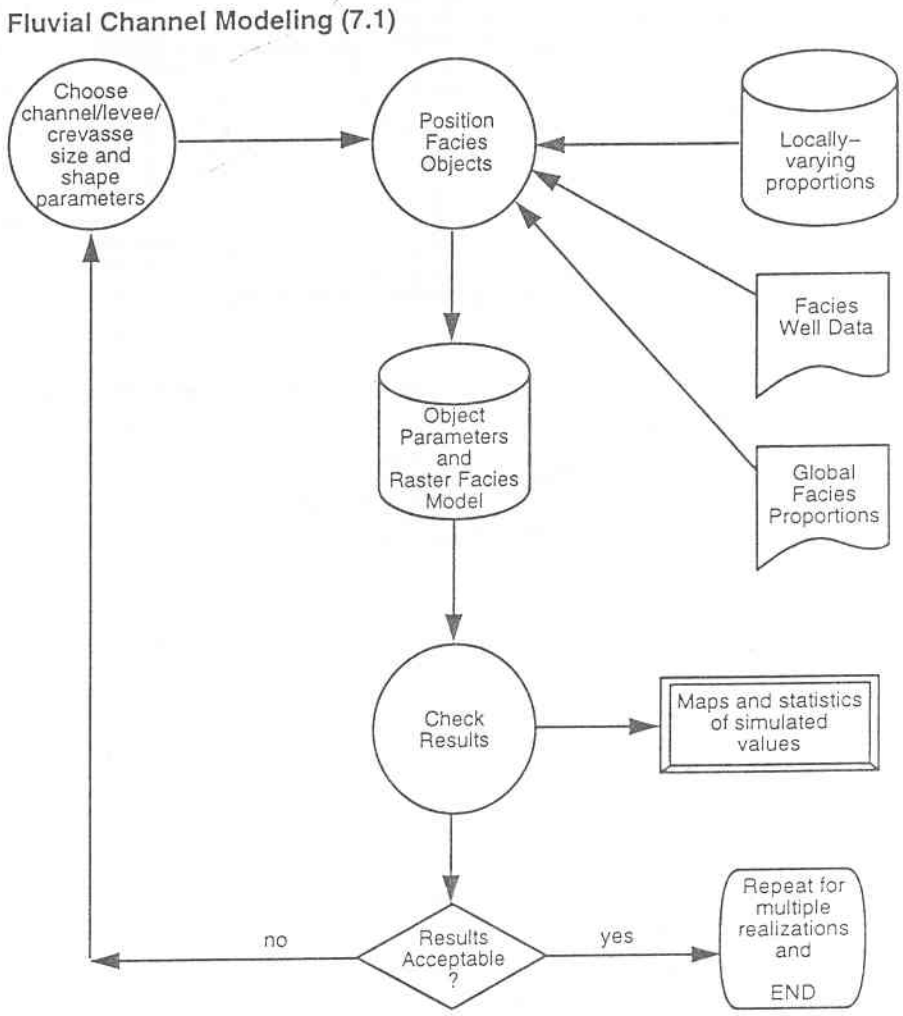


Figure 7.14: Work Flow 7.1: fluvial modeling.

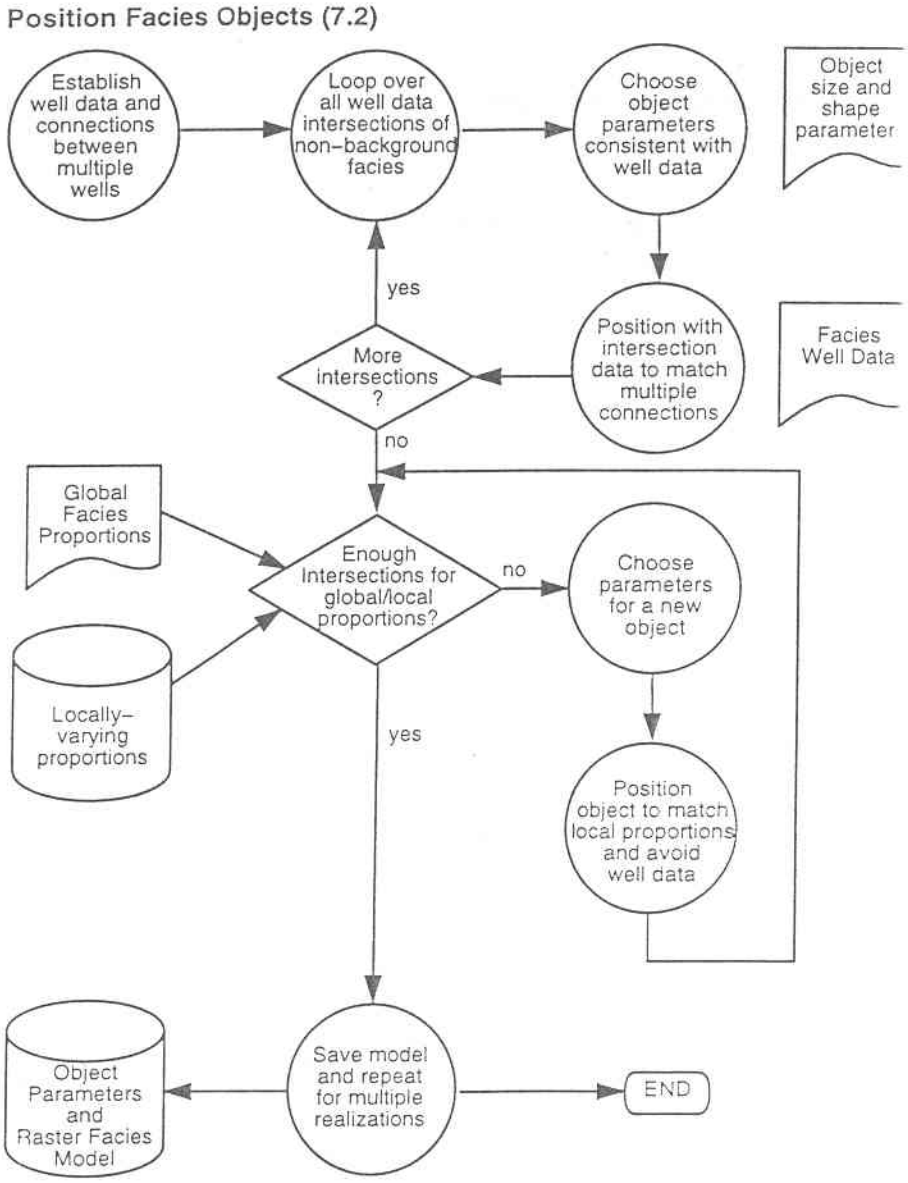


Figure 7.15: Work Flow 7.2: positioning of facies objects.

Chapter 8

Porosity and Permeability Modeling

Petrophysical properties such as porosity and permeability are modeled within each facies and reservoir layer. Shale and non-net facies may be assigned arbitrary low values; however, the petrophysical properties within most facies must be assigned to reproduce the representative histogram, variogram, and correlation with related secondary variables. The Gaussian and indicator techniques introduced in Chapter 5 are used for this purpose. This chapter discusses practical considerations and implementation details.

Section 8.1 describes the characteristics of porosity and permeability and the considerations for modeling. Porosity is relatively straightforward since it averages linearly and has low variability. Permeability is more difficult. It is more variable, has a highly skewed histogram, does not average linearly, and must reproduce an imperfect relationship with porosity.

Section 8.2 describes how sequential Gaussian simulation can be used to create porosity models. Porosity is simulated within each facies type, defined within the correct layer-specific coordinate system. This Gaussian technique is used extensively because of its simplicity of application and flexibility to create realistic heterogeneities.

Section 8.3 describes basic methods to get permeability with porosity/permeability transforms such as regression, conditional expectations, and simple Monte Carlo simulation from porosity/permeability cross plots. These simplistic techniques only work when (1) there is excellent correlation between porosity and permeability, or (2) the permeability histogram has very low variance. In general, some technique that reproduces the variogram of permeability is required.

Section 8.4 presents Gaussian techniques to cosimulate permeability using porosity as a secondary variable. Gaussian cosimulation is easy to apply and reproduces the basic characteristics of permeability variation. Indicator techniques must be considered when permeability shows complex patterns of

spatial variation such as exceptional continuity of extreme values. Section 8.5 gives details of how indicator techniques are applied to permeabiltiy. The final section, 8.6, presents work flows for porosity and permeability modeling.

8.1 Background

Variables

Porosity and permeability are the main variables needed for reservoir characterization. Other variables such as impedance, velocity, and residual water saturation may also be of interest. In the following we refer to these variables as petrophysical properties, reservoir properties, or just properties.

Porosity is the volumetric concentration of pore space. It is a well-behaved variable since it typically has small variability and averages according to a simple arithmetic average. There are different definitions for porosity. It is common to model the "effective" porosity. The "total" porosity includes volume that could never be accessed through practical recovery mechanisms.

Permeability and some other variables are not intrinsic rock properties; they depend on boundary conditions outside of the volume of measurement or specification. These variables can vary over several orders of magnitude. Effective properties are not simple averages of smaller-scale values. There are some considerations in modeling permeability: (1) the data should be corrected to reservoir fluid and pressure conditions; these conditions are different from the laboratory where the measurements were made and a prior correction using well test-derived permeability data must be considered, (2) horizontal to vertical anisotropy is an important factor; the principal directional components of permeability can be modeled or global vertical to horizontal anisotropy ratios can be applied on a by-facies basis.

As discussed in Section 5.6, vertical and areal trends in the mean and direction of continuity are common. Facies modeling may have removed some of the trend. Deterministic trends must be modeled and some decision must be made about how to account for the trend, that is, a cokriging or a mean plus residual approach. Details are presented in Section 5.6. The "variable" to model may be residuals instead of original units.

Scale Differences

Petrophysical properties must be assigned to every cell in the geological model. As discussed in Chapter 1, the volume difference between different core measurements, well log derived properties, and the geological modeling cell size is not considered. Therefore, we are essentially assigning *point* scale values on a regular grid in stratigraphic coordinate space. These values are associated to the entire geological modeling cell for calculation of fluid volumes and properties for flow simulation. Reservoir modelers will increasingly consider the vast volume difference between well data and the geological mod-

eling cells (see the "missing scale" discussion of Chapter 11); however, the conventional practice presented here has led to reservoir models with significant predictive ability.

Facies

Porosity and permeability are modeled within each facies and reservoir layer. The reservoir layers are distinct due to deposition at different times. Reservoir properties within the *same* facies and *different* layers may have similar characteristics. Nevertheless, they have to be modeled separately because of different layer-specific stratigraphic coordinate systems. The properties within *different* facies in the *same* layer often are significantly different and unrelated.

The properties within the different facies can be modeled independently when the facies are unrelated. The "cookie-cutter" approach is sometimes used: (1) complete 3-D models of the properties are constructed for each facies, and (2) the models are merged using the facies model, that is, the porosity and permeability are taken from the appropriate facies dependent model. This is wasteful. A faster approach is to use the facies model as a template during simulation of the properties. The predictions are limited to where they are needed. This does not introduce any biases or other problems.

Properties in different facies may be related to each other, that is, there is correlation in petrophysical properties across facies boundaries. There are a number of approaches to deal with this correlation:

- Ignore the correlation and model the properties independently in the different facies. The correlation across facies boundaries likely has little consequence on volumetrics and flow performance predictions. We could argue that such correlation is not a first-order effect.

- Consider a form of cokriging with the properties in different facies as secondary variables. Calculation and modeling of the necessary models of coregionalization is difficult. This approach is too awkward for practical application.

- Create the models sequentially, that is, construct the model of porosity and permeability in one facies and then use the result to assist in modeling the other facies. This is similiar to the standard practice of using a *logic matrix* in modeling mineral deposits.

A logic matrix specifies whether to use data from another facies, i, when assigning the current facies, j. Figure 8.1 gives an example. Note that this matrix may not be symmetric. In the example of Figure 8.1, measurements from crevasse sand are to be used in the simulation of channel facies, but measurements from channel sand are not to be used to predict crevasse. The rationale for this may be due to geological considerations or data acquisition considerations. Once data from another facies are used, they are considered

Data in Facies: \ Modeling Cell in Facies:	shale	channel	crev.	conc.
shale	Yes	No	No	No
channel	No	Yes	No	No
crev.	No	Yes	Yes	No
conc.	No	No	No	Yes

Figure 8.1: Illustration of a logic matrix of using data from different facies (the rows) to simulate the petrophysical properties in other facies (the columns).

the same as data from the same facies as the cell being modeled. One could imagine some kind of weighting scheme.

Secondary Data

In many cases there are secondary data to be considered: seismic impedance is negatively correlated with porosity, permeability is positively correlated with porosity, and irreducible water saturation is correlated with permeability. The properties we are modeling often are correlated. Seismic or inversion of production data provides an external source of information.

Typically modeling proceeds sequentially. The secondary data available at each stage may be of multiple types; they are then either (1) redundant or (2) one data type brings the most information. Therefore, most often only one secondary variable is considered for each property. Of course, this is not a theoretical limitation; it is merely practical. Typically, porosity is modeled first using seismic as the secondary variable and then permeability is modeled using the previous simulation of porosity as a secondary variable. Seismic data is indirectly used for permeability through the porosity simulation.

Due to geological complications and inherent limitations in seismic data acquisition, seismic inexactly measures average porosity. Seismic-derived porosity may be correlated with the true porosity with a correlation of 0.5 to 0.7. The specific seismic attribute and degree of correlation must be calibrated for each reservoir. We account for this precision. Seismic-derived porosity represents a volume significantly larger than the typical geological modeling cell. The areal resolution is often comparable. The vertical resolution; however, is 10 to 100 times the resolution of the geological modeling

cells. Current geostatistical models are built at a vertical resolution of 1/3 ft and current seismic data informs a 30/100 ft vertical average. We often account for this scale difference by considering the seismic data in block cokriging or as a local mean (Chapter 5).

Number of Realizations

A single realization of porosity and permeability should be constructed for every facies realization. The facies geometry is typically much more consequential for flow. Therefore, there may be no need to generate too many realizations of porosity and permeability. A few realizations may be generated for validation and the uncertainty from the facies realizations accounted for in subsequent ranking and flow simulation.

Chapter 10 gives more details related to uncertainty management and the number of realizations to generate.

8.2 Gaussian Techniques for Porosity

There are many methods for simulating the spatial distribution of a regionalized variable. The Gaussian formalism, however, is the simplest way to create a realization that honors local data, a histogram, and a variogram. Major abrupt discontinuities in the reservoir are captured by the structural and facies model. Considering a stationary histogram and variogram to quantify the properties within reasonably homogeneous facies is deemed adequate in most cases.

The building blocks of sequential Gaussian simulation (SGS) have been discussed. Section 2.5 discussed the calculation of a representative histogram. Section 5.6 discussed construction of locally varying mean models. Chapter 4 covered all the details of variogram calculation. Section 5.2 covered SGS. There remain some implementation details, particularly with respect to seismic data, that should be discussed.

The porosity data or residual porosity values (obtained by subtracting locally varying mean porosity values) must be transformed to a normal distribution. A variogram must be established. Sequential Gaussian simulation consists of visiting each grid node in a random order and:

1. Finding nearby data and previously simulated grid nodes,
2. Constructing the conditional distribution by kriging, that is, calculating a mean and estimation variance by simple kriging, and
3. Drawing a simulated value from the conditional distribution (normal with mean and variance from 2).

This entire procedure is repeated with different random number seeds to generate multiple realizations.

Seismic Data in SGS for Porosity

Seismic data is areally exensive over the reservoir and is sensitive to porosity variations. The large-scale information provided by seismic will be accounted for first in the facies model. The only additional information to gain is porosity variations within the facies, which may be minor compared to the differences between facies. The calibration procedure will reveal whether seismic should be used for SGS of porosity.

Calibration of Seismic Data to Porosity

Calibration is done on a by-facies basis. The calibration depends on modeling procedures adopted later: block (co)kriging requires the vertical average of porosity to be calibrated against seismic, collocated cokriging requires the small-scale data to be calibrated against seismic.

There are many seismic attributes that could be considered. In general, the most reliable attribute is acoustic impedance. The seismic data should be processed to provide a single seismic attribute for use as a secondary variable for geostatistical modeling. Neural networks, discriminant analysis, rule induction, or many regression-type procedures (not covered here) can be used to arrive at a single seismic- derived porosity. The calibration can appear unrealistically good when the well data used for calibration are the same as used to arrive at the single seismic attribute. Independent well data (held back from the derivation of seismic attribute) should be used for the calibration. Then, after the calibration statistics have been extracted, all of the well data can be used for the ultimate calculation of a seismic-derived porosity.

Calibration considers paired observations of porosity (at the right scale) and seismic attribute have been assembled: $Z_S(\mathbf{u}_\alpha), Z_\phi(\mathbf{u}_\alpha), \alpha = 1, \ldots, n_c$. Both variables must be transformed to standard normal distributions (see Section 2.4). The paired values are then shown on a cross plot with the correlation coefficient. Figure 8.2 shows an example cross plot. Only the correlation coefficient is used for Gaussian techniques; the bivariate distribution is assumed to follow a bivariate Gaussian distribution with elliptical probability contours.

The correlation coefficient is a two-point measure of correlation. Any lateral shift or mismatch in the seismic data could significantly affect the correlation coefficient. In general, the correlation coefficient will be reduced if the well data are being compared to a nearby, but not collocated, seismic data value. Careful data cleaning and some repositioning may be required.

Horizontal Variogram Inference

Regardless of which variant of simulation is used, a full 3-D variogram model is needed to descibe the spatial correlation structure of porosity. The main challenge in variogram inference is the horizontal direction. The seismic data

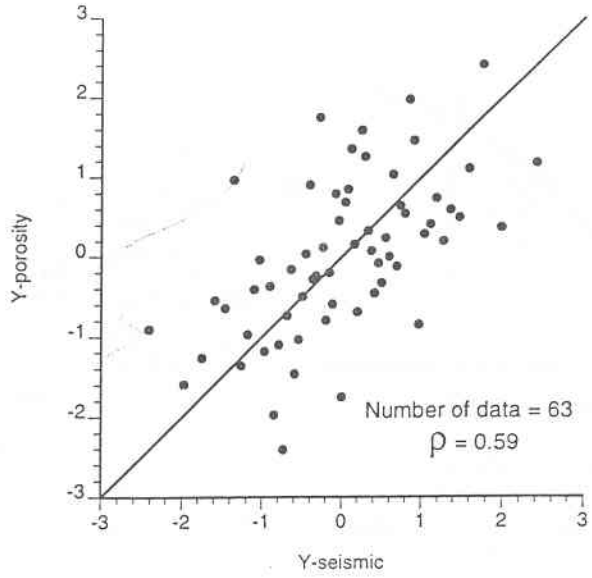


Figure 8.2: Calibration of the normal score of seismic versus the normal score of porosity. There are 63 well data showing an excellent correlation of 0.59.

provides an excellent source of information to guide the selection of horizontal range parameters. As discussed in Section 4.4, we must assess the presence of any zonal anisotropy in addition to adopting the horizontal range from the seismic data.

Locally Varying Mean

Rather than cokriging, we can convert the seismic data to the units of porosity and then use them as a local mean. The first step is to calculate the average porosity at all locations from the seismic attribute, see Figure 8.3 for two schematic illustrations. Then, simple kriging is used in the Gaussian simulation using the seismic-derived mean values. The kriged estimate is calculated as:

$$y^*(\mathbf{u}) - m(\mathbf{u}) = \sum_{\alpha=1}^{n} \lambda_\alpha \cdot [y(\mathbf{u}_\alpha) - m(\mathbf{u}_\alpha)]$$

$$y^*(\mathbf{u}) = \sum_{\alpha=1}^{n} \lambda_\alpha \cdot y(\mathbf{u}_\alpha) + \left[1 - \sum_{\alpha=1}^{n} \lambda_\alpha\right] \cdot m(\mathbf{u})$$

where $y^*(\mathbf{u})$ is the kriged estimator or the mean of the conditional distribution for Gaussian simulation, n is the number of local data, $\lambda_\alpha, \alpha = 1, \ldots, n$ are the kriging weights, $y(\mathbf{u}_\alpha)$ are the local transformed porosity data, and $m(\mathbf{u})$ is

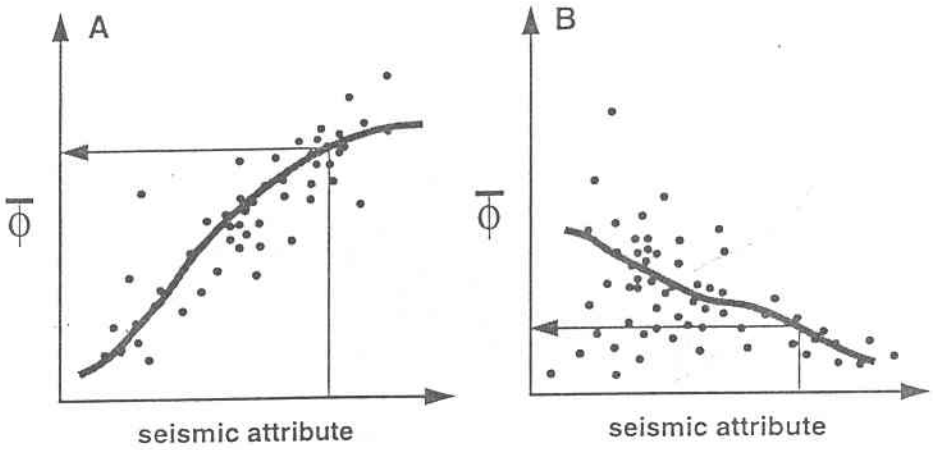


Figure 8.3: Two examples of converting a seismic attribute to average porosity. There can be positive or negative correlation. An example of converting the seismic attribute to average porosity is shown on each cross plot. The correlation can be positive (A) or negative (B).

the seismic-derived mean porosity (transformed to Gaussian units) at location $\mathbf{u}$ being considered. When there are few local data, the sum of the weight to the data is small and the seismic-derived mean porosity receives a large weight. When the weight to local data is high, the seismic-derived mean is given low weight.

Collocated Cokriging

Some type of cokriging must be applied to weight the seismic data in a way that explicitly accounts for the "softness" of the seismic data. Doyen [68] showed one of the first applications of cokriging porosity with seismic data. A large number of refinements and simplificatons have followed [10, 9, 274].

The collocated cokriging formalism was developed in Section 5.1 and further explained, in the context of SIS, in Section 6.2. The sole input required to use the seismic data is the correlation coefficient between the *hard* Gaussian transform of porosity and the *soft* Gaussian transform of seismic.

As mentioned in Section 5.1, keeping only one collocated secondary datum does not affect the estimate (close-by secondary data are typically very similar in value), but it may affect the resulting cokriging estimation variance: that variance is overestimated, sometimes significantly. Since the kriging variance defines the spread of the conditional distribution from which simulated values are drawn, this may be a problem. The collocated cokriging variance should then be reduced by a factor (assumed constant $\forall \mathbf{u}$) to be determined by trial and error.

The variance of the simulated values should be close to 1.0. We start with the variance reduction factor $varred = 1.0$ and run one simulation to

see the variance of the resulting simulated values. The variance reduction is systematically lowered *varred* < 1.0 until this variance is reduced to 1.0. There may be a highly non-linear relation between the variance reduction factor *varred* and the variance of the simulated values. There is more likely to be a problem when the seismic data is much smoother than the porosity data.

Block Cokriging

Collocated cokriging is an approximation that avoids (1) calculation of cross variograms or covariances, (2) fitting a model of coregionalization, and (3) handling the large-scale nature of seismic data. As with SIS, we could do the job right. All variograms and cross variograms could be calculated (see Chapter 3). The corresponding point-scale variograms could be determined with analytical models. A model of coregionalization could be fit (see Chapter 3). Finally, a block cokriging could be performed in SGS that would account for the scale of the seismic data and the "softness" of the calibration from seismic to porosity. The cokriging formalism is classical. Implementation of block cokriging in a conventional SGS program (such as `sgsim` in GSLIB [62]) is straightforward.

An implicit assumption behind this block cokriging formalism is that all variables average linearly. This may be the case for porosity; however, the seismic variable certainly does not average linearly. Notwithstanding this assumption, it may be better than ignoring the measurement scale.

Stochastic Inversion

An alternative to the cokriging approach to integrating seismic data is to consider a form of "stochastic inversion." The idea is to directly account for seismic data was launched in the early 1990s by workers at Elf [27]. The basic idea is to simulate acoustic impedance, process the impedance through a forward seismic model, and choose the impedance model by rejection sampling, that is, the impedance model that reproduces the original seismic data. Porosity is then linked to acoustic impedance.

The original idea proposed by Haas was to generate a number of cross sectional models, say 10 to 100, and perform forward seismic modeling on each cross section. The difference between the forward simulated seismic and the real seismic data is computed for each realization. Then, the simulated cross section with seismic traces closest to the original seismic data is retained. Each cross section is simulated conditional to all well data and previously simulated cross sections. Dubrule, Haas, and coworkers extended the original approach to work on each 1-D column one at a time [115, 76]. This speeds convergence.

There are different versions of geostatistical or stochastic inversion appearing in the literature [125, 173, 218, 245]. Many of the more recent approaches start with an initial model and loop over every cell in the entire model some

number of times (10/15). At each iteration the facies or acoustic impedence is retained according to some decision rule. The decision rule is either based on simulated annealing (see next chapter) or the result that matches the seismic data the closest.

The geostatistical inputs to stochastic inversion are conventional, that is, a global histogram and variogram. These must be supplemented by petrophysical relationships that link acoustic impedence to porosity and the wavelet that links acoustic impedence to the original seismic response.

Small Example

The following example is based on Chu et al. [37]. A 10400-ft by 10400-ft by about 50-ft thick reservoir layer of mixed carbonate siliciclastic sediments is considered. Figure 8.4 shows a gray-scale image of the 2-D grid seismic data and a cross plot between the normal score transform of porosity and the normal score transform of the seismic variable. The correlation coefficient is 0.4, which is high given the small scale of the porosity data. As shown on Figure 8.2, the correlation increases with averaging to almost 0.6. The correlation coefficient can increase of decrease depending on the direct and cross variogram structure.

There are 61 wells in the area of interest, see location map on Figure 8.5. The porosity and permeability logs at one well location are also shown in Figure 8.5; extra detail including the other well logs have been omitted for brevity. No explicit trend modeling is considered since the areal trends are captured by the seismic data. The vertical and horizontal variograms of the normal score transform of porosity are shown in Figure 8.6. The directional variograms are N-S (the one with the lower sill) and E-W (the upper one). The stratigraphic coordinate was taken proportional between the upper and lower bounding surfaces of the reservoir layer.

Sequential Gaussian simulation with collocated cosimulation using the seismic data was performed. Gray-scale images of a horizontal slice and E-W cross section are shown in Figure 8.7. The higher porosity values in the northeast corner can be seen on the porosity slice. This illustrates a simple and effective procedure for property modeling. Permeability can be modeled using porosity as a secondary variable, multiple realizations can be constructed, and the models can be scaled for flow simulation.

8.3 Porosity/Permeability Transforms

Porosity is modeled on a by-facies basis. The porosity models are used directly to assess pore volume, pore volume uncertainty, and connectivity based on porosity. Multiple realizations can be ranked and selected on the basis of hydrocarbon pore volume; see Chapter 10. However, permeability is required in addition to porosity for flow simulation.

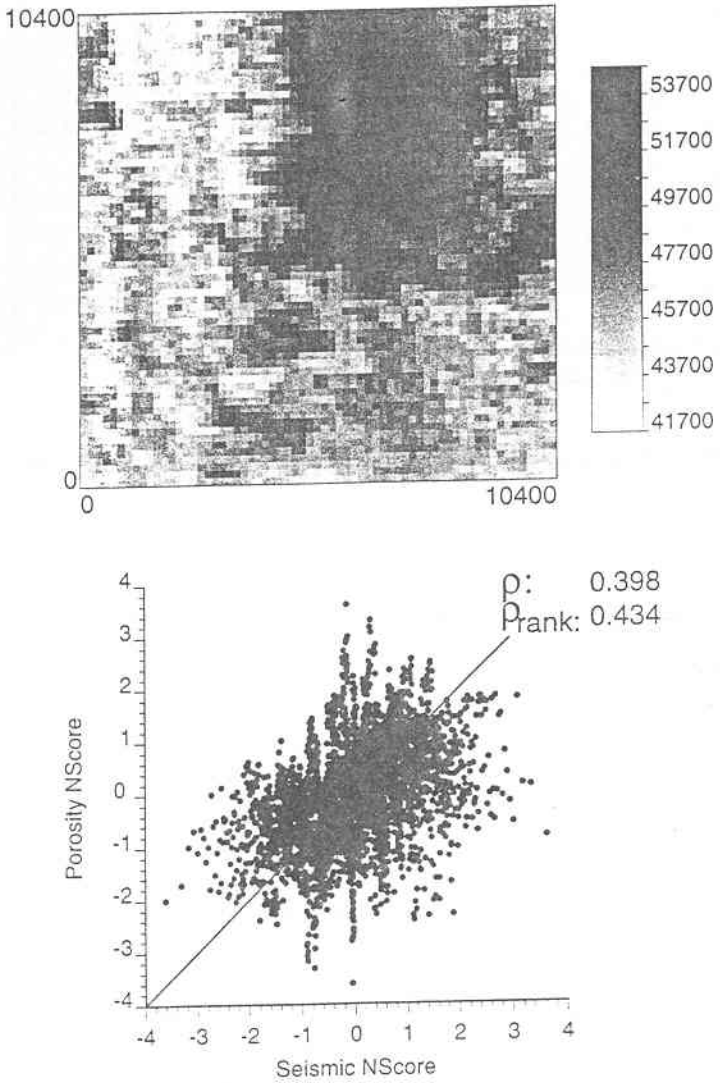


Figure 8.4: 2-D grid of seismic variable over a 10400 ft by 10400 ft area and the correlation of the seismic variable with porosity over a 1-ft vertical interval. The correlation between the normal score of the seismic and normal score of porosity is 0.4, which is very good given the scale of the porosity data. In this case, the correlation increases with porosity averaging.

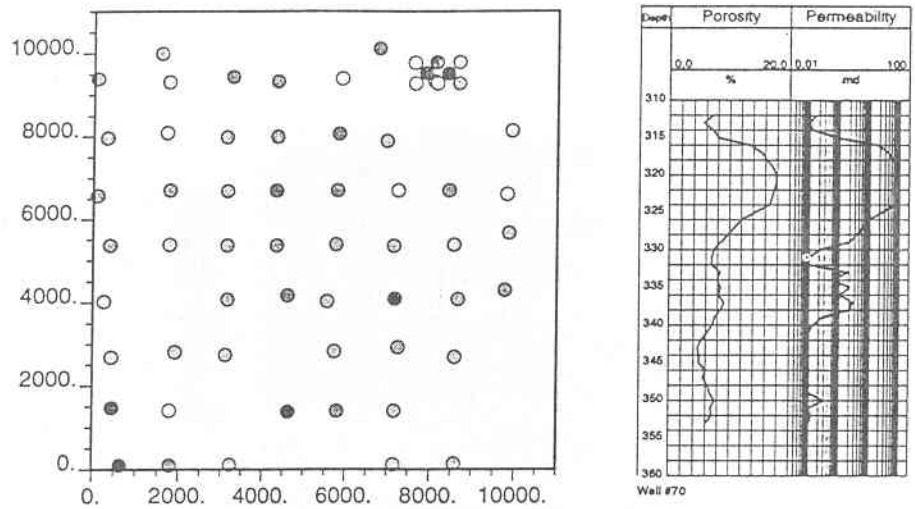


Figure 8.5: 2-D map of well locations and the porosity and permeability logs at one well location.

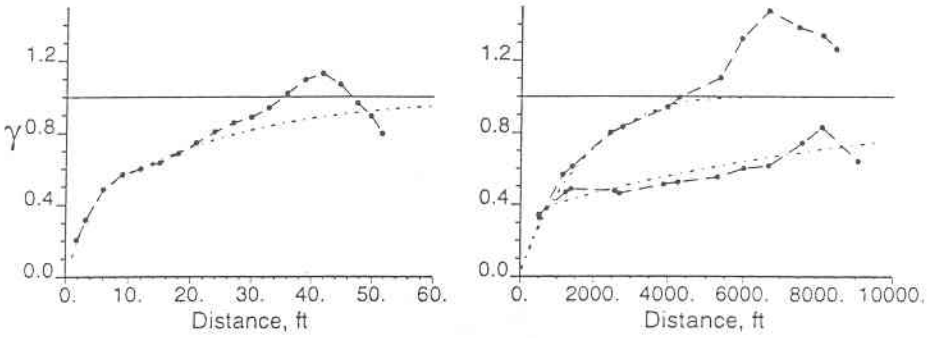


Figure 8.6: Vertical and horizontal variograms with fitted models for 3-D porosity data of Figure 8.6.

Three approaches will be developed in this chapter for permeability modelign. This section (Section 8.3) on porosity/permeability transforms, Section 8.4 on Gaussian techniques, and Section 8.5 on indicator techniques. These are in order of increasing complexity and flexibility. The porosity/permeability transforms described below are only rarely used because they fail to reproduce permeability variation patterns. Gaussian techniques are widely used. Indicator techniques are occasionally used when there are sufficient data to infer the complex required statistics.

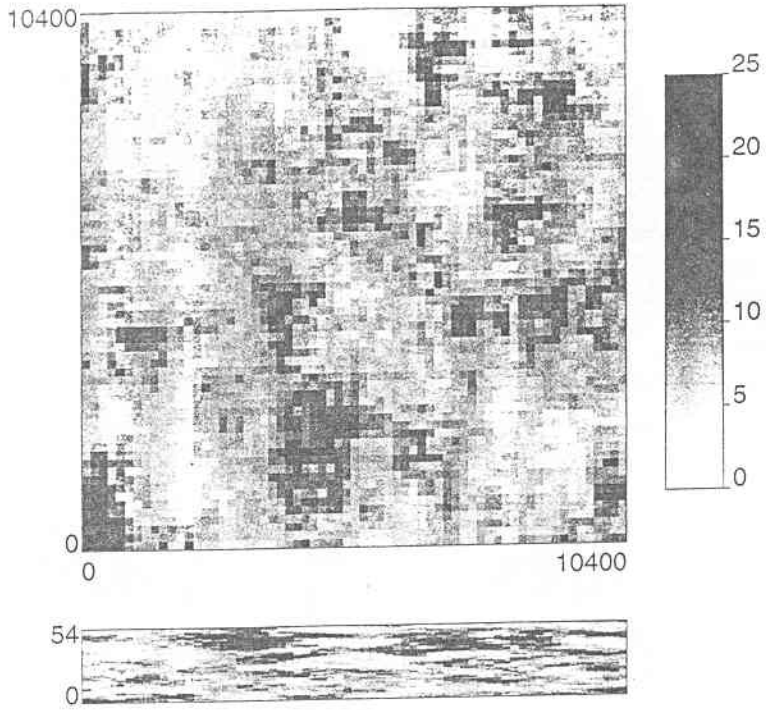


Figure 8.7: Horizontal slice and E-W cross section through porosity model generated by sequential Gaussian simulation. The seismic variable was considered as a secondary variable using collocated cokriging with a correlation coefficient of 0.4; see the seismic data on Figure 8.4.

Regression

Classical regression can be applied to develop a relationship between porosity and permeability. Typically, the logarithm of permeability is used because of the skewed and approximately lognormal character of many permeability histograms. The regression equation takes the form:

$$log(K)^* = a_0 + a_1 \cdot \phi + a_2 \cdot \phi^2 + a_3 \cdot \phi^3 \tag{8.1}$$

where $log(K)^*$ is the predicted logarithm of permeability, $a_i, i = 0, \ldots, n_R$ are regression coefficients, and ϕ is porosity. Log-linear regression is common, that is, only the first two terms are used. The second order term $a_2 \cdot \phi^2$ may be used to capture non-linear features. Only rarely are higher order terms ≥ 3 needed or recommended. The calculation of the regression coefficients $a_i, i = 0, \ldots, n_R$ is automatic in most software; details can be found in most statistics books.

Figure 8.8 gives four real examples of porosity/permeability cross plots and second-order regression curves. These equations can be used directly to predict permeability given porosity. The advantages of this approach include

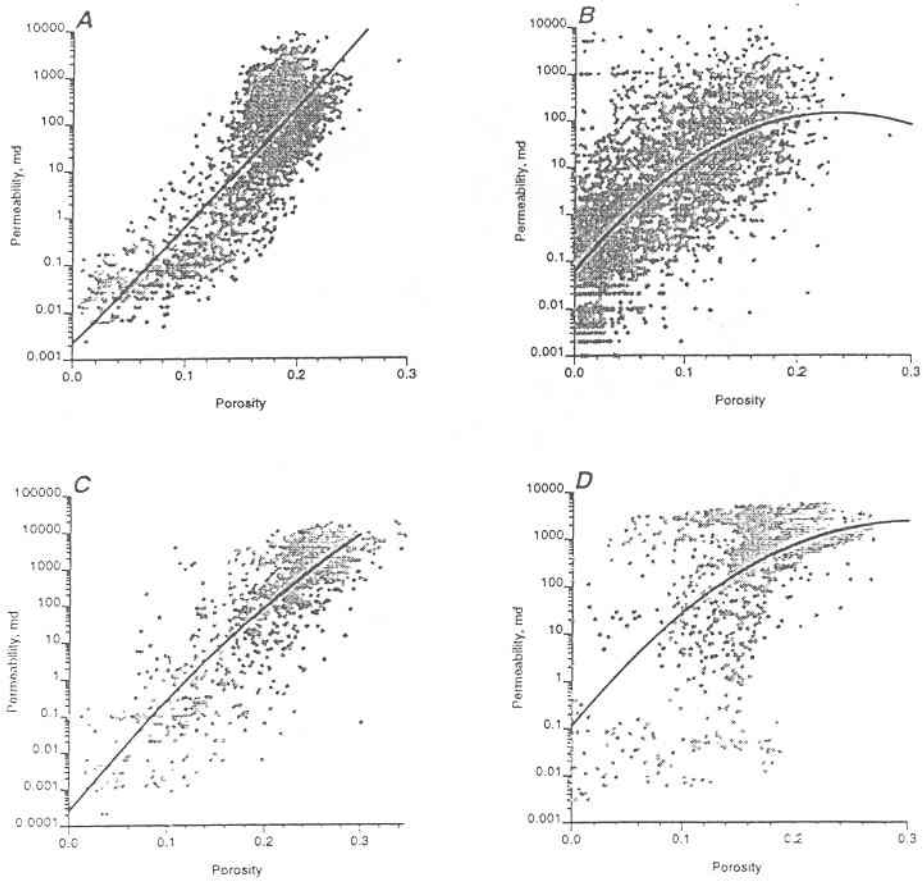


Figure 8.8: Four examples of porosity/permeability cross plots with second-order regression curves.

(1) the approach is simple and almost automatic, and (2) the relation between porosity and permeability is approximately reproduced.

There are disadvantages to regression. The low and high permeability values are smoothed out, that is, it is unlikely to encounter predicted values as extreme as found in the data. The variogram or spatial variability of the predicted permeability values is borrowed from porosity; however, permeability is usually more variable than porosity. Finally, regression-based permeability values do not account for uncertainty specific to permeability, beyond that of porosity. Geostatistical simulation techniques overcome these limitations.

Conditional Expectation

Assuming a parametric relation between (log) permeability and porosity can be limiting. Often, the relation is more complex than the simple polynomial relations used above (8.1). For that reason, the conditional expectation of permeability given porosity is more flexible to capture complex non-linear relationships and the conditional variance of permeability given porosity. The concept of a conditional expectation transform was illustrated in Figure 8.3. Here is the procedure to construct a conditional expectation curve:

- Sort the N porosity-permeability pairs in order of ascending porosity.
- Choose a moving window filter size M on porosity. This number will depend on the number of paired observations available. M must be greater than 10 to avoid erratic fluctuations. M is usually less than $N/10$ to avoid over smoothing.
- Average the porosity/permeability values for moving window averages of M values at a time. This generates $N-M$ pairs of smoothed porosity-permeability values.

Permeability is then predicted by interpolating between the smoothed porosity-permeability pairs. The permeability can be kept constant for each porosity class or linear interpolation can be used. The most critical decision is how to handle the lower tail (below the lowest averaged porosity value) and the upper tail (above the highest porosity value). Figure 8.9 gives the four conditional expectation curves that correspond to the paired porosity/permeability data introduced in Figure 8.8.

There may be too few data points to fully specify the bivariate relationship between porosity and permeability. A bivariate smoothing algorithm could be used to fill in the cross plot [58, 220].

Monte Carlo Simulation from Conditional Distributions

The concept of conditional expectation curves can be taken one step further. The permeability value at a location $\mathbf{u}$ can be drawn by Monte Carlo simulation from the conditional distribution of permeability given the porosity at that location $f(k|\phi(\mathbf{u}))$. A series of conditional distributions are constructed, see the three on Figure 8.10. In general, 10 or more conditional distributions are used. The porosity "windows" used to construct the conditional distributions can overlap. The details of Monte Carlo simulation were covered in Section 2.7.

The histogram of permeability and the full scatter between porosity and permeability is reproduced with this approach. Some spatial correlation is imparted to the predicted permeability values because of correlation with porosity; however, the spatial variability of the predicted permeability values

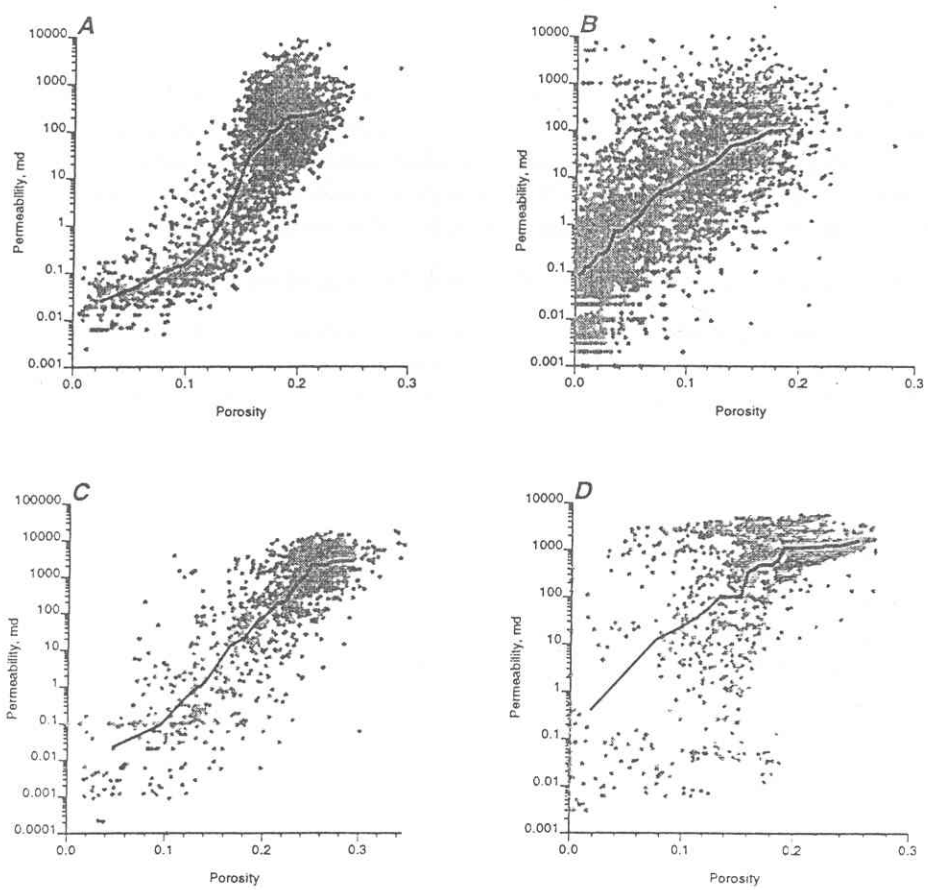


Figure 8.9: Four examples of porosity/permeability cross plots with conditional expectation curves; second-order regression equations are shown in Figure 8.8.

will be too random since previously drawn permeability values are not considered. A geostatistical approach is required to impart the correct spatial correlation to permeability.

8.4 Gaussian Techniques for Permeability

The cokriging procedure described above in Section 8.2 to simulate porosity using seismic data works very well for permeability correlated to porosity. The implicit assumption is that porosity, permeability, and the bivariate relationship of porosity and permeability follow a (bi)Gaussian distribution. The procedure:

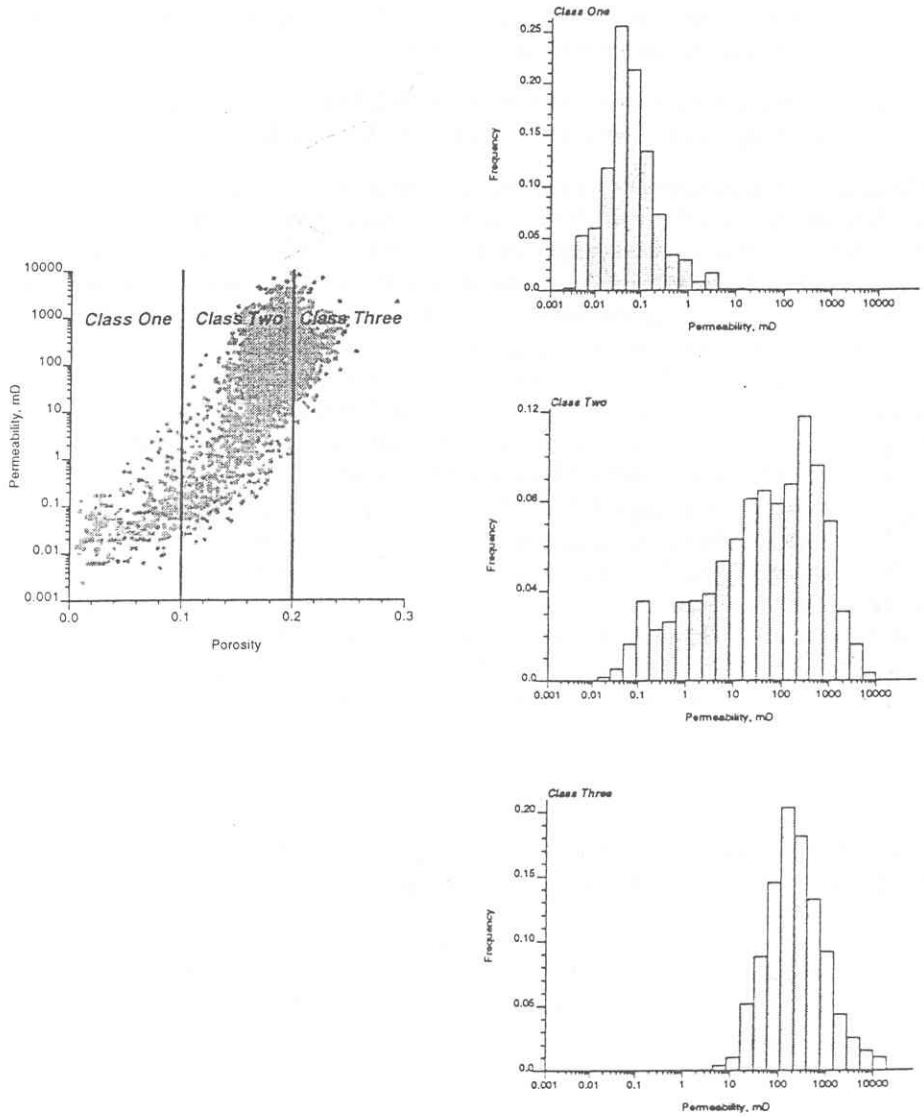


Figure 8.10: Porosity/permeability cross plot with three conditional distributions for Monte Carlo simulation. More conditional distributions would be used in practice; three are shown for illustration.

- Normal score transform porosity and permeabilty. The normal score variogram of permeability and the correlation ceofficient between the normal score transforms of porosity and permeability are needed. Full cokriging would require a linear model of coregionalization between porosity and permeability (see Section 4.6).
- Simulate the normal score transform of permeability using porosity as a covariate. The results are checked and back transformed.

Collocated cokriging works well since previously simulated porosity is present at all locations. Figure 8.11 shows the four examples of normal score porosity permeability cross plots corresponding to previous cross plots in Figures 8.8 and 8.9. The bivariate distributions shown in all four cases do not appear to have the elliptical probability contours required of Gaussian techniques because they are not bivariate Gaussian.

In general, such Gaussian techniques are widely used because they are simple to apply and reproduce most required features. A characteristic feature of permeability, however, is that the continuity of extreme low values (flow barriers) and extreme high values (flow conduits) are very important in subsequent flow modeling studies. As mentioned in Chapter 5, a characteristic feature of Gaussian techniques is that the extreme values are disconnected. For this reason, Gaussian techniques are not always recommended for permeability. Prior facies models would capture large-scale continuity. In presence of a "good" facies model it is unlikely that a more sophisticated permeability model is warranted. Nevertheless, indicator methods, described in the next Section, may be used to account for greater or lesser continuity of extreme values.

Small Example

Figure 8.12 shows two wells 600 m apart and a cross section through a porosity model between the two wells. The cross plot between porosity and permeability for this example is given on Figure 8.11.D. Gaussian simulation was used to generate permeability realizations (after calculating the variogram of the normal scores of permeability). A realization of permeability generated by Gaussian simulation using porosity is shown in Figure 8.13. Note the good correlation between porosity and permeability. The sample correlation coefficient is reproduced exactly.

8.5 Indicator Technique for Permeability

The indicator formalism for kriging and simulation is described in Section 5.4. Details on indicator cokriging for facies modeling is given in Section 6.2. Implementation details of indicator methods for permeability are described here. Indicator methods for permeability would be used when the spatial continuity at low and high permeability thresholds does not fit the Gaussian

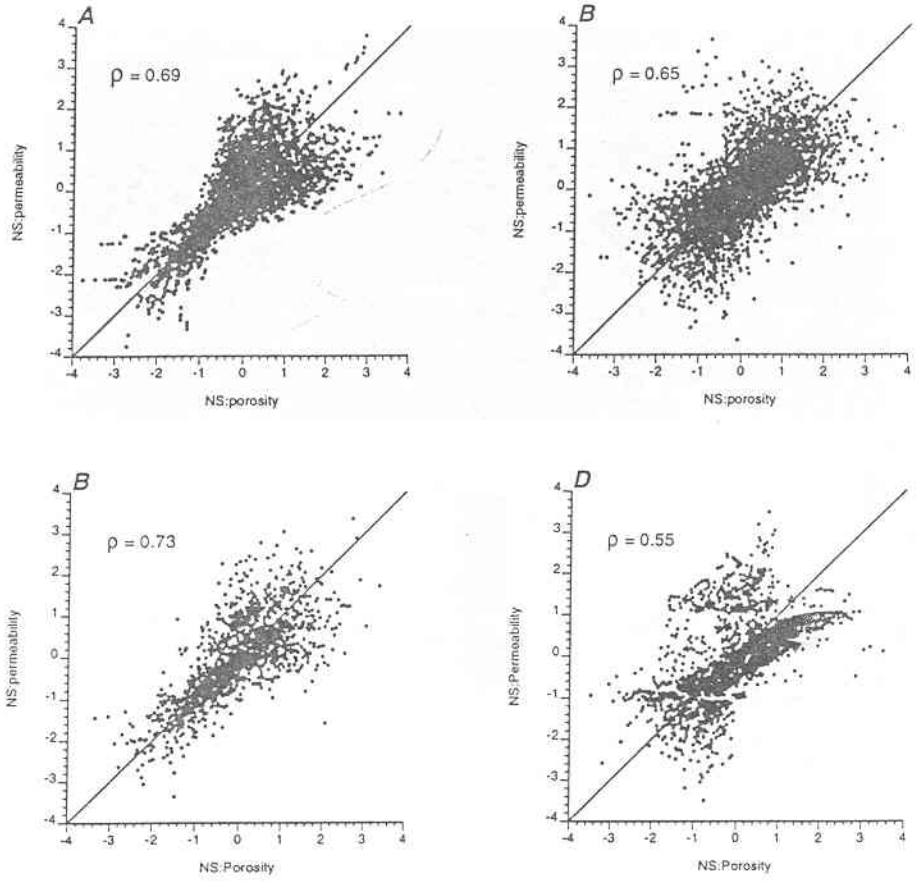


Figure 8.11: Four examples of normal score porosity/permeability cross plots corresponding to previous cross plots in Figures 8.8 and 8.9. Note that the bivariate distributions shown in all four cases do not appear to have the elliptical probability contours required of Gaussian techniques.

model. The price of this greater flexibility is more professional and CPU time for model building. Some form of cokriging is required to account for the correlation with porosity.

Indicator kriging directly estimates the conditional probability distribution at an unsampled location with no explicit Gaussian assumption. The range of permeability variability is divided with a series of threshold values $z_j, j = 1, \ldots, N_J$. The symbols k or K are commonly used for permeability; therefore, j notation is used to avoid confusion. As mentioned in Section 5.4, between 5 to 11 thresholds are considered between the 0.1 and 0.9 quantiles. Choosing the nine deciles is a good starting point. Fewer thresholds could be considered to simplifiy the procedure. Additional thresholds could be added

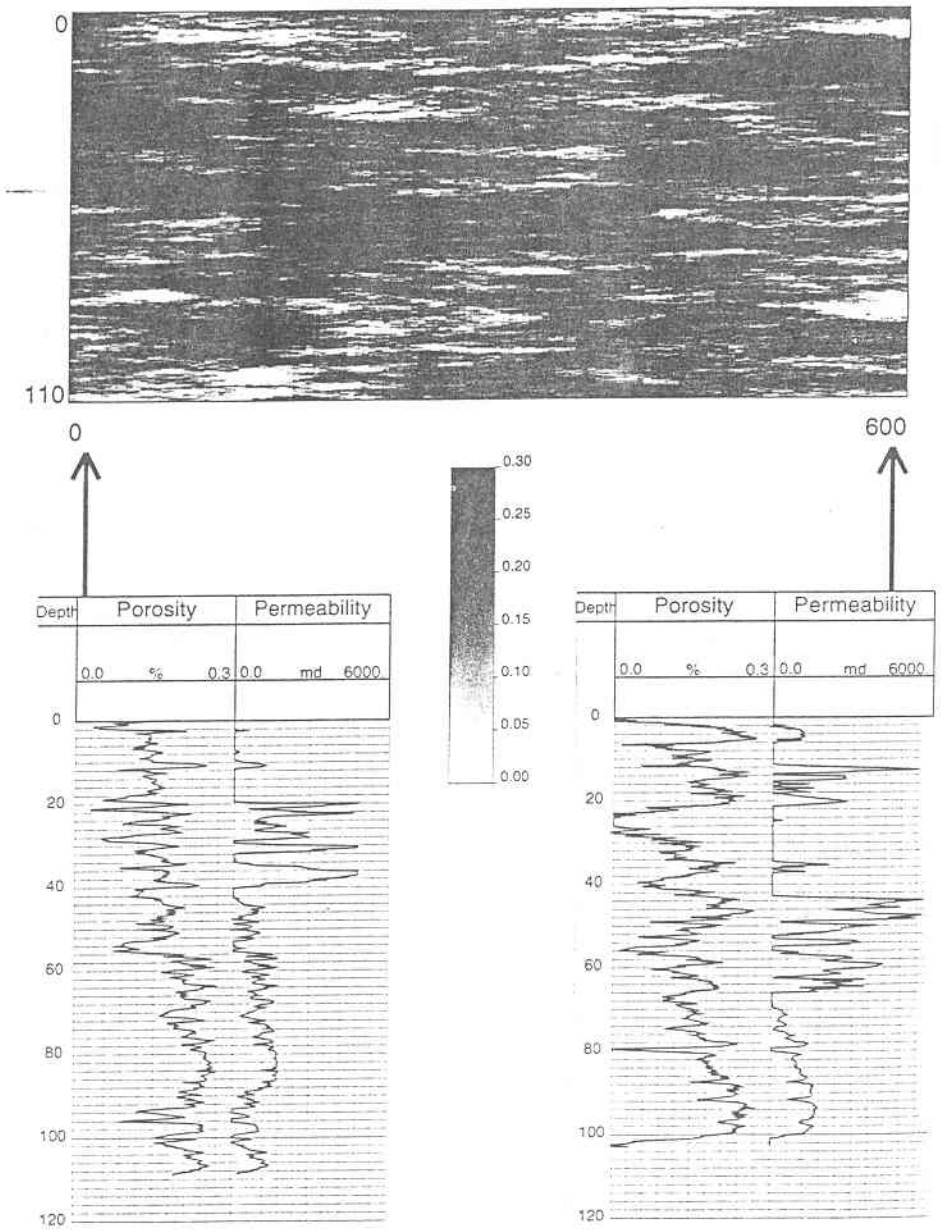


Figure 8.12: Simulated porosity model and well logs for two wells 600 m apart. The porosity model will be used as a covariate for permeability modeling.

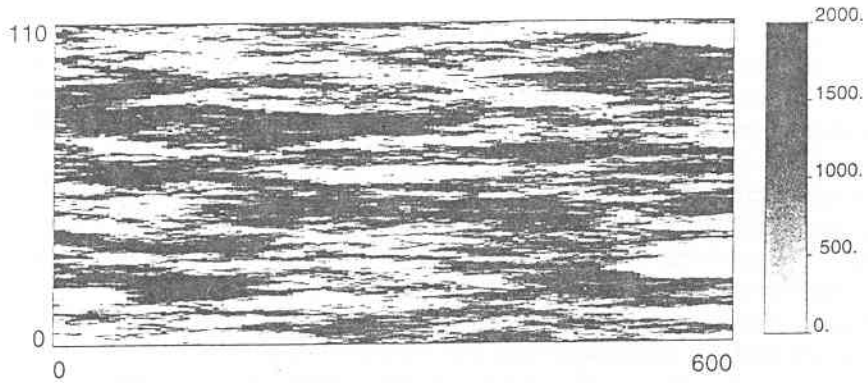


Figure 8.13: A realization of permeability generated by Gaussian simulation using porosity as a secondary variable (see Figure 8.12).

to provide greater resolution. Choosing too many thresholds results in significant order relation violations, which defeats the goal of choosing more thresholds.

A variogram is needed for each threshold. Most software makes this easy. Nevertheless, the same variogram can be used for a number of thresholds, for example, indicator variogram at the 0.1-quantile can be used for the 0.1, 0.2, and 0.3 quantiles. The common origin of the indicator data will impart consistency in the indicator variograms, that is, the parameters such as nugget effect and anisotropy directions of the indicator variograms should vary smoothly from one cutoff to another. Drastically different indicator variograms will be difficult to reproduce in indicator simulation because of order relation violations.

The porosity values must be coded or transformed to pre-posterior probabilities of permeability at the thresholds $z_j, j = 1, \ldots, N_J$ considered for permeability modeling:

$$y(\mathbf{u}; z_j) = Prob\{Z(\mathbf{u}) \leq z_j | \phi(\mathbf{u})\} \quad j = 1, \ldots, N_J$$

A high porosity $\phi(\mathbf{u})$ at a particular location would, in general, entail a high permeability, that is, a small probability $y(\mathbf{u}; z_j)$ to be less than a low threshold z_j. A low porosity $\phi(\mathbf{u})$ would, in general, entail a low permeability; thus, a high probability $y(\mathbf{u}; z_j)$ to be less than a low threshold z_j. These probabilities are calculated using the conditional distributions from a cross plot of collocated values, see details in Section 8.3.

A distribution of uncertainty in permeability at each location $\mathbf{u}$ is then directly estimated through kriging the indicator transform at each threshold:

$$[i(\mathbf{u}; z_j)]^* = \sum_{\alpha=1}^{n} \lambda_\alpha i(\mathbf{u}_\alpha; z_j) + \sum_{\beta=1}^{n'} \lambda'_\beta y(\mathbf{u}_\beta; z_j) \quad j = 1, \ldots, N_J \tag{8.2}$$

See also equation (5.19). There are n hard nearby permeability data with their indicator transforms $i(\mathbf{u}_\alpha; z_j)$ and n' nearby secondary porosity data (choosing the single collocated porosity value is often adequate) with their indicator transforms $i_{soft}(\mathbf{u}_\beta; z_j)$. The weights $\lambda_\alpha, \alpha = 1, \ldots, n$ and $\lambda'_\beta, \beta = 1, \ldots, n'$ are calculated by indicator cokriging.

A linear model of coregionalization could be built betwen the i-permeability indicators and the y-porosity indicators at each threshold, see Section 4.6; however, the simpler Markov-Bayes model [280] is commonly adopted. This was introduced in Section 6.2 in the context of categorical indicator data.

Calibration parameters $B_j, j = 1, \ldots, N_J$ are required for each threshold. These calibration parameters and the i-permeability indicator variograms fully specify the direct and cross variograms needed for the cokriging, see 6.3. The coefficients B_j are obtained by comparing collocated hard and soft indicator data:

$$B_j = E\{Y(\mathbf{u}; z_j) \mid I(\mathbf{u}; z_j) = 1\} - E\{Y(\mathbf{u}; z_j) \mid I(\mathbf{u}; z_j) = 0\} \quad \in [-1, +1] \tag{8.3}$$

$E\{\cdot\}$ is the expected value operator or simply the arithmetic average. The term $E\{y(\mathbf{u}; z_j) \mid I(\mathbf{u}; k) = 1\}$ is close to 1 and $E\{y(\mathbf{u}; z_j) \mid I(\mathbf{u}; z_j) = 0\}$ is close to 0 if the porosity is highly correlated to permeability. The best case is when $B_j \approx 1$. When $B_j = 1$, the porosity indicator data are treated as hard indicator data. Conversely, when $B_j = 0$, the porosity data are ignored, that is, their kriging weights will be zero.

Indicator kriging (IK), that is, estimation of the probability at each threshold with equation (8.2) leads to a distribution of uncertainty. This IK-derived distribution can then be used for stochastic simulation as part of the sequential indicator simulation (SIS) algorithm. The steps for indicator simulation of permeability may be summarized as:

1. Choose permeability thresholds and establish indicator variogram models for each threshold. The same variogram model can be used for multiple thresholds.

2. Transform the porosity values to secondary indicator data using the cross plot between collocated porosity and permeability values (the `bicalib` program of GSLIB can be used for this purpose).

3. Calculate the B calibration parameters for each threshold. The porosity indicator data need not be used if these parameters are all close to zero, which would indicate that there is no correlation between porosity and permeability.

4. Perform sequential indicator simulation to create multiple realizations of permeability that reproduce (1) the permeability conditioning data and histogram, (2) the continuity at different permeability thresholds, and (3) the correlation with porosity.

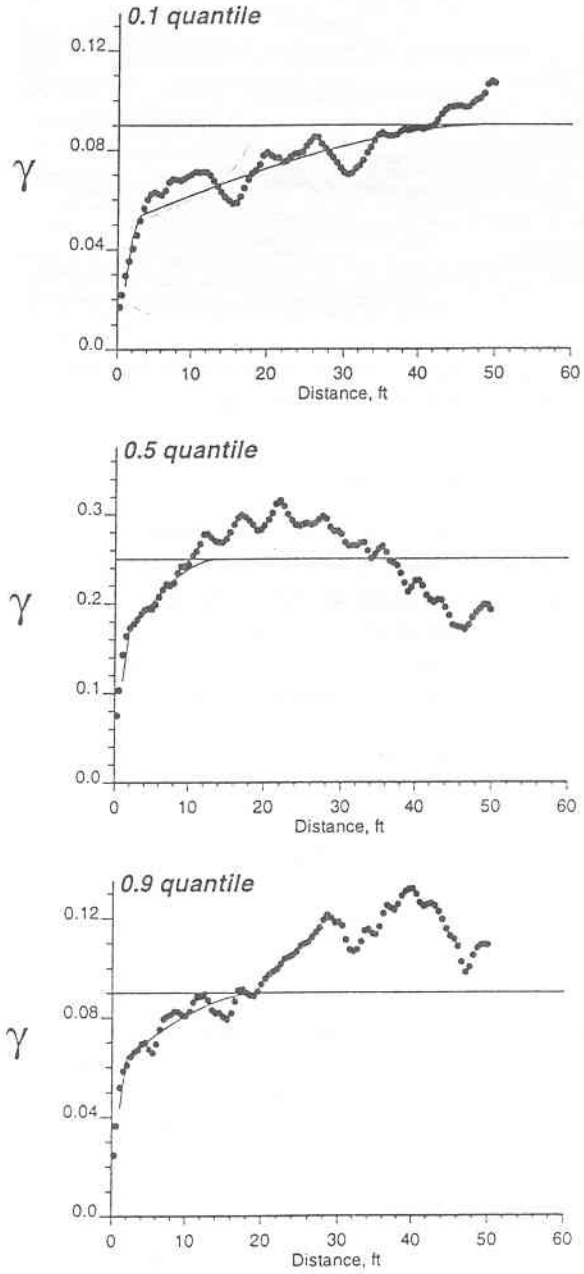


Figure 8.14: Permeability indicator variograms for 0.1, 0.5, and 0.9 quantiles. Note the different spatial continuity, especially the greater continuity of the low values relative to the high values.

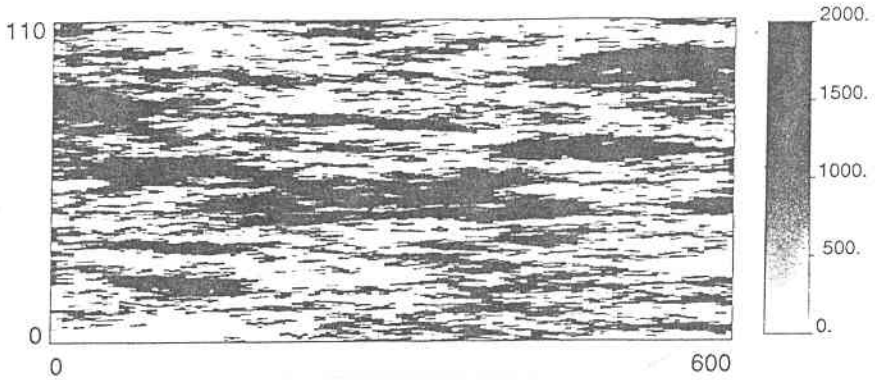


Figure 8.15: A realization of permeability generated by sequential indicator simulation with the Markov-Bayes option for cokriging.

Continuation of Small Example

Gaussian techniques were used for Figures 8.12 and 8.13 shown above. An indicator procedure was considered here. Three indicator variograms for the 0.1, 0.5, and 0.9 quantile thresholds of permeability are shown in Figure 8.14. Note the different spatial continuity, especially the greater continuity of the low values relative to the high values. The Gaussian model has only one degree of freedoom in the input variogram; it would be unable to capture these features.

Figure 8.15 shows a permeability realization generated by sequential indicator simulation. This was generated with the Markov- Bayes approach to secondary data integration (`sisim_gs` in GSLIB). Note that this realization appears more "ordered" or with "less entropy" than the Gaussian realization shown in Figure 8.13.

8.6 Work Flow

Following are the main geostatistical workflows related to property modeling. Figure 8.16 illustrates the work flow to calibrate seismic data to porosity data at well locations. The main considerations are the scale difference and the normal score transform. Attention should be paid to the fact that averaging of porosity is not linear after normal score transformation.

Figure 8.17 illustrates the work flow for porosity modeling by sequential Gaussian simulation with seismic data. This is very similar to work flow 5.1 back in Chapter 5; however, there is often seismic data and trends that must be considered; these additional data inputs and implementation considerations are shown.

Figure 8.18 shows the work flow for permeability modeling by porosity-permeability regression-type transforms. The work flow is straightforward and requires little geostatistical input, which is why it is commonly used in

practice. Figure 8.19 illustrates the work flow for permeability modeling using a porosity model for conditioning. The collocated approach is commonly used for simplicity; it is relevant since a previously simulated porosity value is available for every grid cell where permeability is being simulated.

The continuity of extreme high and low permeability values can be very important in fluid flow predictions. Sequential indicator simulation is a technique to account for this greater continuity. Figure 8.20 illustrates the work flow for sequential indicator simulation of permeability using a prior porosity model.

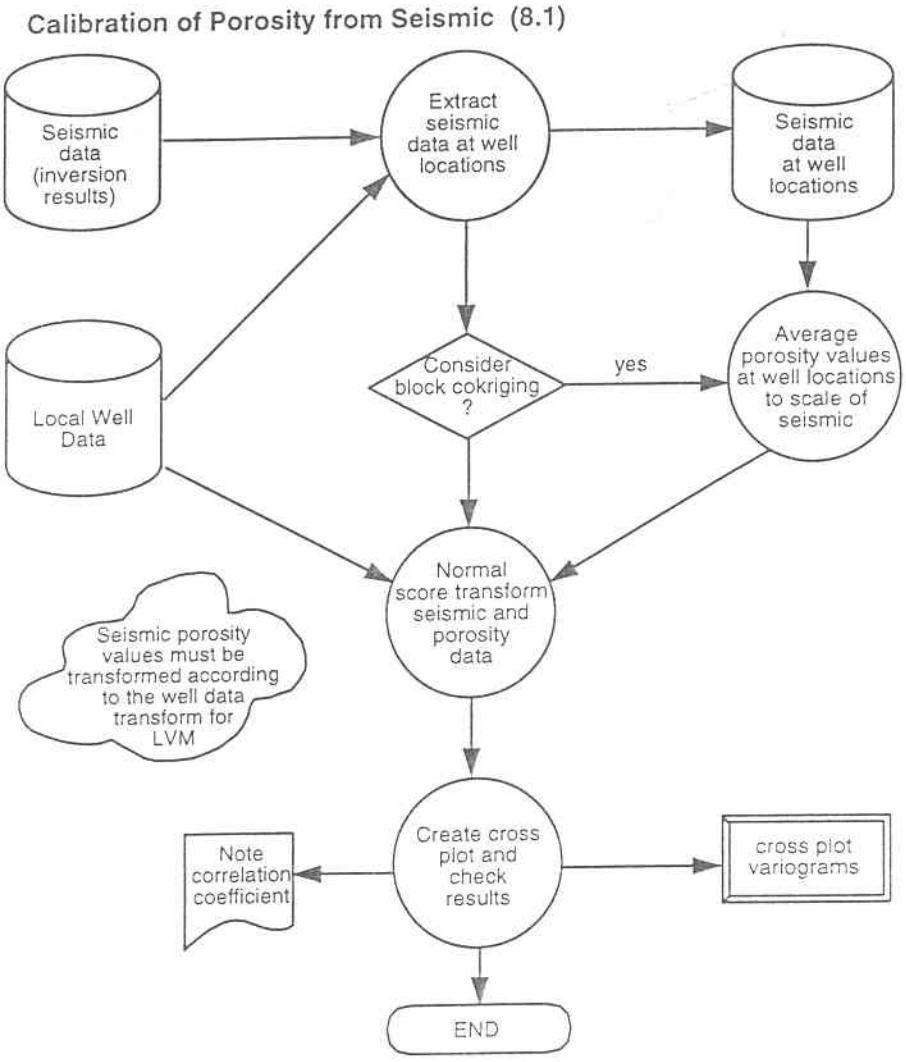


Figure 8.16: Work Flow 8.1: calibration of porosity from collocated seismic data. There are times when the search for nearby seismic data must be expanded to include more than the collocated data.

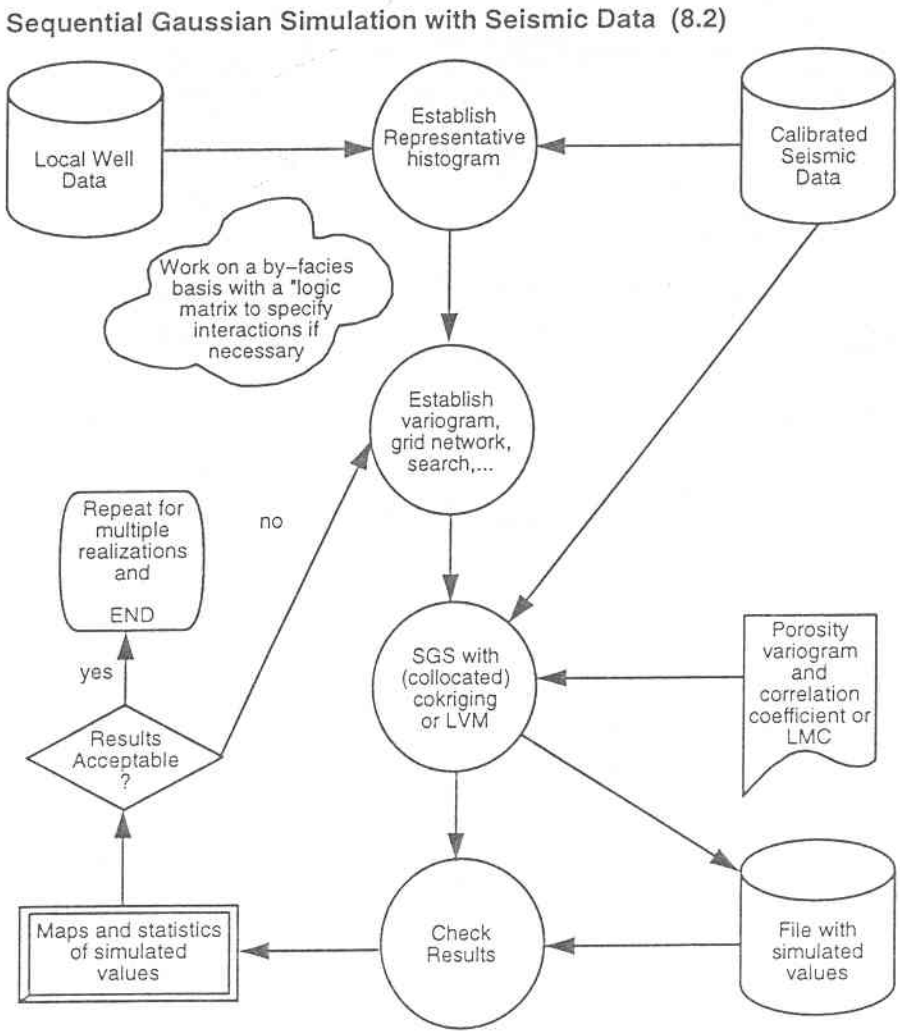


Figure 8.17: Work Flow 8.2: porosity modeling by sequential Gaussian simulation with seismic data.

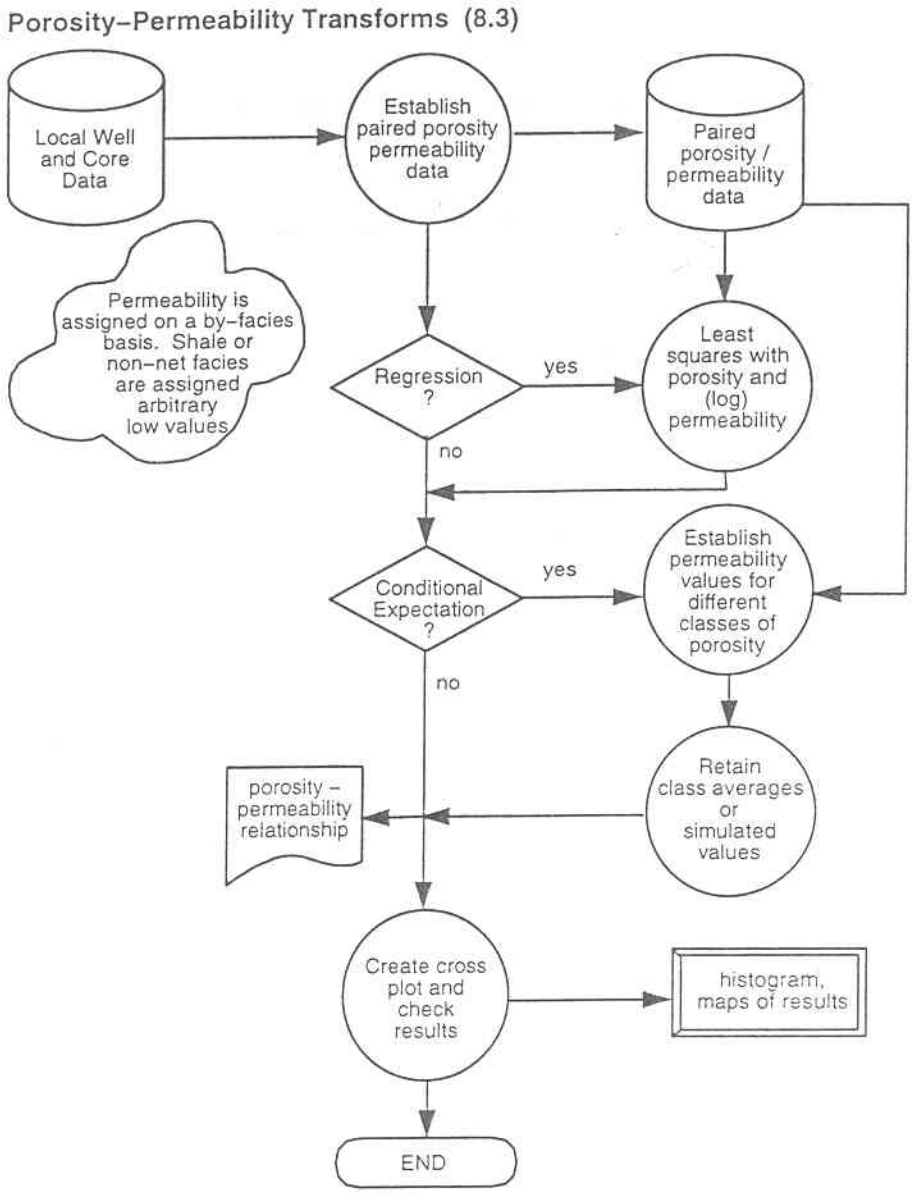


Figure 8.18: Work Flow 8.3: permeability modeling by regression-type transformation using porosity.

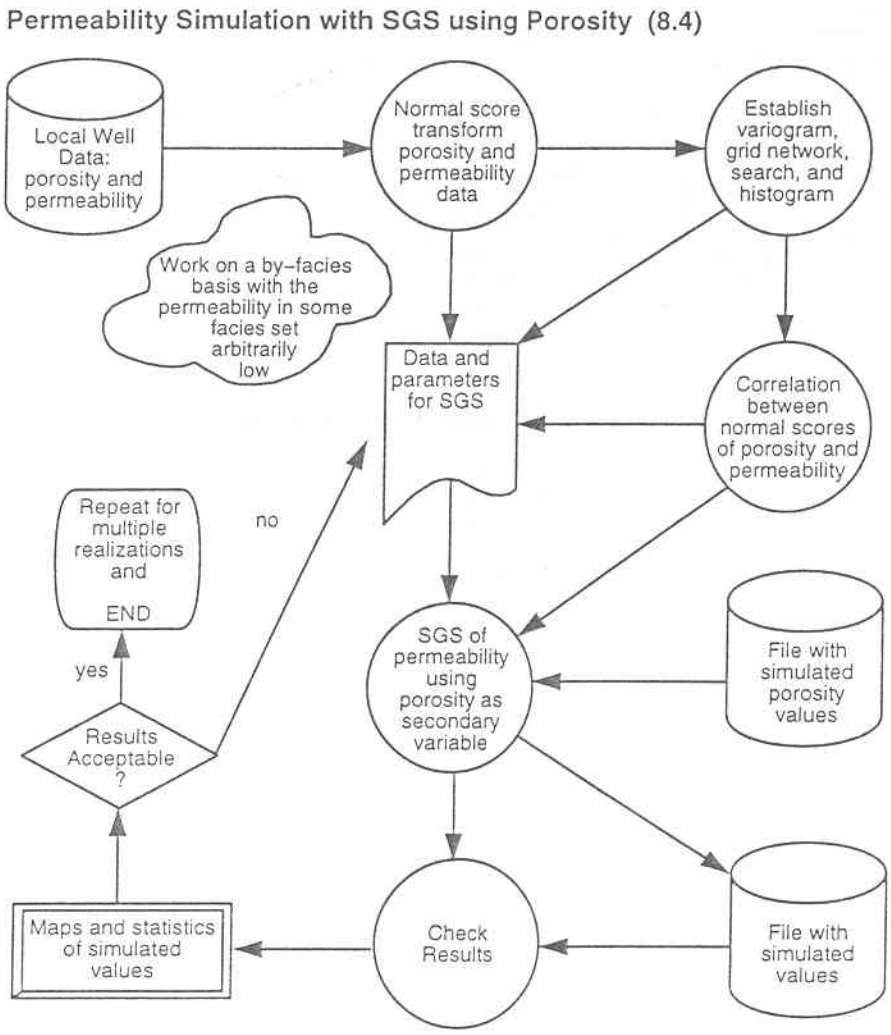


Figure 8.19: Work Flow 8.4: permeability modeling by sequential Gaussian simulation using porosity as a covariate.

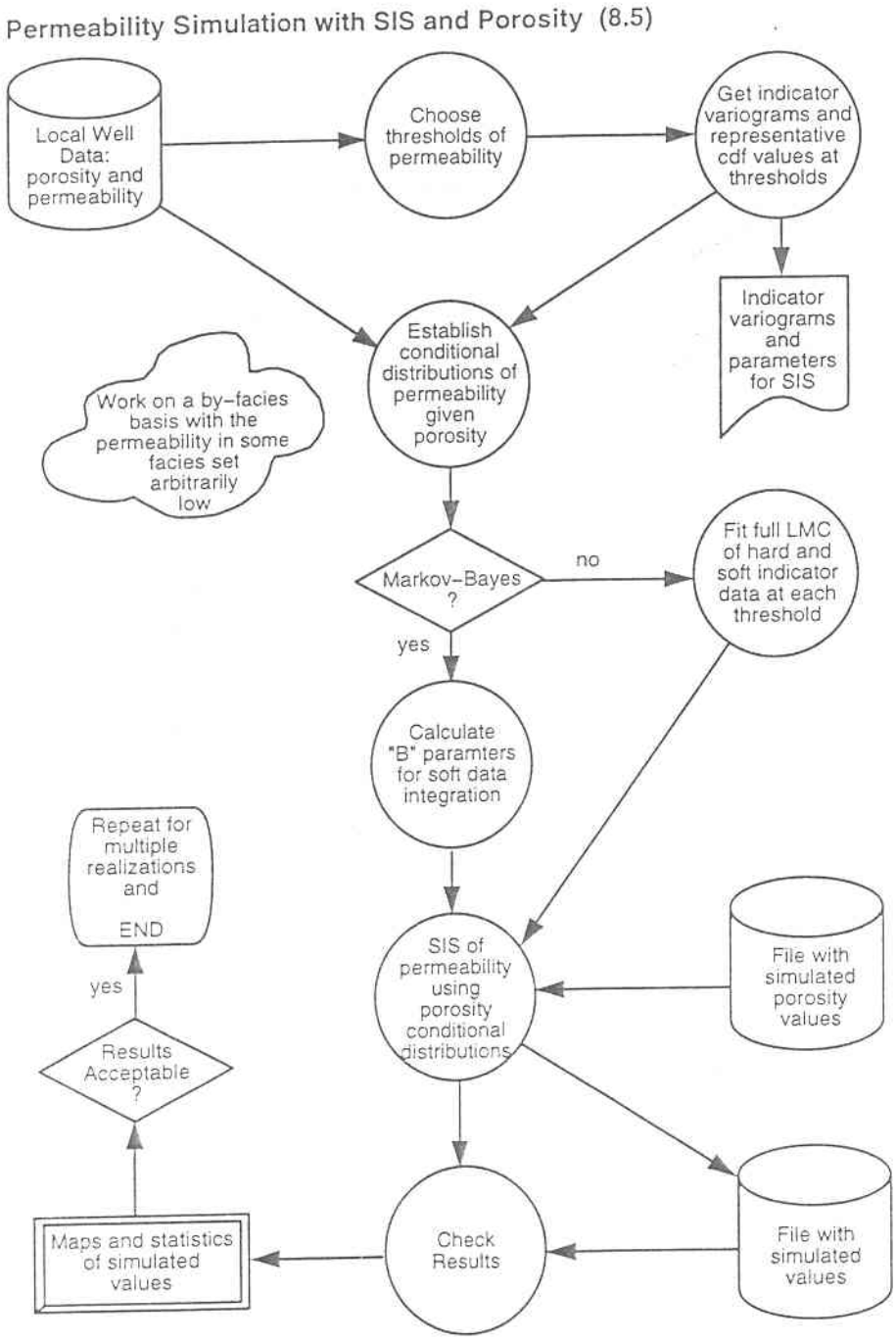


Figure 8.20: Work Flow 8.5: permeability modeling by sequential indicator simulation using porosity as a covariate.

Chapter 9

Simulated Annealing for Geostatistics

Data integration is of fundamental importance in reservoir modeling. The optimization technique of simulated annealing has found application in the integration of diverse data types. The background and theoretical framework for simulated annealing for geostatistical mapping is presented in Section 9.1.

The steps of simulated annealing are described in Section 9.2. The key concept is iterative correction of an initial model to impart desired spatial features. There are a number of issues related to (1) the initial image, (2) the perturbation mechanism, and (3) the objective function used in simulated annealing; these issues are addressed in Section 9.2.

Simulated annealing is a solution method in the field of combinatorial optimization that has been adapted to geostatistical applications; it is not a stochastic simulation algorithm. This requires careful implementation to avoid artifacts such as (1) edge effects, (2) discontinuities at conditioning data, and (3) a restricted space of uncertainty. These problem areas are covered in Section 9.3.

In spite of the problems of simulated annealing, it has found a place in reservoir characterization. Specific examples where annealing is suited to reservoir modeling are discussed in Section 9.4.

9.1 Background

The *method of simulated annealing* is an optmization technique that has attracted significant attention. For practical purposes, simulated annealing has effectively "solved" the famous *traveling salesman problem*, that is,

> Find the shortest cyclical itinerary for a traveling salesman who must visit each of N cities once and only once in a sequential path.

This is a classic combinatorial minimization problem where the solution space is not simply the N-dimensional space of N parameters. There is a discrete but factorially large configuration space, for example, the number of routes increases as $N!$. Another characteristic feature is that there is no notion of "direction" or "downhill." Thus the minimization cannot rely on the calculation of gradients or derivatives, which are commonly used in optimization problems.

The central idea behind simulated annealing is an analogy with thermodynamics, specifically with the way liquids freeze and crystallize, or metals cool and anneal. At high temperatures the molecules can move freely. As the temperature is slowly lowered, the molecules line up in crystals which represent the minimum energy state for the system. The Boltzmann probability distribution, $P\{E\} \sim exp(-E/(k_b T))$, expresses the idea that a system in thermal equilibrium at a temperature T has the energies of its component molecules probabilistically distributed among all different energy states E. The Boltzmann constant k_b is a natural constant which relates temperature to energy. Even at a low temperature there is a probability that the energy is quite high, in other words, a system will sometimes give up a low energy state in favor of a higher energy state [206].

Metropolis and his coworkers [188] extended these principles to numerically simulate molecular behavior. A system will change from a configuration of energy O_1 to a configuration of energy O_2 with probability $p = exp(-(O_2 - O_1)/(k_b T))$. The system will always change if O_2 is less than O_1 (that is, a favorable step will always be taken), and sometimes an unfavorable step is taken: this has come to be known as the Metropolis algorithm. More generally, any optimization procedure that draws upon the thermodynamic analogy of annealing is known as simulated annealing.

In the early 1980s Kirkpatrick et al. [162] and independently Cerny [256] extended these concepts to combinatorial optimization. They formulated an analogy between the objective function and the free energy of a thermodynamical system [2, 206]. A control parameter, analogous to temperature, is used to control the iterative optimization algorithm until a low energy (objective function) state is reached.

One of the first direct applications to spatial phenomena was published by Geman and Geman [100] who applied the method to the restoration of degraded images. About the same time Rothman [216] at Stanford applied the method to non-linear inversion and residual statics estimation in geophysics. Independent research by C. L. Farmer [87] led to the publication of a simulated annealing algorithm for the generation of rock-type models. This triggered considerable interest in the method among geostatisticians [56, 229].

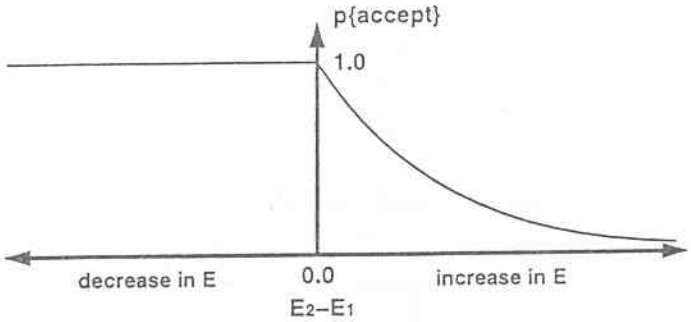


Figure 9.1: Probability of accepting a change to the system in simulated annealing: probability is 1.0 when the objective function decreases and the probability follows an exponential distribution when the objective function increases.

The essential contribution of simulated annealing is a prescription for when to accept or reject a given perturbation. The acceptance probability distribution is given by:

$$P\{accept\} = \begin{cases} 1, & \text{if } O_{new} \leq O_{old} \\ e^{\frac{O_{old}-O_{new}}{t}}, & \text{otherwise} \end{cases} \tag{9.1}$$

This probability distribution is shown in Figure 9.1. All favorable perturbations ($O_{new} \leq O_{old}$) are accepted, and unfavorable perturbations are accepted with an exponential probability distribution. The parameter t of the exponential distribution is analogous to the "temperature" in annealing. The higher the temperature, the more likely an unfavorable perturbation will be accepted.

In the physical process of annealing, the temperature must not be lowered too fast or else the material will be "quenched" (frozen in its evolution) and never achieve a low-energy state. The same phenomenon happens in simulated annealing; the image or realization will get trapped in a suboptimal situation and never converge when the t parameter is lowered too quickly. Time in the context of simulated annealing may be considered as the total number of attempted perturbations. The scheme employed to reduce the temperature with time is known as the *annealing schedule*. The determination of a suitable annealing schedule for different problems remains an outstanding challenge in the application of simulated annealing to geostatistical problems.

9.2 Steps in Annealing

Figure 9.2 illustrates the overall flow chart for simulated annealing applied to geostatistical problems. The steps in annealing will be considered one at a time in the following sections.

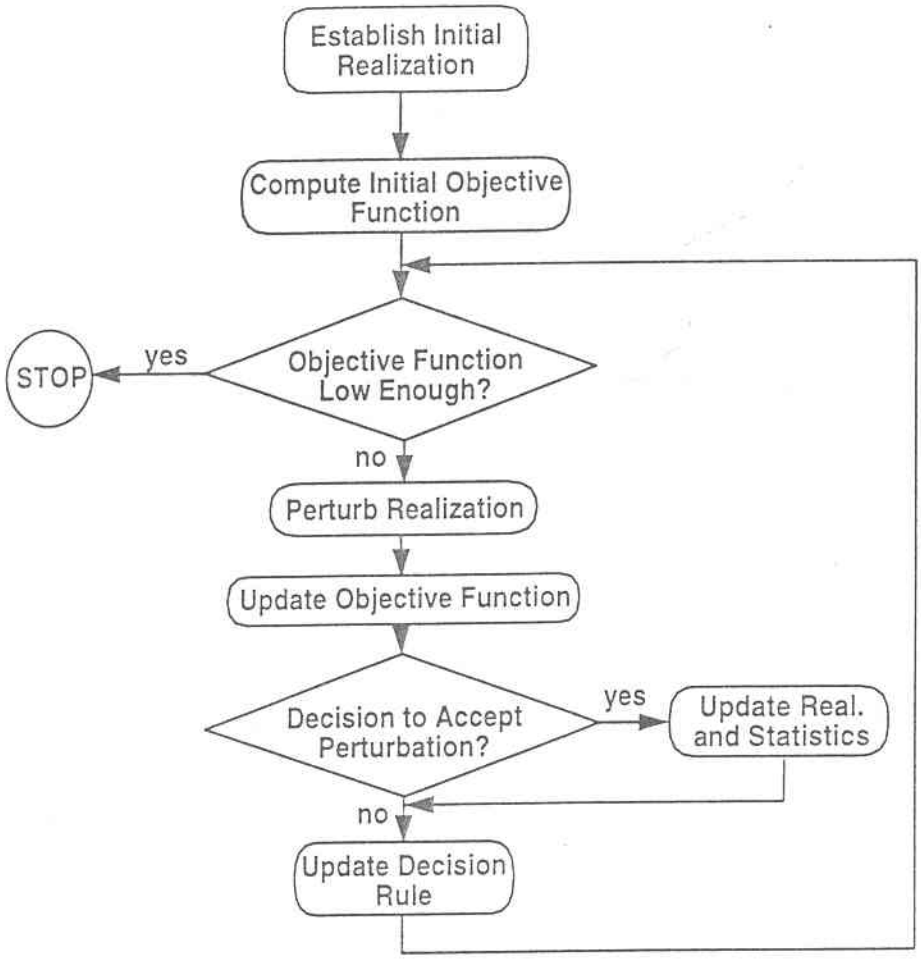


Figure 9.2: A flow chart with the steps in simulated annealing. Each step will be presented in detail in Section 9.2.

Initial Realization

The initial realization is the starting point for the simulated annealing optimization. We commonly consider a facies indicator $i(\mathbf{u})$ or a petrophysical property $z(\mathbf{u})$ on a regular grid, that is, a regular Cartesian grid of size $i = 1, \ldots, nx$ by $j = 1, \ldots, ny$ by $k = 1, \ldots, nz$ (typically, $nx \cdot ny \cdot nz \approx 10^5$) with

$$f_{i,j,k} = \text{facies code taking a value } f = 0, 1, \ldots, n_F$$

$$\phi_{i,j,k} = \text{porosity } \phi \in (0, 1]$$

$$K_{i,j,k} = \text{horizontal permeability } k \in (0, \infty]$$

The use of a regular Cartesian grid is not a condition for simulated annealing; however, a Cartesian grid that conforms to major stratigraphic layering makes updating the objective function more straightforward; see later.

The space of uncertainty or possible configurations is N^K where N is the number of grid nodes and K is the number of allowable outcomes. This space is larger than we can possibly imagine; however, we are only interested in the configurations that reasonably approximate all available data (within their reliability).

We do not have to consider cell-based representations of geologic properties. The "realization" in simulated annealing could be an object-based model of geological facies units such as channels, levees, and crevasse splays. The space of uncertainty would then be the parameters that control the specific positioning and geometry of the geologic facies objects.

Consider two options for constructing the initial realization for simulated annealing: (1) a completely random or even constant value realization, or (2) the result of some faster (but less flexible) simulation algorithm, which imparts some desireable features. Although the second option provides an initial realization that starts at a lower objective function, it is often better to start with a random or non-commital realization. There are a number of reasons for initial random realizations. Firstly, there are no initial large-scale features in the wrong place that are hard to undo. In fact, it takes more time to reconfigure a wrong image than to configure an initially random image. A second reason in support of an initial random realization is the resulting space of uncertainty, which is larger and more realistic with random initial realizations [56].

Simulated annealing "freezes" large-scale features early, at relatively high temperatures, and then converges on details at lower temperatures. It is necessary to set an initially high temperature to permit large-scale features to be established by all components of the objective function. Such high temperature results in the initial image being randomized, which defeats the goal of using an initially correlated image.

Objective Function

Simulated annealing requires an energy or objective function that measures closeness to data or desired features. In general there are multiple data sources that are coded as separate component objective functions. These components are put together in a weighted sum for the total objective function:

$$O = \sum_{i=1}^{N_c} \omega_i \, O_i$$

where N_c is the number of components, $\omega_i, i = 1, \ldots, N_c$ are the weights, and $O_i, i = 1, \ldots, N_c$ are the component objective functions. Some possible component objective functions are listed and then a procedure to determine the weights will be discussed.

Each component objective function O_i is a mathematical expression to enforce some data or impart some desirable spatial character to the realization. All component objective functions are measures of mismatch between a desired property and that of a 3-D model. The $*$ superscript in the following objective functions denotes the property of the 3-D model; no subscript indicates the target property. Some examples:

- All reservoir models must reproduce local facies, porosity, and permeability data coming from core and well log data. These data may be forced into the model by construction, that is, assigning them to the closest grid node location or they may be reproduced as part of the objective function. Three component objective functions for local data:

$$\begin{aligned} O_i &= \sum_{\alpha=1}^{n_{data}} \sum_{k=1}^{n_{facies}} [i(\mathbf{u}_\alpha; k) - i^*(\mathbf{u}_\alpha; k)]^2 \\ O_Z &= \sum_{\alpha=1}^{n_{data}} [Z(\mathbf{u}_\alpha) - Z^*(\mathbf{u}_\alpha)]^2 \end{aligned} \tag{9.2}$$

  where $\mathbf{u}_\alpha, \alpha = 1, \ldots, n_{data}$ are data locations, the $*$ superscript denotes the facies, porosity, or permeability in the candidate model, i (in this context) is the facies indicator, Z is a continuous variable such as porosity, permeability, or water saturation. There may be a different number of data for each variable; the notation has been kept simple.

- The declustered representative histogram of the facies or continuous variable should also be reproduced by the numerical model:

$$\begin{aligned} O_{p_i} &= \sum_{k=1}^{n_{facies}} [p_k - p_k^*]^2 \\ O_{F_Z} &= \sum_{j=1}^{n_q} [F_Z(q_j) - F_Z^*(q_j)]^2 \end{aligned} \tag{9.3}$$

  where p_k is the proportion of facies $k = 1, \ldots, n_{facies}$ and $F_Z(q_j)$ is the cumulative distribution function at a particular quantile $q_j, j = 1, \ldots, n_q$. The number of quantiles n_q would be chosen by the modeler. Once again, the $*$ denotes the property of the candidate model.

- Beyond local data and the correct univariate distribution, reservoir models should reproduce two-point variogram measures of spatial correlation:

$$O_\gamma = \sum_{j=1}^{n_{\mathbf{h}}} [\gamma(\mathbf{h}_j) - \gamma^*(\mathbf{h}_j)]^2 \tag{9.4}$$

  The number of lags, $n_{\mathbf{h}}$, is chosen to "inject" all spatial correlation deemed important. It is unnecessary to consider distances and directions where there is no spatial correlation; all features not imposed by a

component objective function will show no correlation due to the initial random model.

- Kriging techniques cannot easily account for multiple point measures of spatial correlation; however, complex measures of spatial correlation can be coded into an objective function with little difficulty. For example, the N point connectivity in a particular $\mathbf{h}$ direction:

$$O_C = \sum_N [C(\mathbf{h}; N; z_c) - C^*(\mathbf{h}; N; z_c)]^2 \tag{9.5}$$

  where C is a multiple N-point covariance for unit lag $\mathbf{h}$ at threshold z_c. Different multiple point statistics of arbitrary configuration and size could be considered; however, N cannot be too big in practice because there are rarely sufficient data for reliable inference of the target multiple-point statistics.

- As mentioned in Section 5.5, deterministic trends are important in reservoir modeling. Such trends could enter as component objective functions. For example, in the case of vertically varying facies proportions:

$$O_v = \sum_{k=1}^{n_{facies}} \sum_{z=1}^{nz} [p_k(z) - p_k^*(z)]^2 \tag{9.6}$$

  where $p_k(z)$ is the proportion of facies k at areal slice z; $k = 1, \ldots, n_{facies}$ the number of facies, and $z = 1, \ldots, nz$ the number of areal slices. Areal variations could be considered in a similar manner. Areal and vertical variations in the average of a continuous variable may also be expressed in a similar manner.

- Multiple petrophysical properties are almost always considered, for example, facies, porosity, and permeability. The relationships between the variables must be reproduced by the reservoir models for reliable predictions. As an example, consider the bivariate distribution between porosity and permeability:

$$O_{\phi/K} = \sum_{i=0}^{n_\phi} \sum_{j=0}^{n_K} [f(\phi_i, K_j) - f^*(\phi_i, K_j)]^2 \tag{9.7}$$

  where the bivariate distribution is discretized into $i = 1, \ldots, n_\phi$ porosity classes and $j = 1, \ldots, n_K$ permeability classes, and $f(\phi_i, K_j)$ is the probability of a porosity/permeability pair falling into a particular class. The correlation coefficient could be considered as a simpler measure of correlation; however, bivariate probabilities are more flexible to capture non-linear and complex relationships.

- Seismic data is of significant importance when well data is sparse. The bivariate relation between seismic data and average porosity could be

captured as bivariate probabilities, as above in equation (9.7); however, there are often too few calibration wells to infer the full bivariate relation. In many cases, we can only infer the correlation coeficient between the porosity and permeability with any degree of reliability. This correlation correlation between seismic data and porosity could be imposed at the correct scale:

$$O_\rho = \left[\rho_{\overline{\phi},Impedance} - \rho^*_{\overline{\phi},Impedance}\right]^2 \tag{9.8}$$

where $\rho_{\overline{\phi},Impedance}$ is the classical correlation coefficient.

- Another important source of data is well test and historical production data. Such data must be reproduced by numerical reservoir models for reliable future predictions. There is a large variety of production data that is handled differently; Chapter 11 discusses this dynamic data in more detail. As an example, we could construct a component objective function for a well test derived effective permeability:

$$O_{wt} = \sum_{i=1}^{n_{well}} \sum_{t=1}^{n_{time}} \left[\overline{k}(\mathbf{u}_i, t) - \overline{k}^*(\mathbf{u}_i, t)\right]^2 \tag{9.9}$$

  where $i = 1, \ldots, n_{well}$ is the number of wells with well test intervals, $t = 1, \ldots, n_{time}$ are the time intervals, and $\overline{k}(\mathbf{u}_i, t)$ is the effective permeability for well i and time t. The link between time and space is commonly made by Oliver's weighting function [199].

The component objective functions must be positive and decrease to zero when data or spatial statistic is reproduced by the realization.

In general, the number of required perturbations to achieve convergence is between 10 and 1000 times the number of variables in the system. Each component objective function must be updated or recalculated after each perturbation. As discussed below, we should aim at being able to locally update the objective function after a local perturbation to the 3-D model. Global recalculation of an objective function is avoided because of CPU cost.

The units of each component objective function are different. Moreover, the rate at which each component objective function decreases to zero is different. It is necessary to weight each component to ensure that all components i are considered in the global objective function. Weights are first determined to ensure that each component has an equal importance. Then, we may increase the importance of a particular objective function as a subjective choice or in the presence of conflicting objective function components.

The absolute magnitude of the objective function is not used in making a decision; the difference is considered:

$$\Delta O = O_{new} - O_{old}$$

$$\Delta O = \sum_{i=1}^{N_c} \omega_i \Delta O_i = \sum_{i=1}^{N_c} \omega_i \left[O_{i_{new}} - O_{i_{old}}\right]$$

Our goal is to establish weights $\omega_i, i = 1, \ldots, N_c$ so that, on average, each component contributes equally to a change to the global change in the objective function ΔO, that is,

$$\omega_i = \frac{1}{\overline{|\Delta O_i|}}, \quad i = 1, \ldots, N_c$$

The average change of the objective function for each component can be determined by proposing a some number M of changes to the system and calculating the average change. Approximate by:

$$\overline{|\Delta O_i|} = \frac{1}{M} \sum_{m=1}^{M} |O_i^{(m)} - O_i|, \quad i = 1, \ldots, N_c$$

These weights could be updated throughout the annealing process; however, that has not been found necessary [56].

Stop Criteria

The iterative procedure should be stopped when the objective function is low enough to ensure that all input data are reproduced within their level of certainty. The component objective functions should be designed to go to zero. In presence of soft or imprecise data, the objective function should include the imprecision as in (9.8) above where both the scale and imprecision of seismic data is captured in the objective function. Thus, each component objective function should get close to zero. How close to zero? Less than 1% of the starting value would normally ensure close reproduction of the input statistic. The results can be checked if there is any question about how closely the input statistics are being reproduced.

In practice, the procedure is also stopped once some reasonable CPU limit is exceeded. In this case, the reproduction of each component objective function must also be checked. Slow convergence may be due to (1) a large and difficult problem that truly requires many perturbations, (2) a slow annealing schedule, which could be changed (see below), (3) conflicting objective function components, which should be resolved by retaining only the most reliable. Conflicting objective functions may not be known in advance; however, they may become evident by noting the objective functions that do not decrease to zero.

We do not search for the global minimum; the input statistics are not known so precisely and we would not have sufficient CPU power. A realization that reproduces all known data within an acceptable tolerance is a candidate realization for subsequent reservoir management decision making. We can find many realizations that reproduce all known data. Taken together these realizations represent a space of uncertainty implicit to the random function defined by the simulated annealing problem formulation.

Perturbation Mechanism

Initial applications of simulated annealing considered "swapping" as the perturbation mechanism [87]. This has the advantage that the histogram of the initial image is never changed, so there is no need to include the histogram in the objective function; however, this scheme is inflexible in the presence of complex conditioning data. Secondary data may be available in the form a component objective function that would lead to a different histogram than was used for the initial realization.

Recent applications consider proposing a change at one location at a time [62]. The locations may all be considered before reconsidering a particular location. This scheme is sometimes referred to as the "heat bath" algorithm [48]. Some advantages of the heat bath over random grid node selection have been reported.

After selecting a grid node location to perturb, a new value is drawn from either the global histogram or a local distribution, considering the values at nearby grid node locations. The global distribution is non-commital and is implemented in the `sasim` program of GSLIB [62]. Considering a local distribution, however, can speed convergence.

The local distribution may be determined by kriging weights obtained by some representative variogram and a small local template of grid node locations that excludes the collocated grid node that is being perturbed (see upper left of Figure 9.3). The kriging weights provide a direct estimate of local proportions of categorical variables. For continuous variables, there is a need to provide a continuous ccdf model between the available quantile values. A straightforward linear interpolation between the minimum, the available quantiles, and maximum is considered.

In summary, the implementation steps for this perturbation mechanism are illustrated (Figure 9.3) and described below:

1. Establish kriging weights for a template of points (excluding the location being perturbed) using a representative variogram. The median indicator variogram would be a good choice.

2. Sort the data within the template in ascending order, $z_{(1)}, z_{(2)}, z_{(3)}, \ldots, z_{(n)}$, with kriging weights: $w_{(1)}, w_{(2)}, w_{(3)}, \ldots, w_{(n)}$.

3. Calculate the cdf values for each datum

$$cp_{(i)} = \sum_{1}^{i} w_{(i)}, \quad i = 1, \ldots, n$$

   where the weights $w_{(1)}, w_{(2)}, w_{(3)}, \ldots, w_{(n)}$ sum to 1, $cp_{(0)}$ at z_{min} is 0.0, and $cp_{(n+1)}$ at z_{max} is 1.0.

4. Establish intermediate ccdf values:

$$F(z_{(i)}) = \frac{cp_{(i-1)} + cp_{(i)}}{2}, \quad i = 1, \ldots, n$$

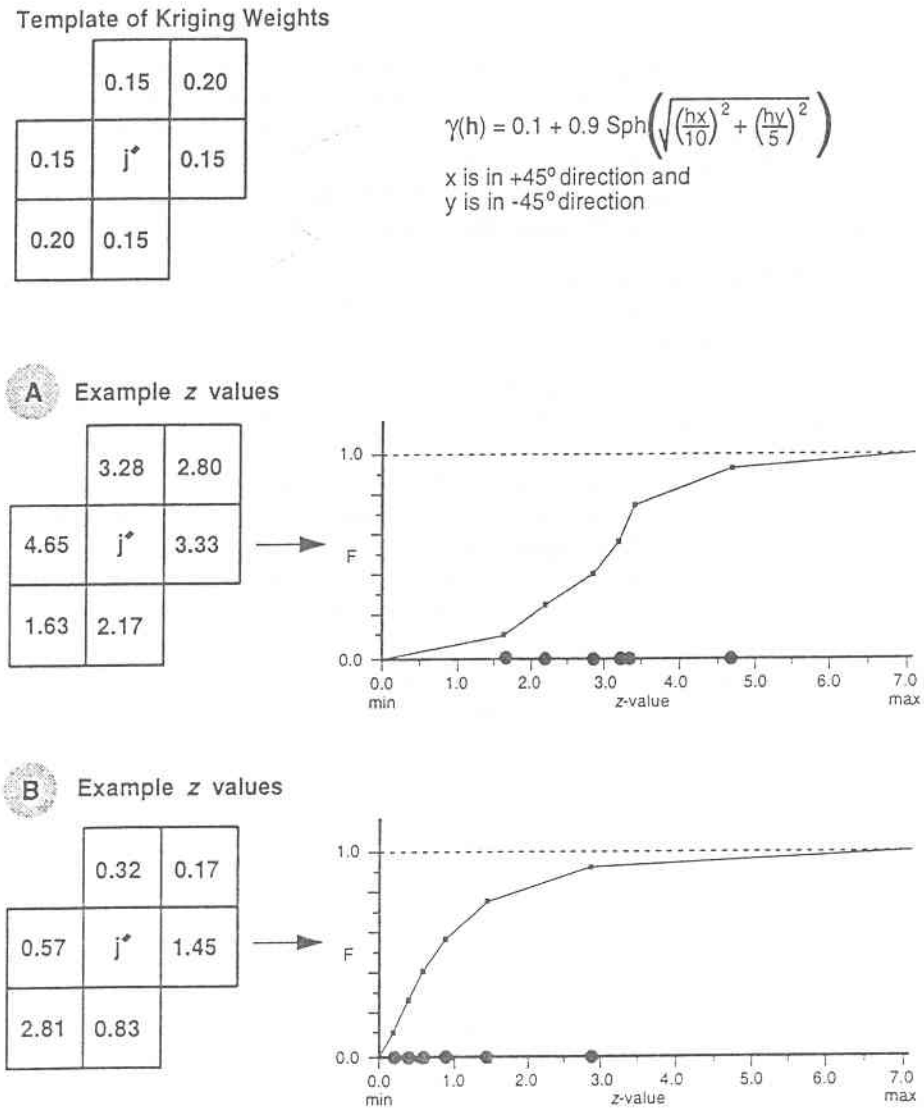


Figure 9.3: An illustration of how local distributions at location j' are constructed by applying kriging weights to local data. Two examples are shown: (A) with values that are in a relatively high area, and (B) with values that are in a relatively low area.

5. Linear interpolation allows a complete specification of the relation between $F(z)$ and z. More elaborate "tail" extrapolation methods could be considered for highly skewed data distributions with limited data [62].

6. Lastly, a new candidate value $z^{new}(\mathbf{u}_j)$ is drawn for location $\mathbf{u}_j$ from this local ccdf. This candidate value is more likely to be accepted than a value drawn from the global distribution, since it is consistent with other cell values in the local neighborhood. Moreover, since it is drawn from a local conditional distribution built by kriging, the variogram of the perturbed model will more likely be improved.

   Note that a candidate value $z^{new}(\mathbf{u}_j)$ will be drawn from the global distribution if the variance of the local distribution is greater than that of the target histogram.

The weights at the top of Figure 9.3 consider an anisotropic spherical variogram model. Two different sets of local data, hence conditional distributions, (A and B) are shown in Figure 9.3.

The only additional parameter required to use this scheme for constructing a local distribution is the size of the template, n_{tem}. There is a small CPU penalty for using a larger template, that is, it takes some CPU time to assemble the local data into a distribution. A perturbation with $n_{tem} = 4$ takes 5% more CPU time to consider than simply drawing from the global distribution. A perturbation with $n_{tem} = 40$ takes 45% more CPU time. From this perspective, a smaller template is to be preferred. A larger template, however, considers more spatial information and requires fewer perturbations to achieve convergence. The CPU time versus template size can be plotted to determine a minimum CPU value. Typical results indicate that templates of size 8/12 yield optimal results. The CPU time is decreased by at least an order of magnitude.

Update Objective Function

The objective function should be updated after every attempted perturbation. There are many attempted perturbations. As mentioned, there are normally between 10 to 1000 times the number of variables or grid nodes. For a 1 million grid node problem there are between 10 million to 1 billion attempted perturbations. Updating the objective function only occasionally slows convergence, but is the last resort for difficult objective functions. In general, all components of the objective function must be updated after each perturbation. It would be very CPU-expensive to recalculate the objective function after each perturbation. In almost all cases, however, it is possible to update the underlying spatial statistics without global recalculation.

Consider the traditional variogram as an example. The variogram is defined as: $\gamma(\mathbf{h}) = \frac{1}{2N(\mathbf{h})}\sum_{i=1}^{N(\mathbf{h})}[z(\mathbf{u}) - z(\mathbf{u}+\mathbf{h})]^2$ with the global cost of cal-

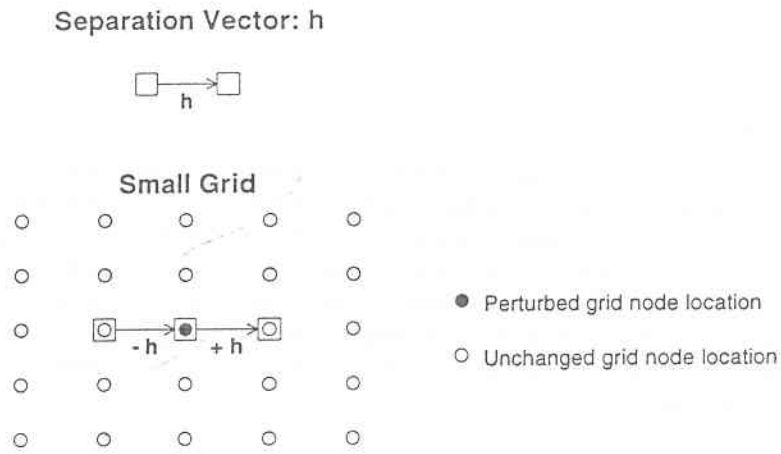


Figure 9.4: Illustration of the lags that need updating in the variogram calculation.

culation proportional to $\propto (nx-1)\cdot ny\cdot nz$ operations. The variogram, however, may be updated by few CPU operations:

$$\begin{aligned}\gamma(\mathbf{h})_{new} = \gamma(\mathbf{h})_{old} \quad &- \quad [z(\mathbf{u}) - z(\mathbf{u}+\mathbf{h})]^2 \\ &- \quad [z(\mathbf{u}-\mathbf{h}) - z(\mathbf{u})]^2 \\ &+ \quad [z'(\mathbf{u}) - z(\mathbf{u}+\mathbf{h})]^2 \\ &+ \quad [z(\mathbf{u}-\mathbf{h}) - z'(\mathbf{u})]^2\end{aligned}$$

where value $z(\mathbf{u})$ is being perturbed to $z(\mathbf{u}')$. The local cost of updating is two operations. As illustrated in Figure 9.4 the $+\mathbf{h}$ and $-lag$ contribution for a particular lag $\mathbf{h}$ must be updated; there are two operations regardless of grid size.

Virtually all spatial statistics may be updated and not globally recalculated including the example objective functions given above: (1) local data, (2) global distributions, (3) variogram measures of correlation, (4) multiple point spatial statistics, (5) vertical and areally varying proportions or averages, (6) bivariate distributions between two variables, and (7) the correlation coefficient between two variables at any scale.

There are some objective functions that cannot be locally updated. In particular, anything that is a function of *all* cells *simultaneously*, for example, dynamic flow data. In this case, it is desirable to replace the complex calculation by a simpler proxy if possible, for example, consider inverted large-scale features or multiple point measures of connectivity. The inversion or "spatial coding" of dynamic data is discussed in detail in Section 11.1.

When complex data cannot be represented by a spatial statistic, we can resort to global calculation. Of course, the difficult component objective function could be updated less frequently than the other objective function

components. Implementation becomes difficult and must be tackled by trial-and-error.

Decision Rule

The temperature parameter t in the simulated annealing decision rule must not be lowered too fast or else the image may get trapped in a suboptimal situation and never converge. However, if lowered too slowly then convergence may be unnecessarily slow. The specification of how to lower the temperature t is known as the "annealing schedule." There are some annealing schedules that ensure convergence in very particular cases [2, 100]; however, they are much too slow for practical application. One form of an annealing schedule that ensures convergence:

$$t(p) = \frac{C}{log(1+p)} \quad p = \text{ pass through system}$$

Under certain assumptions it is possible to show that, with correct choice of C, the system will converge. This is too slow for practical application.

The following empirical annealing schedule is one practical alternative [87, 206]. The idea is to start with an initially high temperature t_0 and lower it by some multiplicative factor λ whenever enough perturbations have been accepted (K_{accept}) or too many have been tried (K_{max}). The algorithm is stopped when efforts to lower the objective function become sufficiently discouraging. The following parameters describe this annealing schedule (see also the chart of Figure 9.5):

t_0: the initial temperature.

λ: the reduction factor $0 < \lambda < 1$ usually set to 0.1.

K_{max}: the maximum number of attempted perturbations at any one temperature (on the order of 100 times the number of nodes). The temperature is multiplied by λ whenever K_{max} is reached.

K_{accept}: the acceptance target. After K_{accept} perturbations are accepted the temperature is multiplied by λ (on the order of 10 times the number of nodes).

S: the stopping number. If K_{max} is reached S times then the algorithm is stopped (usually set at 2 or 3).

ΔO : a low objective function indicating convergence.

Figure 9.6 shows the typical reduction in the global objective function versus the number of attempted perturbations. The rate of decrease in the objective function is non-linear. The objective function remains quite high until a "critical temperature," then it drops quickly until convergence, and finally the objective function drops slowly. Note the logarithmic scaling on both axes of Figure 9.6.

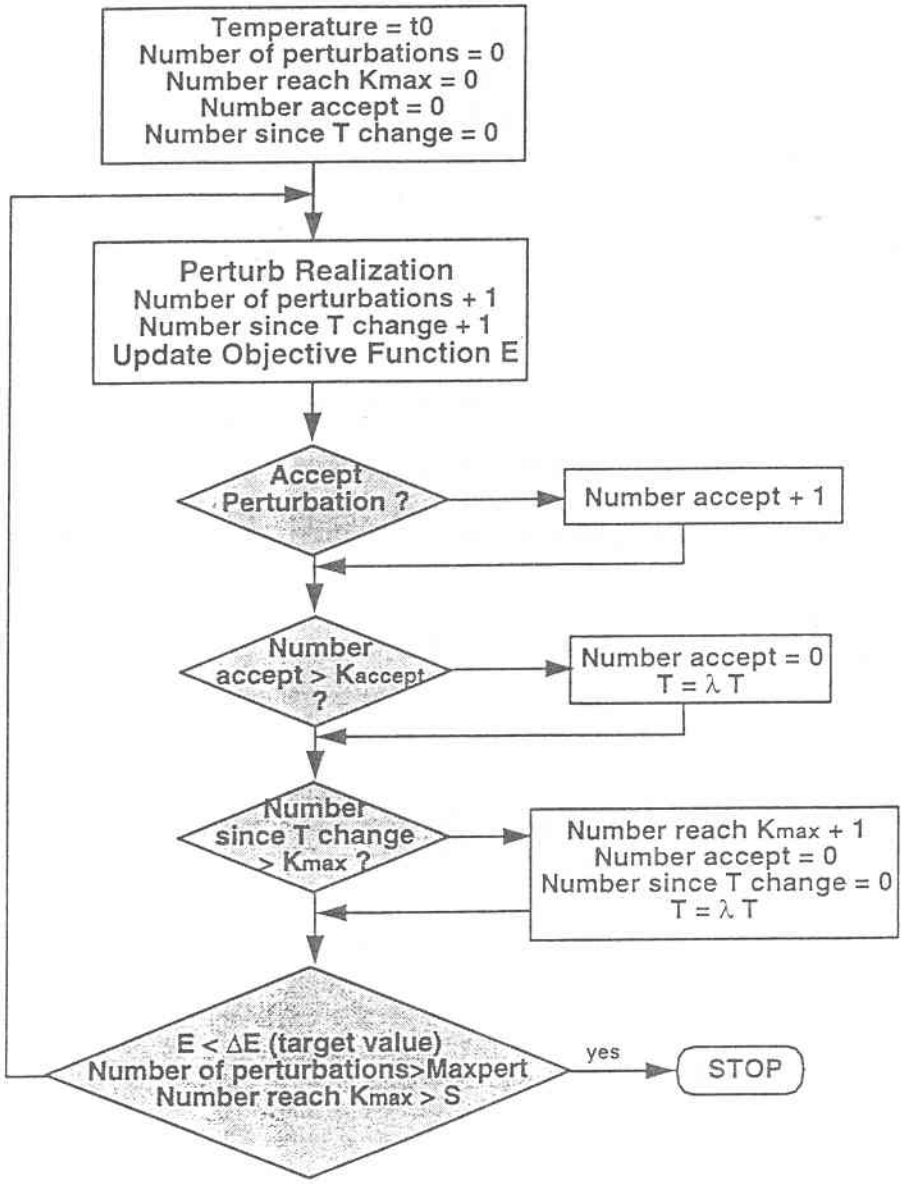


Figure 9.5: Illustration of annealing schedule.

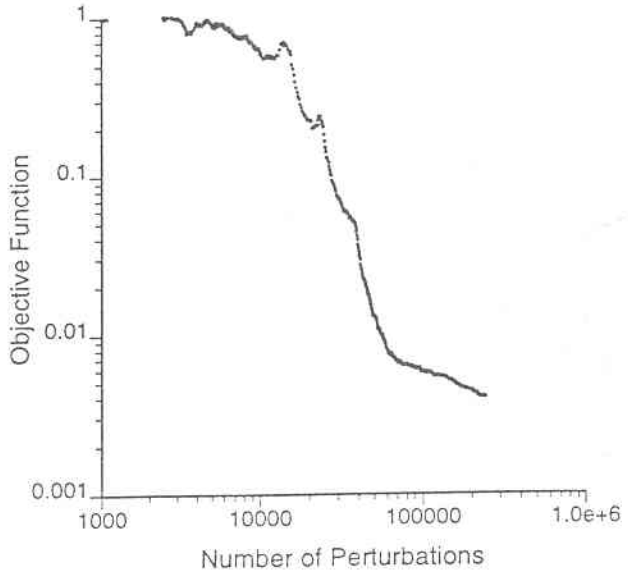


Figure 9.6: Illustration of typical reduction in objective function versus number of perturbations.

The Maximum A Posteriori (MAP) Variant

Bayesian classification schemes provide a variant to the simulated annealing algorithm [14, 24]. The classic works of Geman and Geman [100], Besag [24], and others have recently been applied in geostatistics by Doyen [70]. The same basic algorithm is followed; however, all favorable perturbations are retained and all unfavorable perturbations are rejected. Convergence is not achieved for difficult problems.

Threshold Accepting (TA)

A third published stochastic relaxation technique is a method called threshold accepting (TA) [77]. Threshold accepting differs from annealing and the MAP approaches only in the decision rule to accept or not an unfavorable perturbation. An unfavorable perturbation will be accepted if the change in the objective function is less than a specified threshold. The threshold is lowered in much the same way as the temperature parameter t is lowered in simulated annealing.

Many different variants of the same basic annealing algorithm are possible by considering different objective functions, ways to create the initial image, perturbation mechanisms, and decision rules.

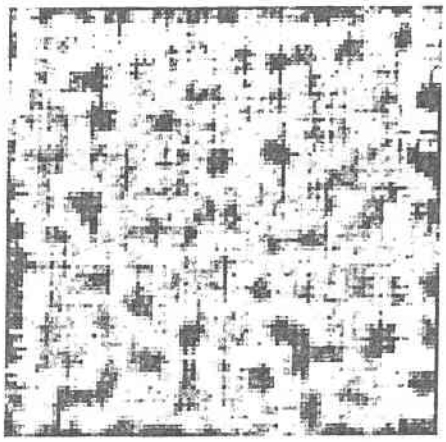


Figure 9.7: Illustration of thermodynamic edge effects caused by too-fast cooling and no special weighting of edge pairs.

9.3 Problem Areas

There is no equivalent to the metallurgical cooling in applications of simulated annealing. Nevertheless, there is an artifact that appears as "thermodynamic edge effects"; see Figure 9.7. Extreme values (black and white colors in Figure 9.7) are preferentially located at the edge of the grid when the temperature is lowered quickly. This is due for example to the variogram component of the objective function. The pure nugget variogram of the initial random realization can be improved quickly by placing high (and low) values at the grid edge where they only enter a variogram pair once and not for $+\mathbf{h}$ and $-lag$. The histogram is reproduced and the variogram is improved by such placement. Of course, this is unsatisfactory and a special weighting of edge pairs can be considered in software.

Another artifact of fast annealing is discontinuities near well data; see Figure 9.8 where there is artificial discontinuities near two well locations. The objective function is near zero for such an image: the data are honored albeit with a discontinuity, the proportion of white and black facies is reproduced, and the variogram is reproduced. The problem is that the variogram pairs involving the conditioning data are overwhelmed by the pairs of the remaining grid nodes. Convergence takes place without particular attention to the conditioning data, which are reproduced by construction. One solution to this problem is to place extra weight on the variogram pairs involving conditioning data [62].

The potential for a quasi-exact reproduction of input statistics, see Figure 9.9, is another problem with simulated annealing. Uncertainty in input statistics must be accounted for explicitly. There are "ergodic" fluctuations in kriging-based techniques, which somehow account for the variability of the

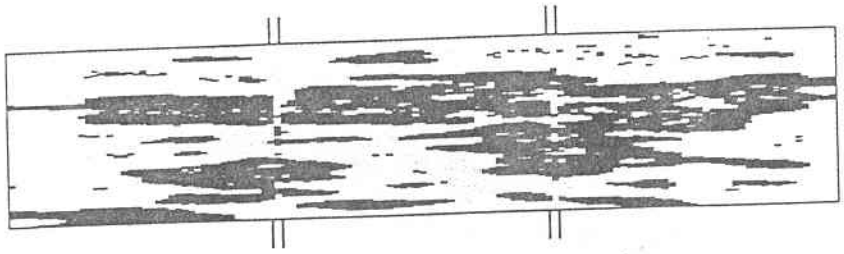


Figure 9.8: Illustration of discontinuities at conditioning data locations caused by too-fast cooling and no special weighting of well locations.

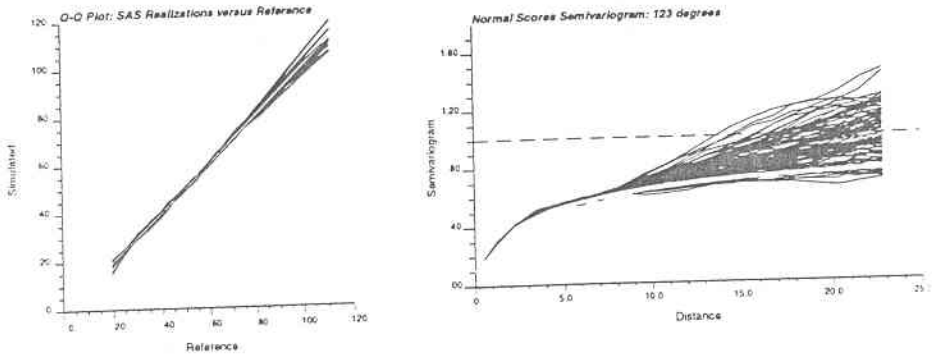


Figure 9.9: Illustration of quasi-exact reproduction of input statistics, which could lead to a too-narrow space of uncertainty.

input statistics for the region being simulated. There is a concern in simulated annealing that such close reproduction of input statistics artificially reduces the final "space of uncertainty." A number of case studies [56, 57] show that this is not a concern; in fact, the space of uncertainty of simulated annealing is larger than that of kriging-based techniques. Nevertheless, the challenge remains to reasonably account for the uncertainty in input statistics, this is of particular concern in reservoir modeling with few well data.

The CPU time requirements for simulated annealing methods can become excessive particularly when there are many perturbations required with a complex objective function. In attempts to reduce the CPU requirements, one often embarks upon the time-consuming process of adjusting the tuning parameters that constitute the weighting function and the annealing schedule. The tradecraft required to do this successfully is a significant problem.

9.4 Place of Simulated Annealing/Work Flow

Simulated annealing is a powerful optimization algorithm that can be used for reservoir modeling; however, it is more difficult to apply than kriging-based methods because of difficulties in setting up the objective function and choosing many interrelated parameters such as the annealing schedule. Therefore, the place of simulated annealing is not for conventional problems where kriging-based simulation is adequate.

Simulated annealing is applicable to difficult problems that involve: (1) dynamic data, (2) large-scale soft data, (3) multiple point statistics, (4) geo-object placement, and (5) special continuity of extremes. In these cases, kriging-based techniques do not have the flexibility to account for all data.

Figure 9.10 illustrates a work flow for 3-D model construction with simulated annealing. The use of simulated anenaling is not standard and defies simple presentation on a workflow. Correct setup of the component objective functions and their updating is a challenge.

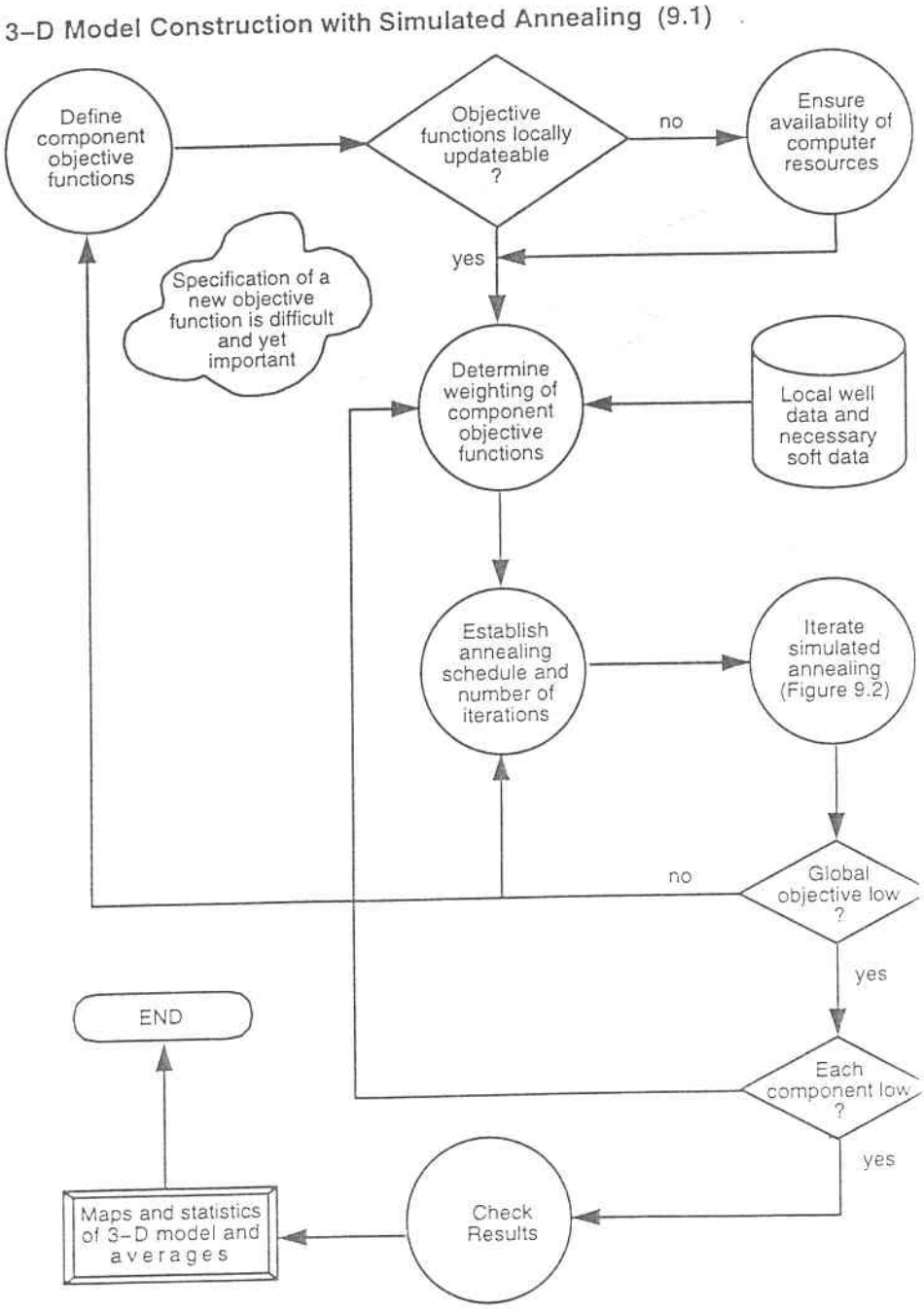


Figure 9.10: Work Flow 9.1: 3-D model construction with simulated annealing.

Chapter 10

Uncertainty Management

Multiple geostatistical realizations can be generated that reproduce the input data equally well. This chapter addresses the inevitable questions related to the space of uncertainty, validating geostatistical models, the number of realizations required, the value of ranking realizations, and decision making in the presence of uncertainty.

Section 10.1 addresses the space of uncertainty represented by geostatistical reservoir models. Fixing the conceptual geological model and the input statistics for all realizations may be unrealistic. An increasingly popular alternative is to consider alternative scenarios to more completely represent our state of incomplete knowledge.

Geostatistical estimation may be checked by setting each well aside and predicting the corresponding reservoir properties with the remaining data. Estimation (Section 10.2) is different from simulation (Section 10.3), which provides an entire distribution of uncertainty rather than a single estimated value. Continuous variables are different from categorical variables. Section 10.3 describes cross validation of simulation techniques for both continuous and categorical variables.

The number of realizations required for a particular reservoir modeling study depends on the aspect of uncertainty being characterized and the precision with which the uncertainty must be known. Section 10.4 presents guidelines to determine the number of required realizations.

Often, flow simulation can only be performed on a limited number of realizations due to CPU requirements and the need to consider uncertainty in fluid properties, well locations, and other aspects of the reservoir development plan. Section 10.5 presents techniques for ranking and selecting realizations for flow simulation. Section 10.6 addresses decision making in the presence of uncertainty. Finally, Section 10.7 presents some work flow diagrams for uncertainty assessment and ranking.

10.1 Models of Uncertainty

A geostatistical reservoir model is a "set" of spatially distributed parameters including the (1) structural definition of each stratigraphic layer, (2) facies within each stratigraphic layer, and (3) petrophysical properties such as porosity, permeability, and residual saturations on a by-layer and by-facies basis. Each model consists of *one* realization of each parameter. For example, multiple porosity realizations are not built for the same facies realization; one porosity realization is associated to each facies realization.

Multiple geostatistical realizations provide an assessment of the "model" or "space" of uncertainty. Each realization is a Monte Carlo sample from the space of uncertainty defined by all decisions implicit to the modeling approach. As discussed in Chapter 1, there is no objective or correct space of uncertainty. The space of uncertainty created by multiple realizations is realistic when the conceptual geological framework and statistical parameters, such as the variograms and size distributions, are well known. These parameters are *not* well known early in the lifecycle of a reservoir; therefore, there is more uncertainty than measured by a set of geostatistical realizations generated with the same set of underlying parameters.

A more realistic space of uncertainty is determined by a combination of a scenario-based approach and conventional geostatistical modeling. Figure 10.1 gives a schematic illustration. In this example, the reservoir could fit any of three different conceptual geological models (M-I, M-II, or M-III) with probability 0.5, 0.3, and 0.2. There may be different sets of modeling parameters (for example, variograms or size distributions) for each conceptual geological model. There are 3, 2, and 3 sets of possible parameters for the conceptual geological models in Figure 10.1. Once again the probability of each set of parameters must be specified. Then, a set of geostatistical realizations is generated for each of the eight scenarios. These are combined for a more realistic space of uncertainty.

Conditional probabilities are needed at every step of the scenario "tree." For example, in Figure 10.1, given the scenario M-I the probabilities for parameter sets P_1, P_2, and P_3 are 0.2, 0.5, and 0.3, reflecting the expert judgment that parameter set P_2 is the most probable given this conceptual model. These probabilities are necessarily subjective and must be specified by the asset team or relevant expert. The conditional probability values must sum to 1 at every level.

Recursive application of Bayes law is used to combine uncertainty at all levels, that is, assign the probability to each geostatistical realization. Recall Bayes law from basic statistics:

$$Prob\{A|B\} = \frac{Prob\{A \text{ and } B\}}{Prob\{B\}}$$

where A and B are events. $Prob\{A|B\}$ is the conditional probability of event A given that B has occurred, $Prob\{A \text{ and } B\}$ is the joint probability of A

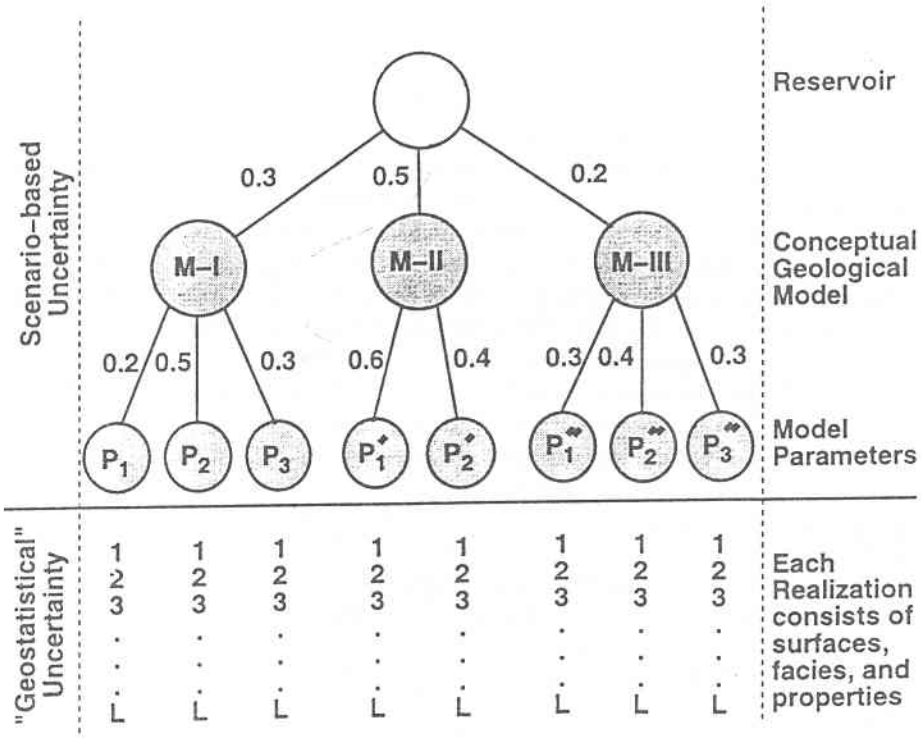


Figure 10.1: Schematic illustration of combining scenario-based and geostatistics-based models of uncertainty. A number of realizations, L, for each scenario is constructed. The probability of each realization is calculated by recursive application of Bayes relation.

and B happening together, and $Prob\{B\}$ is the marginal probability of event B occurring. Bayes law tells us that the joint probability at different levels is established by multiplication, for example, the probability of parameter set (P_1) and facies type one (I) would be:

$$\begin{aligned} Prob\{P_1 \text{ and } MI\} &= Prob\{MI\} \cdot Prob\{P_1|MI\} \\ &= 0.3 \cdot 0.2 = 0.06 \\ &= 0.06 \end{aligned}$$

This relation may be applied recursively for any number of scenarios and any level of complexity.

At the lowest level (Figure 10.1) there are L realizations. Each realization is equally probable given the set of parameters used to create it. Once again, by Bayes law the probability of each realization is calculated by multiplying the probability of each realization ($1/L$) by the probability of each scenario. The sum of all the probabilities is 1. The probabilities of the eight scenarios in Figure 10.1 are 0.06, 0.15, 0.09, 0.30, 0.20, 0.06, 0.08, and 0.06 from left

to right adding up to one. A combined distribution of uncertainty in any particular variable, such as oil in place or recovery, is assembled pooling all of the realizations with the correct probability. All summary statistics and probabilities are calculated using these probability weights.

A reasonable definition of scenarios and assignment of conditional probabilities is critical. The scenarios can reflect different aspects of uncertainty, for example, depositional style (estuarine channels versus tidal dominated), fault seal (sealing versus partial sealing), level of fracturing, and facies definition. The scenarios could also reflect uncertainty in critical statistical parameters such as the net-to-gross ratio (low, medium, and high) or the horizontal variogram range. There can be many levels to the scenario tree and a different number of levels down each branch of the tree. There could also be a different number of geostatistical realizations for each scenario.

The large-scale discrete aspects of uncertainty are quantified with the scenario-based approach. Uncertainty due to incomplete data is quantified with multiple geostatistical realizations. Reservoir uncertainty addressed with geostatistical realizations *alone* is reasonable late in the reservoir life cycle where there is little large-scale uncertainty; small-scale models of heterogeneity are adequate to reflect uncertainty. The next two Sections 10.2 and 10.3 address validating geostatistical models of uncertainty. This is appropriate before presenting procedures for using the resulting space of uncertainty (Sections 10.4/10.6).

10.2 Cross Validation and the Jackknife

Reservoir modeling consists of many interdependent modeling steps with many opportunities for mistakes. It is impossible to completely validate, in some objective sense, stochastic models. Nevertheless, there are some basic checks that can be used to identify problem data or errors in the geostatistical modeling. This Section describes some validation procedures for "estimation." Section 10.3 describes a complementary set of procedures to validate "simulation" models of uncertainty.

The basic idea is to estimate the attribute value (say, porosity) at locations where we know the true value. Analyzing the error values indicates the *goodness* of our modeling parameters [81]. In cross validation the actual data are deleted one at a time and re-estimated from the remaining neighboring data. The term jackknife applies to resampling without replacement, that is, when an alternative set of data values is re-estimated from another non-overlapping data set. The jackknife is a more stringent check since the non-overlapping data are not used to establish the statistical parameters such as the histogram and variogram. Also, the idea behind the jackknife is that it should be repeated with different non-overlapping data sets to filter out the statistical fluctuations of choosing one data set.

The jackknife requires a significant amount of CPU and professional effort to repeat the estimation procedure, including calculation and fitting of all var-

iograms, with different sets of data; therefore, the jackknife is not commonly used. Implementation of cross validation requires reasonable care to make estimation as difficult as it will be in practice. For example, each well must be removed in turn; removal of each sample individually would lead to unrealistically good results due to very close adjacent samples. Cross validation is only practical with some reasonable number of wells, say five; otherwise, there are too few data for checking.

The result of cross validation and the jackknife is a set of true values and paired estimates, $\{z(\mathbf{u}_i), z^*(\mathbf{u}_i), i = 1, \ldots, n\}$. The first check is to cross plot the true values versus the estimated values. The points should fall as close to the 1:1 line as possible with maximum correlation coefficient. The problem estimates or data to investigate further are those in the upper left and lower right corners where the true high values are estimated low, or true low values are estimated high. The location of these data and the details of estimation should be checked.

An error value can also be calculated for each re-estimated location:

$$e(\mathbf{u}_i) = z^*(\mathbf{u}_i) - z(\mathbf{u}_i) \tag{10.1}$$

these errors can be analyzed in different ways (1) the distribution of the errors $\{e(\mathbf{u}_i), i = 1, \ldots, n\}$ should be symmetric, centered on a zero mean, with a minimum spread. (2) the mean-squared error $MSE = 1/n \sum_{i=1}^{n} [e(\mathbf{u}_i)]^2$ should be minimum, (3) the plot of the error $e(\mathbf{u}_i)$ versus the estimated value $z^*(\mathbf{u}_i)$ should be centered around the zero-error line, a property called "conditional unbiasedness," and (4) the n errors should be independent from each other. This can be checked by contouring or posting the error values. Areas or zones of systematic over- or underestimation are a problem and, perhaps, the trend model should be looked at more carefully.

The results of cross validation must be used with care. Cross validation does not prove the optimality of a given estimation scheme. The analysis of re-estimation scores does not consider the *multivariate* properties of the estimation algorithm, that is, of the estimates taken all together. Cross validation results are also insensitive to the variogram at distances less than the sample spacing, which is an important unknown aspect of estimation. Cross validation should be used to check for mistakes and identify problem data. The techniques to check estimation are straightforward. In the context of simulation, however, there is no single estimate; there are a potentially large number of simulated values. A different approach must be considered.

10.3 Checking Distributions of Uncertainty

It is impossible to rigorously validate a model when applied to unknown values [204]. The true and unknown reality will certainly not follow our relatively simplistic probabilistic models and may not be reliably informed by the data we have at hand. Nevertheless, we can use cross validation to assess geostatistical models of uncertainty.

The result of cross validation and the jackknife applied in a simulation setting is a set of true values and paired distributions of uncertainty, $\{z(\mathbf{u}_i), F_Z(\mathbf{u}_i; z), i = 1, \ldots, n\}$. Different modeling decisions would lead to different distributions of uncertainty, that is, different $F_Z(\mathbf{u}_i; z)$'s. These local distributions or ccdf models may be (1) derived from a set of L realizations, (2) calculated directly from indicator kriging, or (3) defined by a Gaussian mean, variance, and transformation. Checking distributions of uncertainty or comparing alternative approaches requires an assessment of the *goodness* of distributions of uncertainty.

In the context of evaluating the goodness of a probabilistic model, specific definitions of accuracy and precision are proposed. For a probability distribution, accuracy and precision are based on the actual fraction of true values falling within symmetric probability intervals of varying width p:

- A probability distribution is accurate if the fraction of true values falling in the p interval exceeds p for all p in $[0, 1]$.
- The precision of an accurate probability distribution is measured by the closeness of the fraction of true values to p for all p in $[0, 1]$.

A procedure for the direct assessment of local accuracy and precision is now described. In checking the goodness of a probabilistic model, we can compare alternative models, and perhaps fine-tune the parameters of a chosen model.

Direct Assessment of Accuracy and Precision

The first step is to calculate the probabilities associated to the true values $z(\mathbf{u}_i), i = 1, \ldots, n$ using the model of uncertainty:

$$F^*(\mathbf{u}_i; z(\mathbf{u}_i)|n(\mathbf{u}_i)), \quad i = 1, \ldots, n$$

For example, if the true value at location $\mathbf{u}_i$ is at the median of the simulated values then $F(\mathbf{u}_i; z(\mathbf{u}_i)|n(\mathbf{u}_i))$ would be 0.5.

Then, a range of symmetric p-probability intervals (PIs), say, the centiles 0.01 to 0.99 in increments of 0.01 can be considered. The symmetric p-PI is defined by corresponding lower and upper probability values:

$$p_{low} = \frac{(1-p)}{2} \quad \text{and} \quad p_{upp} = \frac{(1+p)}{2}$$

For example, for $p = 0.9$, $p_{low} = 0.05$ and $p_{upp} = 0.95$.

Next, define an indicator function $\xi(\mathbf{u}_i; p)$ at each location $\mathbf{u}_i$ as:

$$\xi(\mathbf{u}_i; p) = \begin{cases} 1, & \text{if } F(\mathbf{u}_i; z(\mathbf{u}_i)|n(\mathbf{u}_i)) \in (p_{low}, p_{upp}] \\ 0, & \text{otherwise} \end{cases} \tag{10.2}$$

The average of $\xi(\mathbf{u}_i; p)$ over the n locations $\mathbf{u}_i$:

$$\overline{\xi(p)} = \frac{1}{n} \sum_{i=1}^{n} \xi(\mathbf{u}_i; p) \tag{10.3}$$

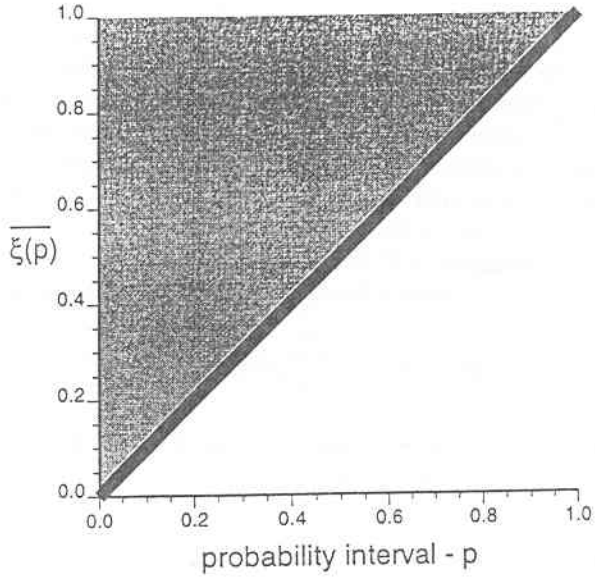


Figure 10.2: Blank (template) accuracy plot: the actual fraction of true values falling within a probability interval p, denoted $\overline{\xi(p)}$, is plotted versus the width of the probability interval p.

is proportion of locations where the true value falls within the symmetric p-PI.

According to our earlier definition of accuracy, the simulation algorithm used to generate the distributions of uncertainty (ccdfs) is accurate when $\overline{\xi(p)} \geq p$, $\forall\, p$. A graphical way to check this assessment of accuracy is to cross plot $\overline{\xi(p)}$ versus p and see that all of the points fall above or on the 45°line. This plot is referred to as an *accuracy plot*. A blank accuracy plot is shown in Figure 10.2. The ideal case is where points fall on the line, that is, the probability distributions are accurate and precise. Points that fall above the line (in the shaded region) are accurate but not precise. Finally, points that fall below the line are neither accurate nor precise.

A distribution is accurate when $\overline{\xi(p)} \geq p$. Precision is defined as the closeness of $\overline{\xi(p)}$ to p when we have accuracy ($\overline{\xi(p)} \geq p$). Quantitative measures of accuracy and precision could be defined; however, visual inspection of the accuracy plot is adequate. The results are acceptable if the points fall close to the 45 degree line. The distributions of uncertainty are too wide when the points fall above the 45 degree line and too narrow when the points fall below the 45 degree line. There are different ways to increase or decrease the spread of the distributions of uncertainty. One way is to increase the spatial correlation, which will decrease the spread of the local distributions of uncertainty. Decreasing the range of spatial correlation will increase the width of the local distributions of uncertainty.

Good probabilistic models must be both accurate and precise: however, there is another aspect of probabilistic modeling: the spread or uncertainty represented by the distributions. For example, two different probabilistic models may be equally accurate and precise and yet we would prefer the model that has the least uncertainty. Subject to the constraints of accuracy and precision, we want to account for all relevant data to reduce uncertainty. For example, as new data become available our probabilistic model may remain equally good and yet there is less uncertainty.

The overall uncertainty of a probabilistic model may be defined as the average conditional variance of all locations in the area of interest:

$$U = \frac{1}{n} \sum_{i=1}^{N} \sigma^2(\mathbf{u}_i) \tag{10.4}$$

where there are n locations of interest $\mathbf{u}_i, i = 1, \ldots, n$ with each variance $\sigma^2(\mathbf{u}_i)$ calculated from the local ccdf $F(\mathbf{u}_i; z|n(\mathbf{u}_i))$.

Reservoir Case Study

To further illustrate the direct assessment of accuracy and precision, consider the "Amoco" data consisting of 74 well data related to a West Texas carbonate reservoir. Figure 10.3 shows a location map of the 74 well data and a histogram of the vertically averaged porosity for the main reservoir layer of interest. The vertically averaged porosity data were transformed to a Gaussian distribution and all subsequent analysis is performed with transformed data. This transform is reversible so the results would be the same were we to consider back transform porosity values.

A number of procedures were checked for building distributions of uncertainty at the 74 well locations. The variable being considered was the vertical average of porosity over the vertical thickness of the reservoir. A "leave-one-out" cross validation approach was considered. A number of alternative geostatistical methods were considered to build distributions of uncertainty in the vertical average of porosity:

1. Gaussian distribution of uncertainty built by kriging of normal score values with the variogram modeled using the complete 3-D set of normal score porosity data. Figure 10.4 shows the variogram and accuracy plot. The variogram of the vertical average data should have been considered: this variogram applies to the small scale data and is an incorrect choice for the 2-D vertical average. The average uncertainty in this case is 0.757 (relative to 1.0 for no local information).

2. Gaussian distribution of uncertainty built by kriging of normal score values with isotropic variogram from the set of normal score transforms of 2-D average porosity. Figure 10.5 shows the variogram and accuracy plot. The average uncertainty is 0.433, which is significantly better that the incorrect choice of the 3-D variogram.

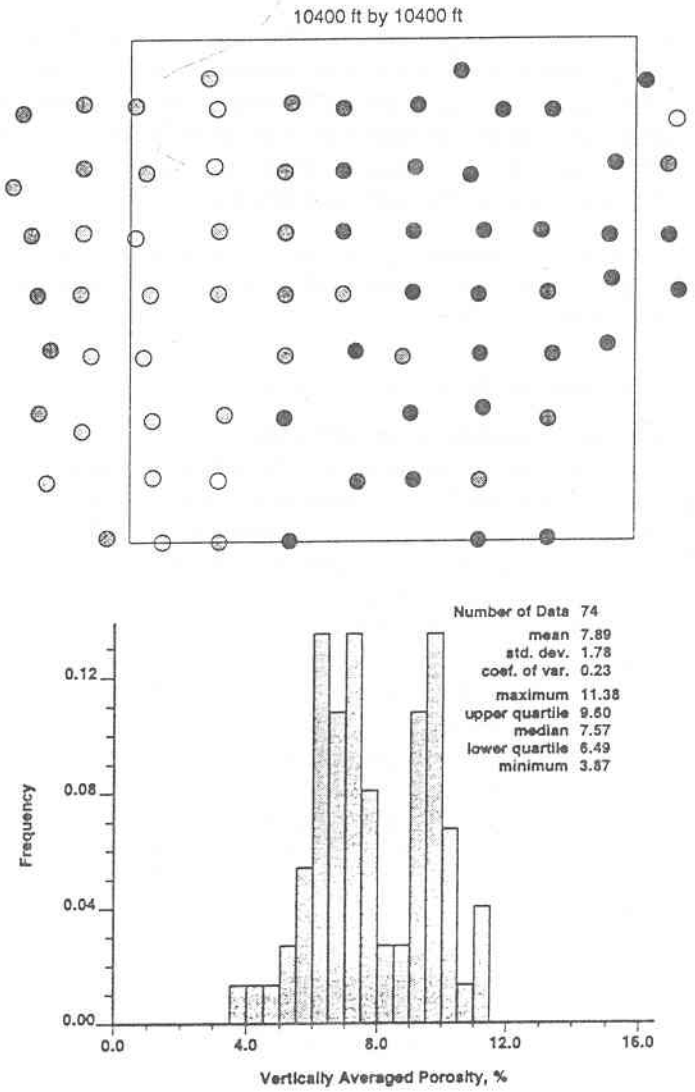


Figure 10.3: Location map and histogram of 74 well data.

3. Gaussian distribution of uncertainty built by kriging of normal score values with an anisotropic variogram model built from the 74 vertically averaged values. Figure 10.5 shows the variogram and accuracy plot. The average uncertainty is 0.385, which is slightly better than the less correct choice an isotropic variogram.

4. An indicator-based approach (see Chapter 8 for details given in the context of permeability) with nine thresholds reduces uncertainty to 0.244 with reasonable accuracy. This improvement in results is at the cost of greater effort: professional time to calculate and model nine anisotropic indicator variograms rather than just one and nine times the CPU effort for the full indicator kriging.

The accuracy plot and uncertainty can be used together to assess local distributions of uncertainty and compare alternatives. The indicator approach appears suited to this example.

Checking Categorical Variable Distributions

The discrete nature of categorical variables and the lack of an ordering require different cross validation techniques than continuous variables. The prediction of facies may be checked by (1) closeness of the estimated probabilities to the true facies, and (2) accuracy of the local probabilities. Cross validation in a categorical variable setting leads to a set of true facies types and predicted probabilities, that is, $\{s(\mathbf{u}_i), p^*(\mathbf{u}_i), k = 1, \ldots, K\}, i = 1, \ldots, n$ where K is the number of facies and n is the number of data locations.

All variants of sequential indicator simulation provide direct estimates of the facies probabilities at the locations being (re)estimated. The cross validation of object-based methods is quite tedious; the facies modeling must be repeated for every well to be (re)estimated. The distributions of uncertainty must be built up directly by simulation without taking the shortcut of kriging indicator transforms. Although tedious, this procedure permits a quantitative comparison with other methods and an ability to fine-tune the object shape and size parameters.

The predicted probabilities, $p^*(\mathbf{u}_i), k = 1, \ldots, K, i = 1, \ldots, n$ can be directly calculated for truncated Gaussian simulation without repeated simulation. The distribution of uncertainty of the Gaussian variable and knowledge of the locally variable thresholds allow us to calculate these probabilities analytically. The predicted probabilities may be checked in the same way as the direct estimates from indicator kriging methods and the simulation result of object-based methods.

The challenge is to summarize the results in a way that permits easy comparison. Three measures are suggested:

1. *Closeness to the true facies* measures how large the predicted probabilities of the true facies are. Ideally, the probability of the true facies $p^*(\mathbf{u}_i; s(\mathbf{u}_i))$ should be close to one.

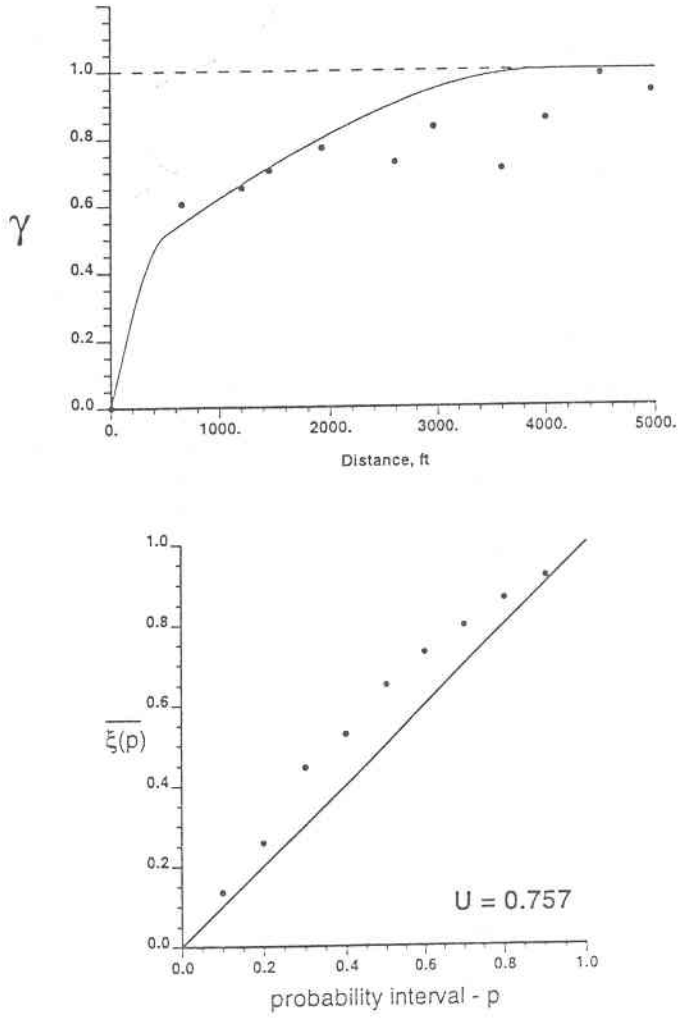


Figure 10.4: Horizontal normal scores semivariogram from the 3-D porosity data (calculated using stratigraphic coordinates) and the corresponding accuracy plot considering cross validation with the normal scores of the 74 well data in Figure 10.3.

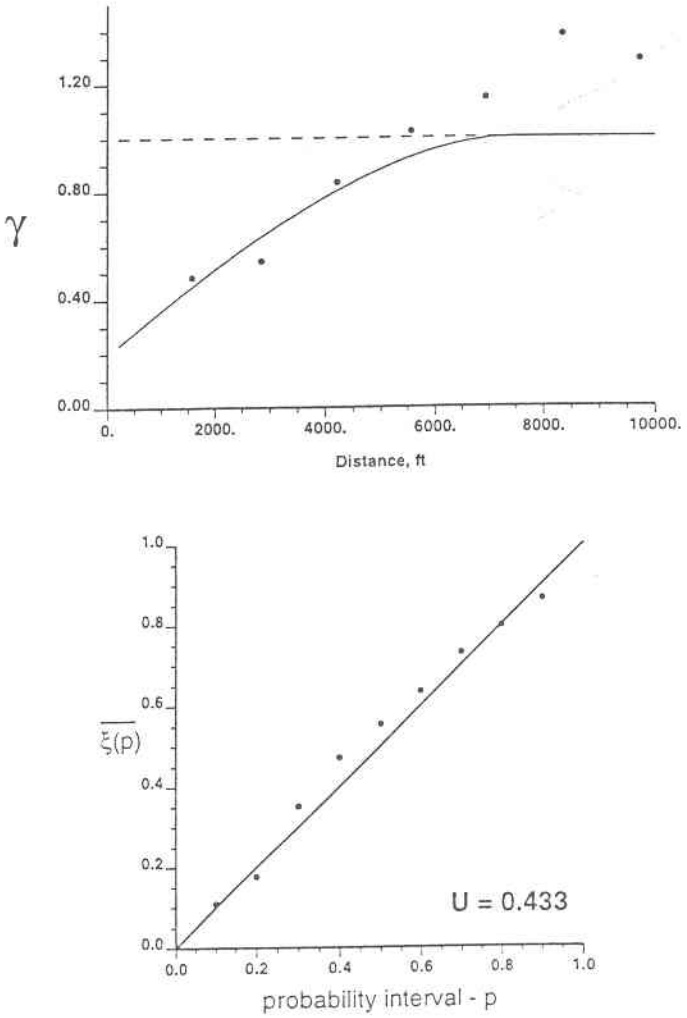


Figure 10.5: Horizontal normal scores semivariogram from the vertically averaged porosity and the corresponding accuracy plot considering cross validation with the normal scores of the 74 well data in Figure 10.3.

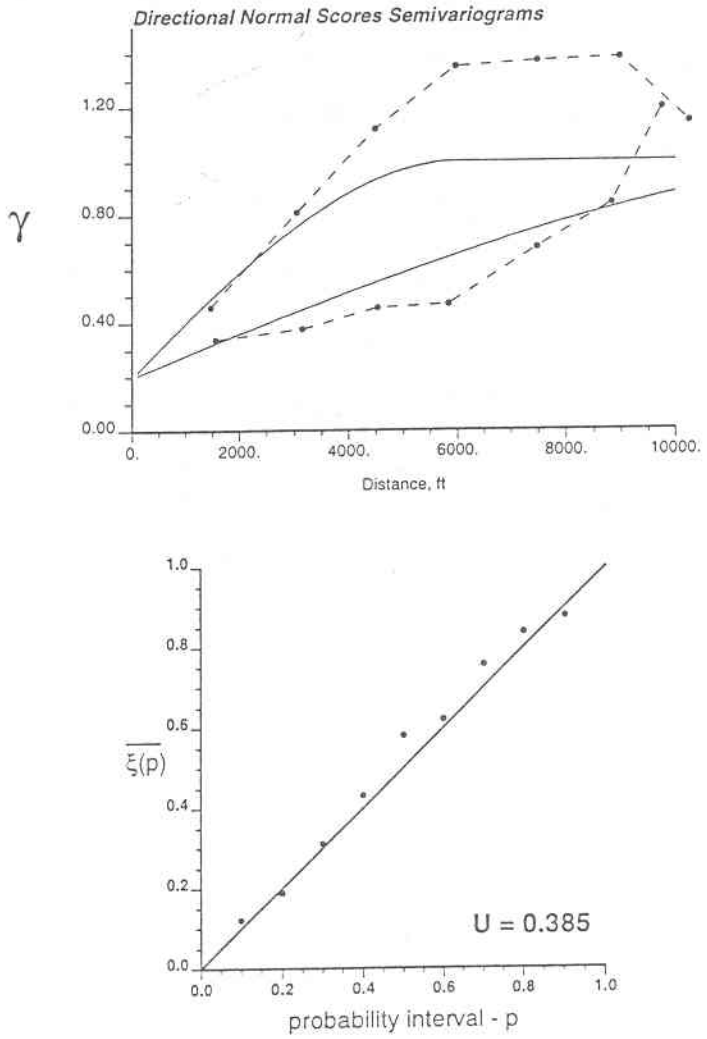


Figure 10.6: Directional variograms from vertically averaged porosity and accuracy plot.

2. *Fuzzy closeness to the true facies* accounts for the fact that the consequences of misclassification are different, that is, assigning the wrong type of sand is less consequential that assigning a sand facies as non-net shale.

3. *Accuracy of the local probabilities* measures the reliability of the predicted probabilities, that is, 80% of the locations assigned an 80% chance to be facies s_k *should be* facies s_k.

A quantitative measure of closeness to the true probabilities may be summarized by:

$$C = \frac{1}{N} \sum_{i=1}^{N} \{p(\mathbf{u}_i; k) | \text{ true } = k\} \tag{10.5}$$

This is interpreted as the average predicted probability of the true facies. This probability would be 1.0 in an ideal case of complete information and $1/K$ in the case where no information is available.

The closeness measure C described above quantifies how "right" the probabilities are for the true facies; however, it gives no credit when a similar facies is predicted nor does it penalize the case when drastically different facies are predicted. A closeness matrix could be defined to account for these consequences. The closeness value $c(i, j)$ specifies how close facies i is to facies j. By definition the closeness $c(i, j)$ is one for $i = j$, the same facies, and $c(i, j)$ is zero for completely different facies, say i = clean channel sand and j = shale. The exact values of $c(i, j)$ must be determined qualitatively considering the impact of misclassification. The need for a qualitative assignment is a shortcoming; however, some means must be used to account for the fuzzy closeness of facies.

The third check is the accuracy of the local probabilities. The local probabilities are "accurate" or "fair" if they accurately reflect the true fraction of times the predicted facies occurs. For example, consider all locations where we predict a 65% probability of facies 1 ($p_1^* = 0.65$); then 65% of those locations should be in facies 1. There is nothing special about 0.65; we check a range of predicted probabilities. Cases where the actual fraction departs significantly from the predicted probability are a problem. These results may be plotted like the accuracy plot discussed above.

Five wells used for a small example to check facies prediction. Seismic data is available to improve the predicted models of uncertainty. Indicator kriging (IK), Bayesian updating (BU), and a full block cokriging (BC) of the facies using the seismic data as a secondary data were considered. The closeness statistic C is 0.320, 0.358, and 0.349 for the three methods (note that a statistic of 1.0 would be ideal). The Bayesian updating approach appears best. We have an indication that the block cokriging approach is more sensitive to the larger number of input parameters. Figure 10.7 shows a cross plot of actual probabilities versus predicted probabilities.

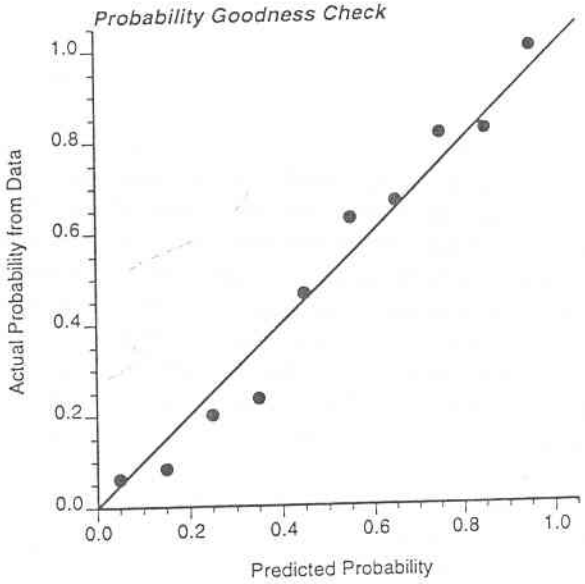


Figure 10.7: Example cross plot of actual probabilities versus the predicted probabilities. The closeness of the results to the 45 degree line attests to the goodness of the probabilities. These results are for block cokriging in all facies. We could also look at the results on a by-facies basis.

The main uses for the diagnostic tools presented here are: (1) detecting implementation errors, (2) quantifying uncertainty, (3) comparing different simulation algorithms (for example, Gaussian-based algorithms versus indicator-based algorithms versus simulated annealing based algorithms), and (4) fine-tuning the parameters of any particular probabilistic model (for example, the variogram model used). These tools provide basic checks, that is, necessary but not sufficient tests. They do not assess the multivariate properties of the simulation. Care is needed to ensure that features that impact the ultimate prediction and decision making, such as continuity of extreme values, are adequately represented in the model.

10.4 How Many Realizations?

An important contribution of geostatistical reservoir modeling is a framework for uncertainty assessment through Monte Carlo sampling from the space of uncertainty in reservoir responses. Monte Carlo sampling proceeds by (1) drawing L realizations from a probabilistic model, (2) processing the L realizations through some performance calculation, and (3) assembling a histogram of the L responses to represent a distribution of uncertainty in the output(s). Classical Monte Carlo simulation requires the L realizations to

be drawn randomly; therefore, they each go into the distribution of uncertainty with equal probability. The question of how many realizations must be addressed.

Significant CPU time is required to process realizations through a flow simulator, which is the most common performance calculator. Often, a small number of realizations are considered for this reason alone. The following discussion on the required number of realizations is for those cases that we *can* consider more realizations for a better assessment of uncertainty.

The required number of realizations depends on the (1) aspect of uncertainty or the "statistic" being quantified, and (2) the precision with which the uncertainty assessment is required. Few realizations may be required to assess an average statistic such as the average porosity. A large number of realizations are required, however, to assess the 1% and 99% percentiles of the oil in place distribution.

A quantitative basis must be established to determine the required number realizations. The reporting procedures of environmental standards and political opinion polls will help us define the basis for the results of geostatistical analysis. Consider a poll to answer the question "Have recent price increases for energy caused any financial hardship for you or your household?" The answer would be reported as percentages of yes and no responses, for example, 56% yes and 44% no. Responsible polling agencies would add a caveat to convey the uncertainty in this result. They may report the number of respondants or, more likely, they would apply basic statistics to more completely communicate uncertainty in the response, for example:

> The true percentage is within 3% of this reported result (56% yes) 18 times out of 20.

These two additional numbers are used to summarize uncertainty in the reported result. The left side of Figure 10.8 shows a schematic example of the uncertainty in the statistic of 56%. If more people were polled (analogous to more realizations being generated) the uncertainty would be less, see right side of Figure 10.8. This notion of "uncertainty in reported statistics" will be extended to quantify the uncertainty in our results for a fixed number of realizations.

The reported statistics from geostatistical analysis are particular quantiles of some response variable (for example, oil in place), say, the $F(p) = 0.1, 0.5,$ and 0.9 quantiles. As introduced above, two tolerance parameters are also required: (1) a tolerance for the quantiles, say $\Delta_F = \pm 0.01$, and (2) the minimum probability of being within probability Δ_F, say $t_F = 0.8$ or 80%. A typical case would be to require the P_{10}, P_{50}, and P_{90} of recovery factor within 2%, 80% of the time. A more stringent requirement would be the P_1 and P_{99} of net present value within 0.1%, 95% of the time ($\Delta_F 0.001$ and $t_F = 0.95$).

Thus, two parameters are used: (1) Δ_F, which is the deviation, in units of cumulative probability, from the reported quantile, and (2) t_F, which is

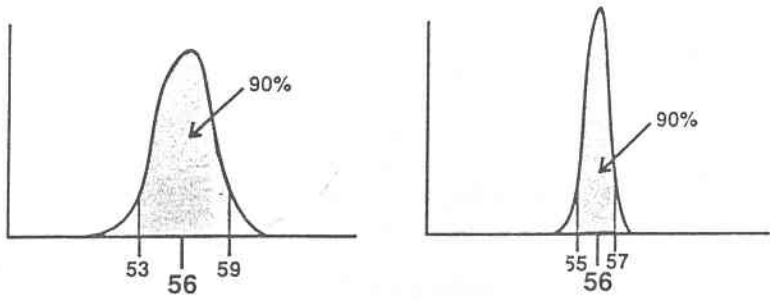


Figure 10.8: Uncertainty in a reported statistic of 56%. The left side figure corresponds to the example in the text where the true value is reported to be within 3% of this number 18 times out of 20 (90% of the time). The right side figure shows reduced uncertainty if a larger number of samples were taken (more realizations available) for uncertainty reporting.

the fraction of times the true value will fall within $\pm$ of the reported statistic. In this context, we can derive the required number of realizations L to meet a certain specified standard, that is, a specified Δ_F and t_F. Alternatively, the parameters Δ_F and t_F can be established for a given L, that is, it is possible to calculate the uncertainty of our results due to a small number of realizations.

The number of realizations L to meet specified Δ_F and t_F can be calculated. The sampling distributions for the cdf values tend to normal distributions. This is not surprising since the cdf value is the sum of a large number of values (the indicator transform at the correct threshold) that are independent (the realizations are random) and identically distributed. The central limit theorem tells us that the sampling distribution in this case tends toward a normal distribution as the number increases. The number we are considering (L) is large for the central limit theorem; therefore, it is expected that the distribution will be normal. This has been verified numerically.

The mean and variance of the sampling distribution for a cdf value (say, the first quantile of interest $F = 0.1$) can be determined theoretically. The sampled cdf value F^* is the average of an indicator value:

$$F^* = \frac{1}{L}\sum_{l=1}^{L} i(u_l)$$

where the indicator function is 1 if random drawing u_l is less than or equal to the value F and 0 otherwise. The mean or expected value of F^* is the true underlying cdf value, that is, F. The variance of F^* is calculated as:

$$\sigma^2_{F^*} = \frac{1}{L}\sigma^2_I = \frac{1}{L}E\left\{I^2\right\} - E\left\{I\right\}^2 = \frac{1}{L}F - F^2 = \frac{F(1-F)}{L}$$

This analytical result for the uncertainty of a cdf/quantile has been verified by numerical results. The simplicity of these results makes it straightforward to calculate the t_F statistic for specified L, Δ_F, and F values.

$$\begin{aligned} t(L, \Delta_F, F) &= Prob\{F^* \geq (F - \Delta_F)|L\} - Prob\{F^* < (F + \Delta_F)|L\} \\ &= G\left(\frac{\Delta_F}{\sqrt{F(1-F)/L}}\right) - G\left(\frac{-\Delta_F}{\sqrt{F(1-F)/L}}\right) \\ &= 2 \cdot G\left(\frac{\Delta_F}{\sqrt{F(1-F)/L}}\right) - 1 \end{aligned} \tag{10.6}$$

There are numerous $G(y)$ functions available, for example, the gcum and ginv functions in GSLIB. This is an important result. The precision for a given number of realizations can be directly calculated. We can invert this relationship to get an equation that gives us the number of realizations L to achieve a certain precision specification:

$$L = \frac{F(1-F)}{\left(\frac{\Delta_F}{G^{-1}(t+1)/2)}\right)^2} \tag{10.7}$$

The number of realizations for a specified precision can be calculated directly.

Figure 10.9 shows curves of t_F versus the number of realizations L for Δ_F = 0.1, 0.2, 0.3, 0.4, and 0.5. These five curves were also calculated numerically and the results match. These curves relate to the 0.5 quantile. These relations tell us "how many realizations are needed for a specified precision" and "how good are probability values for a given number of realizations?"

Practical Results: How Many Realizations?

The analytical result presented above should be used if there is a requirement for quantitative measures of uncertainty. This may lead to an unrealistically large number of realizations. In general, a staged approach should be considered.

- Enumerate the geological scenarios to be considered and generate a single geostatistical realization of all variables for each scenario. Validate that each realization honors the input data within acceptable statistical fluctuation. Calculate the response variables (oil in place, flow performance, and so on) with each realization.

- Generate five realizations for all important scenarios. Important scenarios have a large probability of occurrence or response variables that are unusually low or high. The expert probabilities assigned to each scenario and the response variable from the first realization must be used to make this judgement. Calculate the response variables with these realizations.

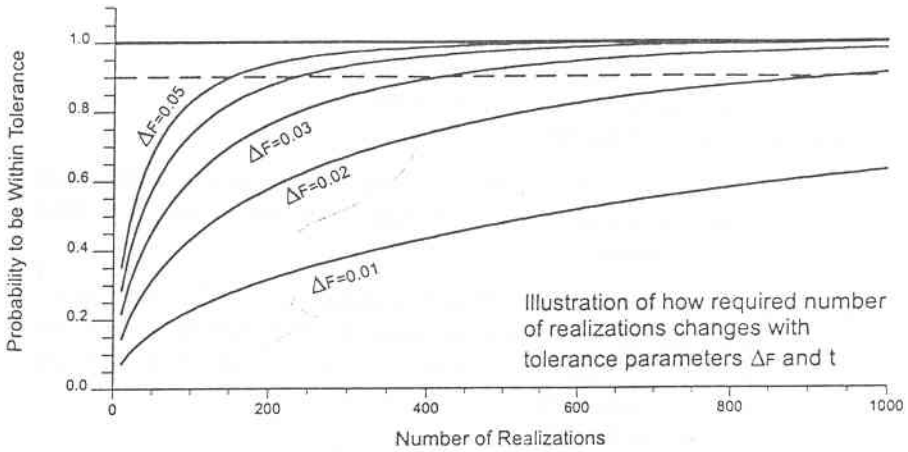


Figure 10.9: Illustration of how the t parameter is related to the number of realizations for $\Delta_F = 0.1$, 0.2, 0.3, 0.4, and 0.5. The precision of the uncertainty assessment increases as more realizations are considered.

- Consider more realizations for those scenarios where the first five random realizations are quite different from each other. The response variables could be calculated on all of these additional realizations.

Calculating the response variables of interest can involve running a flow simulator at significant CPU and professional cost. The generation of geostatistical realizations is relatively quick. This leads to the idea of generating a large number of realizations, ranking them by some fast-to-calculate statistic, and then processing a limited number through the full flow simulator.

10.5 Ranking Realizations

Each geostatistical realization reproduces the input data equally well and is equally probable given the scenario or random function model under consideration. Nevertheless, the realizations may be ranked according to some criteria or measure not used as data. For example, the volume of hydrocarbon-filled pore space connected the well locations can be used to rank a large number of geostatistical realizations.

There is no unique ranking index when there are multiple flow response variables and no ranking measure is perfect. Nevertheless, the value of ranking realizations will be a reduction in the number of realizations that must be processed to arrive at the same level of uncertainty.

A good ranking statistic correctly identifies low and high realizations. Before describing a number of ranking measures, consider some cases where ranking is problematic or unnecessary:

1. when each realization leads to nearly the same answer;

2. when the aspect of uncertainty being assessed is easy to calculate, for example, the uncertainty in oil-in-place may be simply assessed by calculating the pore volume of all realizations;

3. when there are many independent reservoir responses of interest, that is, it is impossible to conceive of a ranking index that would lead to a unique reliable ranking.

There are times, however, when significant professional or CPU time is required to evaluate each realization and the number of realizations considered must be limited. In these situations, it is worthwhile to consider ranking the realizations to limit the number of fine-scale simulations and yet to obtain an idea of uncertainty in the flow response.

Most simple ranking measures require each cell in all geostatistical realizations to be assigned a *net* indicator. This indicator is one if the cell is reservoir quality and zero otherwise. Reservoir quality is achieved in certain facies by porosity and permeability simultaneously above specified thresholds. The net-to-gross ratio for each realization can be easily calculated and used as a simple ranking measure. Another simple ranking measure is the total oil volume, that is, the product of the net-to-gross indicator, the oil saturation, and the pore volume. This accounts for porosity and saturation variations.

Flow performance depends on the connectivity of the reservoir-quality rock; there are many measures of connectivity that may be quickly calculated without running a full simulator [18, 65, 23]. A first measure of connectivity is available by determining the sets of *net* geological modeling cells that are connected in 3-D space. There are fast algorithms to scan through 3-D binary net/non-net realizations aggregating those reservoir-quality cells that are connected net cells. The result will be a 3-D specification of the number N_{geo} of geo-objects or connected objects each with an associated volume $V_{geo,j}, j = 1, \ldots, N_{geo}$. Figure 10.10 shows the geo-objects for three SIS facies realizations.

The geo-objects of a set of realizations can be used in a number of ways to rank the realizations:

- The fraction of cells within the first, say, 5, geo-objects could be used as a ranking measure. A realization is more "connected" when this fraction is large.

- The connected hydrocarbon volume within some radius of the production wells could be considered if the well locations are known. This ranking measure accounts for local information about connectivity.

- The connected volume between "pairs" of injection and production well pairs could be used if the well pair locations are known.

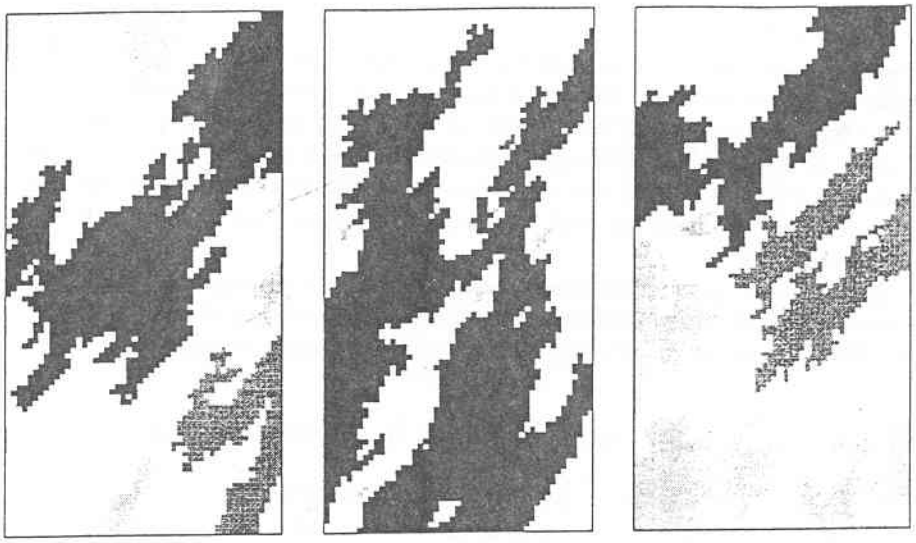


Figure 10.10: Slices through three facies models. The black region on each realization is the largest geo-object. Smaller geo-objects are shaded in gray.

An alternative use of geo-objects is for the selection of well locations. An optimization scheme could be devised to select the well locations that maximizes the geo-object volume within some drainage radius.

Geo-objects measure static connectivity, that is, they do not account for tortuosity, permeability, attic oil, and the interaction between multiple producing well locations. There are some more sophisticated alternatives that measure dynamic connectivity and, yet, are still simpler than running the full flow simulator. There are random walk algorithms that measure "dynamic" continuity between injecting and producing locations. These methods often call for a solution to the pressure field (single phase flow) given assumed well rates. Particles are then tracked through the media and the distribution of "times" or "lengths" between injecting and producing wells provides a measure of connectivity that could be used for ranking.

There are other relatively simple and fast flow models including (1) tracer simulation, (2) simulation based on a network of 1-D streamlines, and (3) a water flood simulation in lieu of a more complex miscible or compositional-type simulation. The time of 5% watercut is likely a good measure for the breakthrough of other miscible components in a more complex process.

Another ranking approach is to use the correct physics or flow equations but with the geological models scaled to such a coarse resolution that the computer time is acceptable. The coarseness of the underlying grid will compromise the direct usefulness of the results. Nevertheless, the relative ranking of the results may be used to rank the underlying geological realizations.

Practical Result: Ranking

Ranking for quantitative statements of probability should be avoided. Published applications of ranking consider the ranking measure to be "quantile preserving," for example, the P_{10} of the full performance assessment is determined by running the full assessment on the P_{10} of the ranked realizations. There is no statement of how close the resulting P_{10} is to the true value. This procedure will work when the correlation between the "rank of the ranking measure" and "rank of the full assessment" is large; however, that is not usually the case.

Ranking can be used to select realizations to get a rough idea of "low," "expected," and "high" performance. Flow-based ranking measures can be used if tools and experience exist for their reliable application.

10.6 Decision Making with Uncertainty

Making the best decision in face of geological uncertainty is an important and complicated topic. Some general ideas are given below; the reader will be required to go to the literature for more information.

Petroleum exploration and production are inherently risky activities. Decisions regarding those activities depend on forecasts of future hydrocarbon production revenue. Such forecasts are uncertain because of (1) uncertainty about the reservoir geometry and the spatial distribution of petrophysical properties, (2) uncertainty about the fluid properties, (3) uncertainty about the actual behavior of the rock and fluid when subjected to external stimuli, (4) model limitations, and (5) uncertainty about costs and future prices. This book addresses the uncertainty of the geological model due to sparse sampling of the reservoir.

Exploration and production are sequential. First, exploration finds a promising geologic structure using knowledge of the sedimentary basin and seismic data. Then, a well is drilled to prove the existence of a hydrocarbon reservoir. Other exploratory wells may be drilled to delimit the reservoir, depending on the field size. Next, a development plan is generated to provide the necessary data for the production cash flow analysis. If the company decides to invest in that project, the development plan is implemented and hydrocarbons are produced.

There are three main types of decisions involved in exploration and production: (1) the decision of whether or not to drill an exploratory well, (2) the technical decision of selecting the best development plan to optimize the profitability of the reservoir production, and (3) the business decision to invest in a project or not. Decision making in exploration typically accounts for uncertainty through decision tables which relate alternative actions to various outcomes [123].

After exploratory reservoir delimitation, a reservoir development plan is devised that determines the number, type, and location of additional wells

and presents the rig work schedule and the curves for injection and production of fluids. Once the development plan is defined, it is possible to transfer some aspects of data uncertainty to the production forecasts. The probability distribution of flow responses can be used to evaluate the expected monetary value of the project and to guide the business decision of whether or not to invest in a particular project.

Accounting for geological uncertainty in the selection of the best development plan will be presented briefly. This is by far the most important decision in reservoir management.

The "conventional" approach to define the development plan ignores geological uncertainty. The conventional procedure consists of (1) building a single deterministic geological model of the reservoir, (2) defining some different possible production scenarios (numbers of wells, configuration for each number of wells, vertical or horizontal, producer or injector, fluid to inject, and so on), (3) running a flow simulator for each scenario to generate the respective production and injection curves, (4) performing a cash flow analysis for each scenario and (5) selecting the scenario that provides the maximum profit.

Well Site Selection

A number of approaches have been developed for well site selection: (1) integer programming using geo-objects [255], (2) optimization combined with the results of simulating with one well at a time [42], and (3) experimental design and response surface methodology [1, 51, 258]. This optimization problem is particularly difficult due to the large combination of well locations.

Development Plans

Optimization of the reservoir development plan is also particularly difficult due to the large combination of parameters to select. The flow simulator has been combined with classical optimization procedures for a single deterministic geological model [26]. Other optimization procedures have been used to account for multiple realizations [43]. Experimental design and response surface methodology have been applied [45, 83, 142] to obtain the distribution of flow responses for each scenario and to retain the best scenario in presence of uncertainty; however, the consideration of a reduced number of realizations required by those methodologies yields an incomplete assessment of the uncertainty in the flow responses [238].

A comprehensive study was performed by Paulo Sérgio da Cruz [43]. A general procedure for decision making in presence of geological uncertainty was presented:

1. Generate L geostatistical realizations of the geological model $l = 1, ..., L$. The notation for the geological model "l" is intentionally simple but actually l is a spatially distributed vector of numerical models representing structural surfaces, facies, porosity, permeability and fluid saturations.

2. Define the possible reservoir management scenarios: $s = 1, ..., S$. Each scenario is a complete specification of one possible solution for the problem. For example, one scenario would define the number of wells, their locations, the completion intervals, the surface facilities. The number of scenarios could be very large, but an inspection of the L realizations and prior sensitivity flow analysis based on just one realization should reduce this number substantially. Note that these scenarios are *not* the geological scenarios discussed at the beginning of the chapter.

3. Establish a quantitative measure of profit P to be maximized. The measure of profit would increase with increased hydrocarbon production and would decrease as more wells and facilities are required. The profit depends on the related costs, hydrocarbon prices, and taxes. A good unit to measure the profit is the present value of the discounted cash flow.

4. Calculate the profit for each scenario and each realization: $P_{s,l}$, $s = 1, ..., S$, $l = 1, ..., L$. The fluid production and injection curves are obtained by running a flow simulator, and the profit measure is calculated from the scenario specifications and curves for each case (s and l).

5. Calculate the expected profit $\overline{P}_s$ for each scenario, based on the average profit over all realizations.

6. Define the optimal scenario s^* as the scenario that has the maximum optimal estimate of profit $\overline{P}_s$.

Implementation of this procedure requires care in the specification of the scenarios and the execution of many flow simulations.

Loss Functions

In the preceding Section, the optimal scenario was defined as the one that maximized the expected value of the profit over the geostatistical realizations. Taking the expected value or average profit implicitly assumes that the consequences of under- and overestimation are the same, more specifically, a least squares criterion is used. In practice, however, the consequences of under- and overestimation are rarely the same and rarely follow a least squares relationship.

The idea of a "loss function" quantifies the notion of different "loss" or consequences for under- and overestimation. The unknown true profit P will be estimated by a single value p^* with a given error $e = p^* - P$. The function $Loss(e)$ must be specified by the organization or person in charge of the economic decisions in the company. The error e is not known, however,

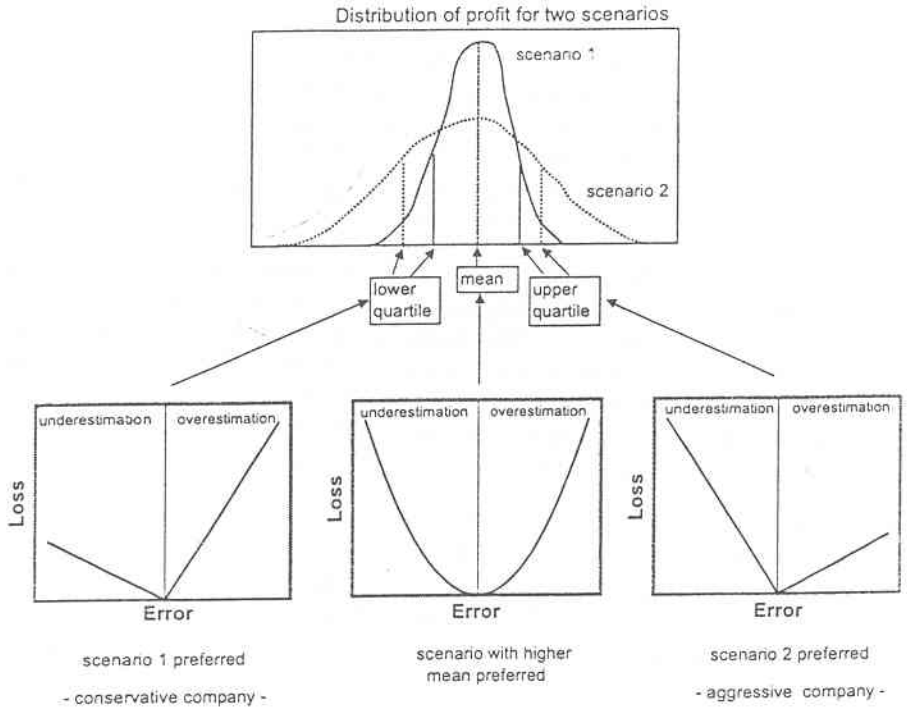


Figure 10.11: Example of probability distribution of profit for two scenarios and three types of loss function that yield different values of the retained profit value for each scenario and different decisions about the best scenario. Modified from Paulo Sérgio da Cruz [43].

because the true value P is not known; therefore, an expected loss value for each scenario s can be determined using the distribution of P:

$$E\{Loss\}_s = \frac{1}{L}\sum_{l=1}^{L} Loss(p_s^* - P_{l,s}) \tag{10.8}$$

recall that s is a particular scenario, L is the number of geostatistical realizations, and $P_{l,s}$ is the profit for scenario s using realization l. The best estimate of profit for the scenario s is $\widehat{P_s}$ such that the expected loss is minimum when taking $p_s^* = \widehat{P_s}$.

Figure 10.11 shows a small example. The distributions of profits for two scenarios are shown at the top. Both scenarios have the same average profit, but scenario 1 has less uncertainty that scenario 2. Some remarks:

- The least squares loss function, shown in the center, would not distinguish these scenarios because the mean values are the same.

- The loss function to the left shows a greater penalty for overestimation than underestimation. This corresponds to a "conservative" company that does not want to overestimate the value of the reservoir. In this case, the profit value retained would be below the mean, for example, the lower quartile of the distribution is retained if overestimation is three times as consequential. Between two scenarios with the same mean profit, a company using this type of loss function would prefer the one with smaller uncertainty (scenario 1).

- The loss function to the right penalizes underestimation more than overestimation. This corresponds to an "aggressive" company that is seeking profit and would take more risk to achieve that profit. In this case, the profit value retained would be above the mean, for example, the upper quartile of the distribution. Between two scenarios with the same mean profit, a company using this type of loss function would prefer the one with greater uncertainty (scenario 2).

Defining the correct loss function is critical. The optimal decision can be easily determined with a loss function $L(e)$, a distribution of uncertainty ($l = 1, \ldots, L$ realizations), and a set of possible scenarios ($s = 1, \ldots, S$). The challenge is to quantify the most appropriate attitude toward risk, that is, the loss function.

10.7 Work Flow

Figure 10.12 illustrates the work flow for quantification of uncertainty using a combination of scenario-based approach and geostatistical realizations.

Figure 10.13 illustrates the work flow for cross validation of an estimation technique. The cross validation of estimation and simulation techniques is quite different.

Figure 10.14 illustrates the work flow for validation of a Gaussian simulation. Once again, any variable could be considered. The cross validation is done in Gaussian space. The transformation back to real units is straightforward (see Section 2.4). The distributions are represented differently with indicator simulation; the principles of checking a distribution of uncertainty remain the same. Distributions of uncertainty from annealing and object-based modeling require multiple realizations with each well left out, which is a great deal of work.

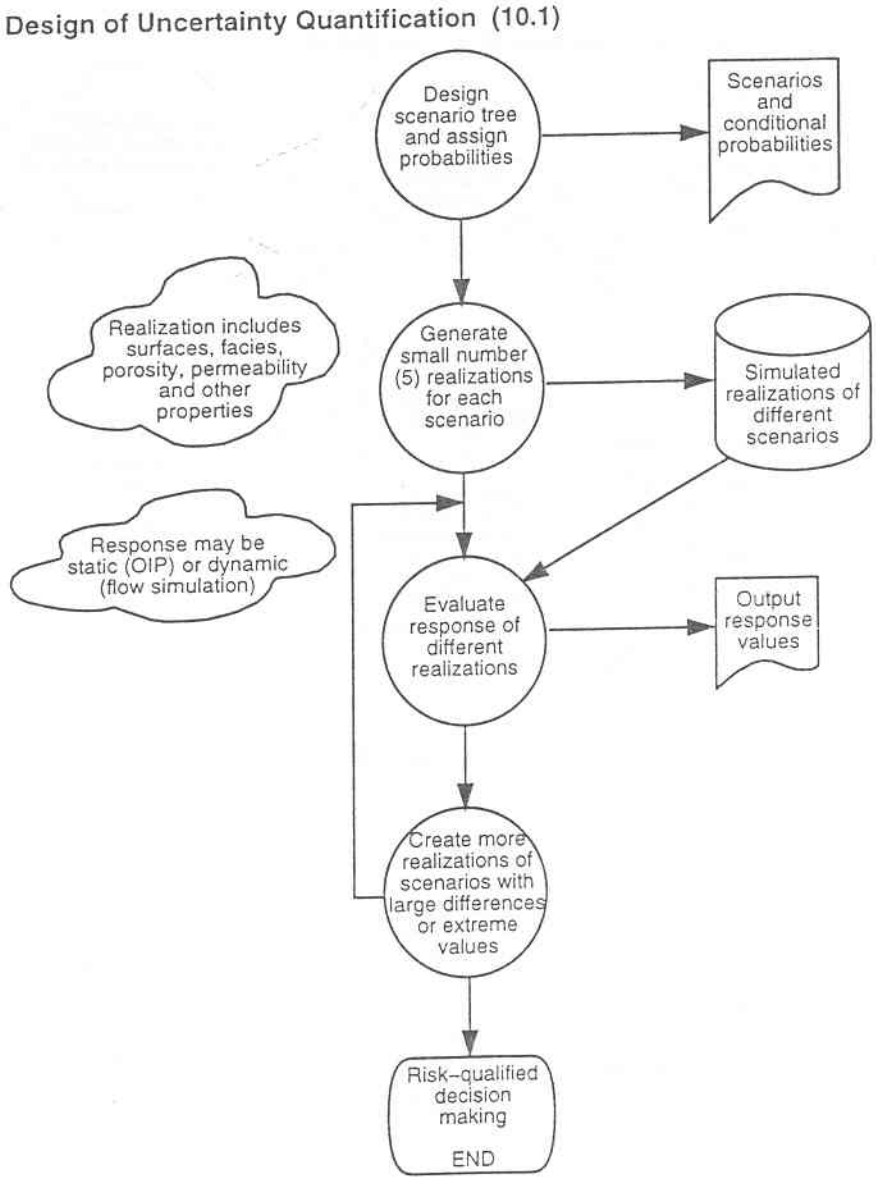


Figure 10.12: Work Flow 10.1: design of uncertainty quantification by scenarios and geostatistical realizations. There are many steps within the "create realizations" bubble.

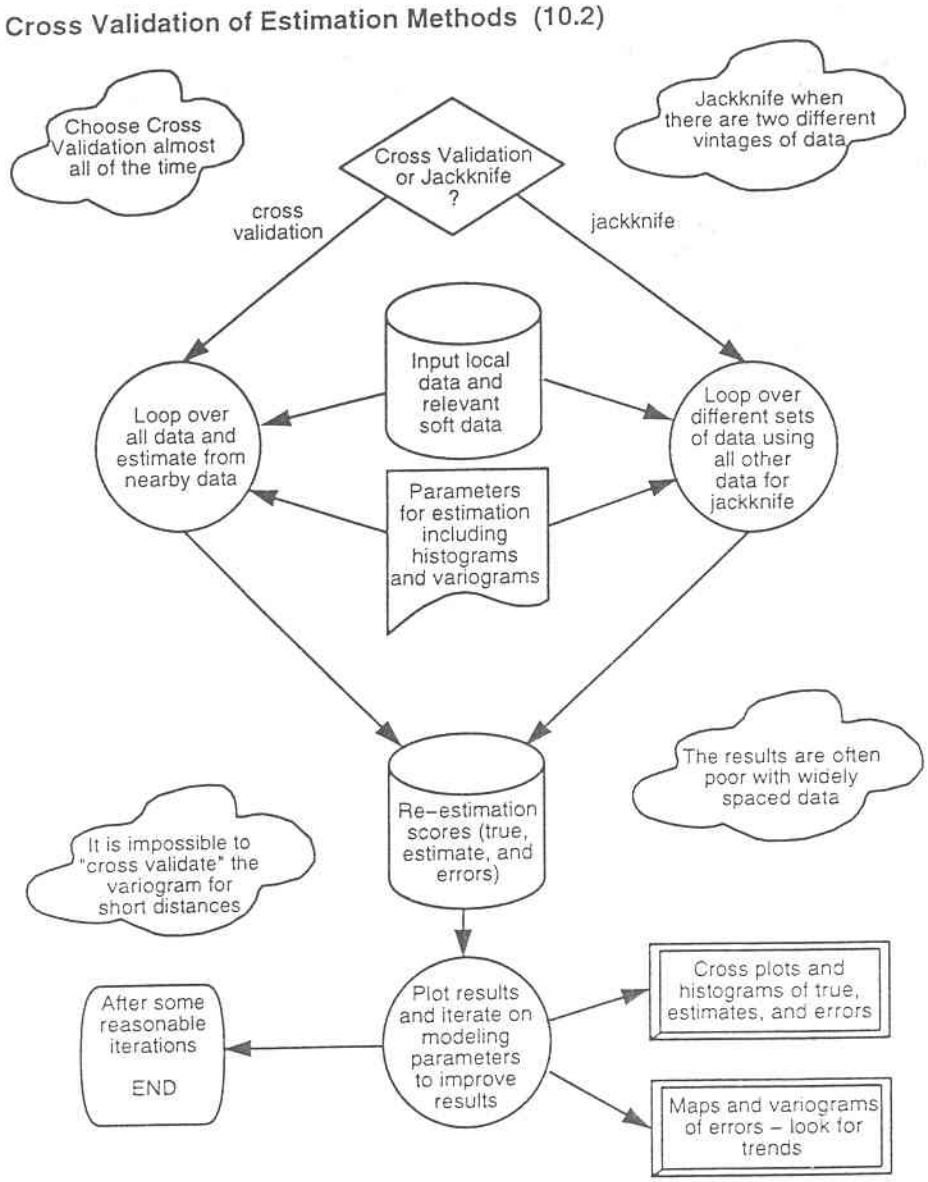


Figure 10.13: Work Flow 10.2: cross validation of an estimation technique.

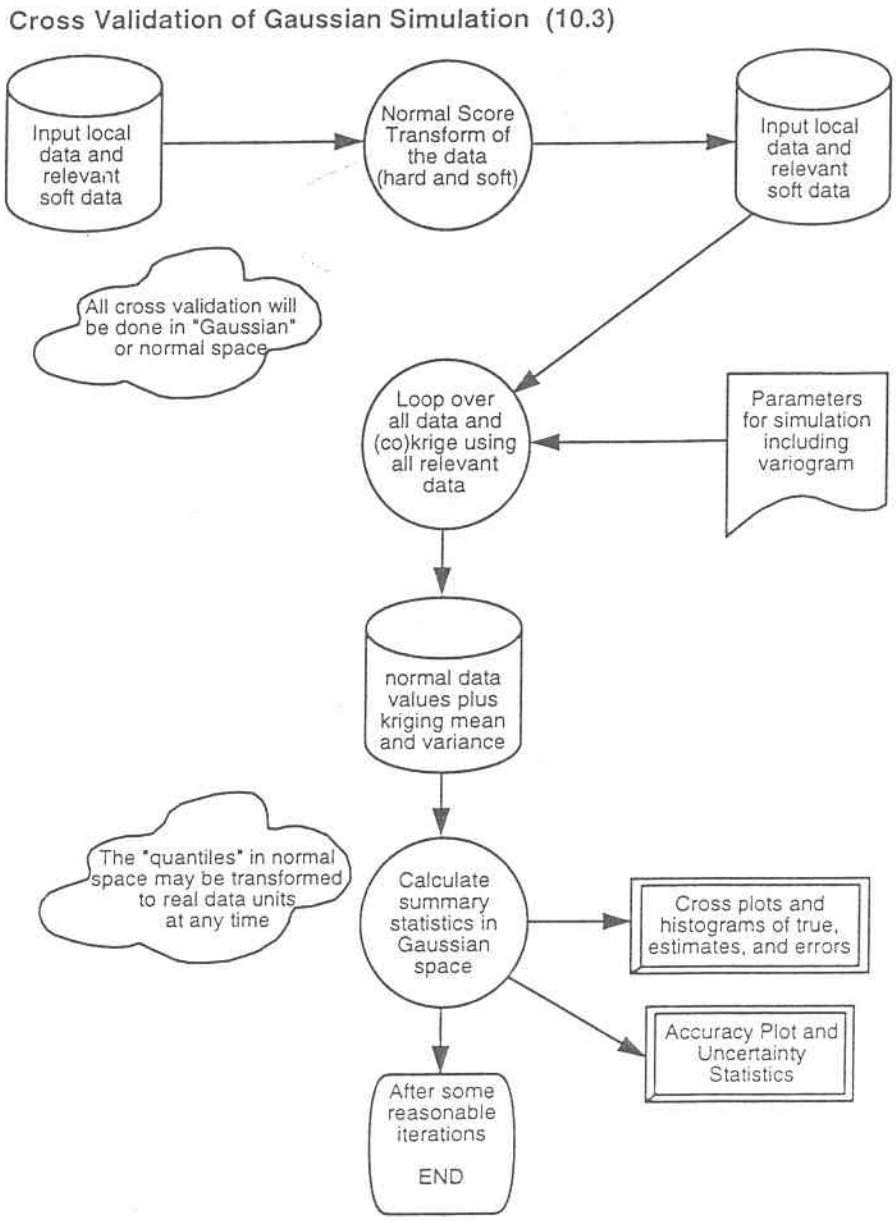


Figure 10.14: Work Flow 10.3: cross validation of Gaussian simulation (with or without a secondary variable).

Chapter 11

Special Topics

Geostatistical reservoir modeling is the subject of ongoing research and development. This chapter presents some frontier areas that are expected to become important in the future.

Section 11.1 presents the problem of the *missing scale* from the core and well log hard data to the scale of the geological modeling cells. This scale difference is typically ignored in conventional geostatistical reservoir modeling; however, the geological variations at this scale could have an important affect on fluid flow.

Geological features were created by the complex succession of numerous deposition and erosion events, which follow physical laws. There are facies relationships and non-linear features that are not captured by conventional geostatistical approaches. Section 11.2 introduces surface-based modeling that has the potential to capture geologically realistic facies and property variations. Section 11.3 presents multiple point statistics that have the potential to go beyond the classical two-point variogram statistics for more geologically realistic reservoir models.

Section 11.4 presents some ideas on the integration of dynamic (time varying) production and 4-D seismic data. Such data are closely related to the reservoir responses being predicted, and hence are very important in the construction of high-resolution geostatistical reservoir models.

Section 11.5 describes some of the issues related to scaling geostatistical reservoir models and transferring them to flow simulation.

Finally, Section 11.6 gives some final thoughts for this introductory book on geostatistical reservoir modeling. Geostatistical methods have many limitations. Nevertheless, they are appropriate for modeling heterogeneity and assessing uncertainty in presence of sparse data.

11.1 Scale Up from Core to Modeling Cell

Reconciling data from different scales is a longstanding problem in reservoir characterization. Data from core plugs, well logs of different types, and seismic data must all be accounted for in the construction of a geostatistical reservoir model. These data are at vastly different scales and it is wrong to ignore the scale difference when constructing a geostatistical model [248].

Geostatistical scaling laws were devised in the 1960s and 1970s primarily in the mining industry where the concern was mineral grades of selective mining unit (SMU) blocks of different sizes. These techniques can be extended to address problems of core, log, and seismic data; however, there are limitations in the context of petroleum reservoir modeling: (1) the ill-defined volume of measurement of well log and seismic measurements, (2) uncertainty in the small-scale variogram structure, and (3) non-linear averaging of many responses including acoustic properties and permeability. The essential results of classical scaling laws are recalled here, but additional research is needed to overcome these limitations.

The first important notion in volume variance relations is the spatial or dispersion variance. The dispersion variance $D^2(a,b)$ [157] is the variance of values of volume a in a larger volume b, that is, in the classical definition of the variance:

$$D^2(a,b) = \frac{1}{n}\sum_{i=1}^{n}\left(\underbrace{z_i}_{\text{Support a}} - \underbrace{m_i}_{\text{Support b}}\right)^2 \tag{11.1}$$

In a geostatistical context, almost all variances are dispersion variances. Often, the volume a is the quasipoint hard data $(\cdot)$ and the volume b is the entire area of interest (A), thus σ^2 could be reported as $D^2(\cdot, A)$.

An important relationship between dispersion variances is the "additivity of variance" relation that is sometimes refered to as Krige's relation:

$$D^2(a,c) = D^2(a,b) + D^2(b,c) \tag{11.2}$$

The additivity of variance notion is explained in detail in any statistics book dealing with analysis of variance (ANOVA). In the context of scaling from core scale to geological modeling scale, consider $a = \cdot =$ the volume of the hard core data, $b = v =$ the size of the geological modeling cells, and $c = A =$ the area of interest (reservoir). Then:

$$D^2(\cdot, A) = D^2(\cdot, v) + D^2(v, A)$$

in words, the variance of core values in the reservoir is the variance of point values in the geological modeling cells plus the variance of the geological modeling cells in the reservoir. It is important to note that this depends on linear averaging, which is correct for porosity and facies proportions, but incorrect for permeability.

Before going further with the development of dispersion variances it is necessary to discuss the meaning and calculation of "gamma bar" values. The "gamma bar" value represents the mean value of $\gamma(\mathbf{h})$ when one extremity of the vector $\mathbf{h}$ describes the domain $v(\mathbf{u})$ and the other extremity independently describes domain $V(\mathbf{u}')$. The size of volumes $v(\mathbf{u})$ and $V(\mathbf{u}')$ need not be the same. The locations $\mathbf{u}$ and loc' need not be collocated. Of course, the volumes and locations could be the same $v(\mathbf{u}) \equiv V(\mathbf{u}')$. In mathematical notation the gamma bar value is expressed as:

$$\overline{\gamma}\left(v(\mathbf{u}), V(\mathbf{u}')\right) = \frac{1}{v \cdot V} \int_{v(\mathbf{u})} \int_{V(\mathbf{u}')} \gamma(y - y') dy dy' \tag{11.3}$$

where v and V are the volumes of $v(\mathbf{u})$ and $V(\mathbf{u}')$, respectively. In three dimensions this integral expression is a sextuple integral, which was infamous among early practitioners of geostatistics.

Although there exist certain analytical solutions to $\overline{\gamma}\left(v(\mathbf{u}), V(\mathbf{u}')\right)$, we systematically estimate $\overline{\gamma}\left(v(\mathbf{u}), V(\mathbf{u}')\right)$ by discretizing the volumes $v(\mathbf{u})$ and $V(\mathbf{u}')$ into a number of points and simply averaging the variogram values:

$$\overline{\gamma}\left(v(\mathbf{u}), V(\mathbf{u}')\right) \approx \frac{1}{n \cdot n'} \sum_{i=1}^{n} \sum_{j=1}^{n'} \gamma(\mathbf{u}_i - \mathbf{u}'_j) \tag{11.4}$$

where the n points $\mathbf{u}_i, i = 1, \ldots n$ discretize the volume $v(\mathbf{u})$ and the n' points $\mathbf{u}_j, j = 1, \ldots n'$ discretize the volume $V(\mathbf{u}')$. The discretization is a regular spacing of points where each point represents the same fractional volume. The number of points n or n' is chosen large enough to provide a stable numerical approximation of $\overline{\gamma}\left(v(\mathbf{u}), V(\mathbf{u}')\right)$. In practice, $n = n' = 100$ to 1000 is more than adequate. 100 points in 1-D, $10 \cdot 10$ in 2-D, and $10 \cdot 10 \cdot 10$ in 3-D is sufficient and provides stable estimates of the gamma bar value.

There are two critical relationships that can be derived from the definition of dispersion variance and gamma bar values:

$$D^2(v, V) = \overline{\gamma}(V, V) - \overline{\gamma}(v, v) \tag{11.5}$$

and the special case where $v = \cdot$ and $V = v$:

$$D^2(\cdot, v) = \overline{\gamma}(v, v) \tag{11.6}$$

These tell us that we can calculate any particular dispersion variance (or volume-related variance) once we have a variogram model for the spatial continuity. The derivation is given in Mining Geostatistics [157].

In presence of variables that average linearly, the change of volume from quasi-point scale $\cdot$ to a geological modeling scale v entails the following:

- mean remains unchanged regardless of scale,
- variance reduces to predictable value $\sigma^2 - \overline{\gamma}(v, v)$, and

- the shape of the histogram of v-size volumes *should* become more symmetric.

The remaining problem is to establish a "change of shape" model. There are a number of analytical models [138] that make specific assumptions about how the distribution changes shape. Experience has shown that these analytical models work when the reduction in variance is small, say, less than 30%. In general, however, a numerical approach must be used to establish the shape of the histogram of geological modeling cell values, that is, we simulate a distribution of quasi-point scale values, average to the desired block scale v, and directly observe the resulting histogram shape.

Relations can also be derived for the change in variogram range and sill parameters, see [63, 157]. As mentioned at the start of this section, there are limitations to these existing procedures that must be addressed by additional research before systematic application in the practice of petroleum reservoir modeling. An important limitation is that these procedures apply to stationary random function models of variables that average linearly and that are characterized by a variogram alone. These are big limitations.

One critical limitation is the assumption that the variogram characterizes spatial variability. Geological features are often much more complex than can be captured by a variogram alone. A striking feature of sedimentary deposits is a geometry defined by stratigraphic bounding surfaces; a promising new area of development is surface-based modeling.

11.2 Surface-Based Modeling

Reservoirs were formed by a succession of depositional and erosional events. Most sedimentary sequences of interest to the petroleum industry were formed along continental margins, where the sedimentation is a result of the interaction of tectonic activity and eustasy [85, 94, 190]. Over long time periods associated with large-scale geological events, major changes in accumulation and erosion result in large changes to the sediment distribution. Within these large-scale changes, there are more frequent events such as the rising and falling of sea level that superimposes cyclical features on the large-scale architecture of the reservoir. Each large-scale sediment package consists of a number of sediment packages at smaller scales, which reflect more frequent depositional events.

Surfaces are boundaries of distinct sediment packages and these boundaries represent the start and end of different geological events. Generally, there are significant changes in the sediment properties at the boundaries due to the change of geological conditions, whereas within the sediment package the petrophysical properties are more homogenous or change with an identifiable trend. Therefore, surfaces provide large-scale connectivity and continuity control of facies and petrophysical properties that is critical for reservoir per-

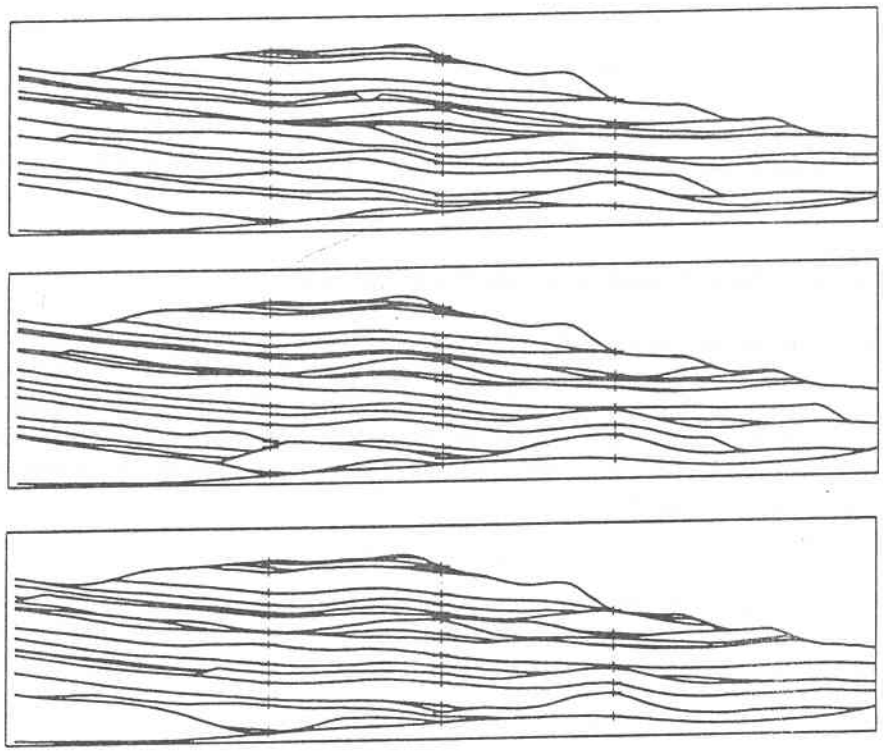


Figure 11.1: Three realizations of stratigraphic surfaces within a larger layer. The realizations are 3-D. These slices are through three wells. Note the reproduction of the well intersections and the different geometry of the surfaces away from the wells.

formance prediction. Considering surfaces in the modeling process provides better constraints for the modeling of facies and petrophysical properties.

Some large-scale time surfaces are visible with seismic data. They are typically modeled deterministically with the aid of well data. These are the stratigraphic layers discussed in Chapter 3. Smaller-scale time surfaces, while not visible with seismic data, can be observed from core or well logs. In general, with few wells, small-scale surfaces may not be modeled deterministically. The purpose of surface modeling is not to get surfaces per se, but to provide constraints for modeling facies and petrophysical properties. Specific facies can be present at surface boundaries. Surfaces may also control grain size and facies trends such as fining or coarsening upward.

Surfaces may be stochastically positioned with parameterized surface templates. Local well data can be explicitly honored at the correct depth. Figure 11.1 gives example realizations that reproduce surface intersections at three well locations.

The place of surface-based modeling in reservoir modeling has not been established. The concepts are important but relatively new; time will tell if the methodology is adopted in practice. Another methodology that has not been established in reservoir modeling is the use of multiple point statistics; however, there is also the promise to create more geologically realistic reservoir models.

11.3 Multiple Point Statistics

Most present applications of the theory of random functions do not consider data-driven multiple point cdfs beyond the two point cdf (4.2). The principal reason is that inference of experimental multiple point cdfs is usually not practical. Thus, random function models have not been developed that explicitly account for multiple point statistics. A particular N-variate or N-point cdf is written:

$$F(\mathbf{u}_1, \ldots, \mathbf{u}_N; z_1, \ldots, z_N) = Prob\{Z(\mathbf{u}_1) \leq z_1, \ldots, Z(\mathbf{u}_N) \leq z_N\}$$

In some applications this multiple point cdf may be summarized by considering binary transforms of $Z(\mathbf{u})$. The multiple point cdf described above appears as the product (non-centered covariance) of N indicator variables:

$$F(\mathbf{u}_1, \ldots, \mathbf{u}_N; z_1, \ldots, z_N) = E\{I(\mathbf{u}_1; z_1) \cdot I(\mathbf{u}_2; z_2) \cdot \ldots \cdot I(\mathbf{u}_N; z_N)\} \quad (11.7)$$

Specific multiple point non-centered indicator covariances have been introduced in the geostatistics literature by Journel and Alabert [152]. In the case of Journel and Alabert [152] the N-points were spatially separated by a fixed lag separation vector $\mathbf{h}$ (that is, $\mathbf{u}_2 = \mathbf{u}_1 + \mathbf{h}, \mathbf{u}_3 = \mathbf{u}_2 + \mathbf{h}, \ldots$). A critical cutoff z_c is considered and the N-point connectivity function $\phi(\mathbf{u}; N, z_c)$ is defined as the expected value of the product of N indicator variables:

$$\phi(\mathbf{u}; N, z_c) = E\left\{\prod_{j=1}^{N} I(\mathbf{u} + (j-1)\mathbf{h}; z_c)\right\} \quad (11.8)$$

this could be interpreted as the probability of having N points, starting from $\mathbf{u}_1$, aligned in the direction of $\mathbf{h}$ being jointly below cutoff z_c.

In the case of a categorical variable, one could also maintain all of the information provided by the probability density function (pdf) representation (an extension of the two-point case presented earlier (4.6)). For example,

$$f(\mathbf{u}_1, \ldots, \mathbf{u}_N, ; k_1, \ldots, k_N) = Prob\{ \begin{array}{ll} Z(\mathbf{u}_1) & \in \text{ category } 1, \ldots, \\ Z(\mathbf{u}_N) & \in \text{ category } N \ \} \\ & k_1, \ldots, k_n = 1, \ldots, K \end{array} \quad (11.9)$$

is the multivariate or multiple point distribution of $Z(\mathbf{u}_1), \ldots, Z(\mathbf{u}_N)$. This type of multiple point distribution, when established from experimental proportions, is referred to as a *multiple point histogram* or an *N-point histogram*.

Given an N-point histogram of a categorical variable that can take one of K outcomes, there are K^N classes. Commonly, the K^N classes are assigned a one dimensional *index* for computer storage and to allow N-point histograms to be displayed in the traditional way, that is, as bar charts of frequency versus class number. Given N categorical variables $z(\mathbf{h}_i), i = 1, \ldots, N$ that can take values $k = 1, \ldots, K$ a unique one-dimensional index may be computed as:

$$index = 1 + \sum_{i=1}^{i=N} [z(\mathbf{h}_i) - 1] \cdot K^{i-1} \tag{11.10}$$

Each index value corresponds to a unique *pattern*, that is, a particular combination of the N points.

Multiple point statistics such as non-centered covariances and multiple point histograms hold out the promise of more realistically representing geological features; however, there are a number of considerations that have slowed practical application. First, reliable inference of the multiple point statistics from sparse well data is problematic. Acquiring a "training image" of the right scale and representative of the 3-D complexity of the reservoir is problematic. Second, enforcing reproduction of the multiple point statistics at the right scale simultaneously with other data can be difficult. These challenges are likely to be met by active research in this area [237].

11.4 Dynamic Data

A number of workers have considered the future of research and applications in reservoir modeling [88, 131, 116, 119]. Important areas for research and development focus on data integration and the management of uncertainty. A critical aspect of data integration is consideration of well test, historical production, and 4-D seismic data. These data are time varying responses that are non-linearly related to the static facies, porosity, and permeability. Consideration of dynamic data integration in reservoir characterization is an active area of research and worthy of treatment in a separate book. Following is a brief presentation of where research and development are going in this important subject area, particularly for production data.

Early approaches to numerical history matching by mathematical inversion have been hampered by: (1) excessive CPU execution time for even relatively small reservoir models, (2) limiting, simplifying assumptions, for example, linear relationship between pressure and reservoir properties, (3) consideration of one data type at the expense of information coming from other data sources such as seismic and geology, and (4) significant difficulty with multiphase flow data. Notwithstanding some research advances in this area, the most common history matching technique procedure is manual iteration.

Production data measurements are essentially at a single point in space that can only yield effective properties over a large volume. Attempts to

build a detailed subsurface model from production data alone are not likely to be successful. A promising avenue is to use the production data at wells to establish coarse-scale maps of permeability. These maps are then used as constraints along with those from other available data in geostatistical reservoir modeling.

Production data could be used to improve the estimation of global reservoir parameters, such as the mean and variogram of reservoir properties. Once estimated, such statistical parameters are used as inputs in geostatistical techniques to construct reservoir models. The contribution of production data lies in the improvement in the estimation of statistical parameters describing the reservoir heterogeneity.

In general, from a provenance perspective, production data may be summarized into three main groups: single-well test data, multiple-well test data, and multiple-well multiple-phase historical production data.

Mathematical inversion methods for single well test pressure data and interpretive tools are largely in place [80, 133, 199, 200, 207, 217]. This is perhaps the subject of most well test analyses research. Some of the typical single well test data are RFT data, drawdown/buildup test data, variable rate test data, production logs, and permanent pressure gauges. The main result of these mathematical inversion methods are large-scale effective porosity and permeability values. To a large extent, these large-scale properties can be accounted for in geostatistical modeling [4, 56, 61].

Compared to single-well test data, multiple-well test data are more extensive in terms of areal coverage and provide specific connectivity information between wells. An important issue is whether sufficient data are available to establish unambiguous connections between multiple locations.

According to production mechanism, historical production data arise from different sources. The classification can be with respect to reservoir depletion with or without water drive, with gas-cap drive, water injection, or gas injection. Each of these has unique implementation and interpretation issues.

A thorough review of the subject of parameter identification in reservoir simulations is not possible. Some reasonable literature reviews include Jacquard and Jain [139], Gavalas et al. [96], Watson et al. [261], Feitosa et al. [89, 90], and Oliver [201]; Yeh [278], and Carrera and Neuman [35] have prepared similar reviews in groundwater hydrology.

11.5 Input to Flow Simulation

The transfer or reformatting of geostatistical reservoir models to flow simulation models must be approached with caution. There are reasonable techniques to calculate effective properties of large flow simulation blocks; however, there is inevitable loss of information and numerical discretization errors. Many people have studied the scale up of high-resolution geostatistical models [52, 53, 104, 181, 246, 267]. A gridding study is always recommended

to assess the consequences of grid discretization in presence of realistic reservoir heterogeneity.

Facies must be scaled up because relative permeability and other saturation functions are often assigned on a by-facies basis in the simulator. The most common facies within each large-scale flow simulation block is taken as the large-scale facies. This may lead to under representation of facies that have small global proportions. Alternative schemes could be considered.

Porosity and saturation are both volumetric concentrations and average linearly. Arithmetic averaging is theoretically correct. The only complication is in low net-to-gross reservoirs where some of the porosity at a small scale may be "ineffective." This complication is not currently considered in scale up.

Permeability

The scaleup of permeability is more problematic. The effective permeability of a flow simulation grid block depends on the spatial arrangement of the constituent geological modeling cells and the flow boundary conditions. The scaling of relative permeability curves and other saturation functions is even more complex than absolute permeability. A very brief discussion on the scale up of absolute permeability is given here. The reader is referred to the literature (primarily SPE papers) for more details.

The three common simple averaging methods are (1) arithmetic averaging, which is correct for linear flow in parallel composites, (2) harmonic averaging, which is correct for linear flow in series composites, and (3) geometric averaging, which is correct for white-noise random media in 2-D with a lognormal distribution of permeability [185]:

$$k_a = \frac{1}{n}\sum k_i, \quad k_h = \left[\frac{1}{n}\sum \frac{1}{k_i}\right]^{-1}, \quad k_g = (\prod k_i)^{\frac{1}{n}}$$

The geometric average can be equivalently calculated as the exponentiation of the arithmetic average of the logarithms of the data. The arithmetic average is the upper bound and the harmonic average is the lower bound of the effective permeability. These may not be the bounds if extreme permeability features such as "zero" permeability shales and "infinite" permeability fractures have not been representatively sampled.

Directional averaging [33] gives much more realistic bounds on the effective permeability. Consider a flow simulation grid block made up of a regular 3-D network of geological modeling cells. The upper bound of the effective permeability in a particular direction can be calculated by a two-step procedure: (1) calculate the arithmetic averages the permeability in 2-D slices perpendicular to the flow direction, and (2) calculate the harmonic average of the 2-D slice averages. The lower bound of the effective permeability is given by a similar two-step procedure: (1) calculate the harmonic averages in each 1-D column of cells in the the flow direction, and (2) calculate the arithmetic

average of the 1-D column averages. Averaging is problematic when these two bounds are very different. A map of the difference can be used to identify areas where the grid size or orientation could be changed.

The conventional averages can all be written in a general "power" averaging [54, 155, 165] formalism:

$$k_\omega = \left[\frac{1}{n}\sum k_i^\omega\right]^{\frac{1}{\omega}}$$

where the arithmetic average is at $\omega = 1$, the harmonic average at $\omega = -1$, and (at the limit)the geometric average is at $\omega = 0$. Studies [54] show the averaging power depends on direction, where it is greater in the horizontal direction and less in the vertical direction. The averaging power depends on the spatial continuity and not on the histogram or relative amounts of high and low permeability values. The averaging power is close to 0.6 for horizontal permeability and -0.4 for vertical permeability for a wide range of geostatistical models. Of course, it should be calibrated to the particular heterogeneity present in each reservoir. The calibration procedure requires the "true" effective permeability to be calculated by single-phase steady-state flow simulation on the flow simulation grid blocks.

These small scale flow simulations are remarkably fast. This leads us to the most common approach to scale up of permeability; just solve for it directly. Boundary conditions must be chosen. No flow boundary conditions on the four faces perpendicular to flow are easy to apply, but can unrealistically confine flow and lead to low permeability. Periodic boundary conditions are often more realistic.

11.6 Final Thoughts

High resolution 3-D models of reservoir properties are required for resource evaluation, flow studies, and reservoir management. This book has presented geostatistical techniques to build such 3-D models. These geostatistical techniques have limitations. A list of the most important limitations would include: (1) relative to conventional methods such as contouring, splines, or distance weighting, geostatistical methods take more time to apply, require more expertise and training, and have the potential for a greater number of mistakes; (2) there is a need for complex and complete input statistics and yet there is often no reliable source for most of these input parameters; and (3) there is an implicit assumption that geostatistical models reflect facies and property variations at all smaller scales, for example, from the core scale through the geological modeling scale. These limitations are being addressed through professional training, improved software, and research.

Even with these significant limitations, geostatistical techniques provide the best tools available for constructing reliable numerical reservoir models

because they (1) provide means to introduce a controllable degree of heterogeneity in facies and reservoir properties, (2) permit geological trends and data of different types to be accounted for by construction rather than laborious "hand" processing, (3) make it easier to perform sensitivity studies, and (4) provide a means to assess uncertainty.

The objective of this book was to collect the important aspects of geostatistical theory and practice applied to petroleum reservoir modeling. That objective has been partially met. There is clearly a need for more specialized books and monographs.

Appendix A

Glossary and Notation

A.1 Glossary

The more complete glossary editted by Ricardo Olea [197] is a useful supplement to the terse presentation below.

absolute permeability: the permeability (see below) of a rock saturated completely with one fluid.

accommodation: a term used in sequence stratigraphy that refers to the volume available for the sediment to accumulate. This term is not used extensively in this book.

aeolian: or eolian sediments are those transported and deposited by wind energy, for example, sand dunes.

analogue data: data from outcrops or more densely drilled reservoirs of similar depositional environment. Analogue data are used to determine statistics (horizontal variograms in particular) for the field of study.

annealing: see simulated annealing.

anisotropy: directional dependence of a parameter, for example, the variogram range. In the context of the variogram, geometric anisotropy is anisotropy that may be corrected by an affine correction of the coordinate system. Zonal anisotropy may not be corrected by an affine transformation and indicates additional variability specific to a particular direction.

affine variance correction: an approach to modify Z values to have lesser variability, that is, to account for a larger-volume support. The affine correction leaves the shape of the histogram unchanged and simply shifts data values closer to the mean, depending on the magnitude of variance reduction, for example, $z' = \sqrt{f} \cdot (z - m) + m$.

bivariate: related to the joint probability distribution of two variables, for example, porosity and permeability.

bootstrap: a statistical resampling procedure whereby uncertainty in a calculated statistic is derived from the data itself. Monte Carlo simulation is used (with replacement) to resample from the data distribution. Strong assumptions of data independence and representativity are required.

bombing model: a particular indicator variogram model based on the result of randomly dropping (or bombing) ellipsoids or any other fixed shapes, coded as "0s" into a background matrix, coded as "1s."

boundary conditions: in the context of determining the effective properties of a grid block, boundary conditions are the pressure and flow rate conditions surrounding the block of interest.

bounding surface: a stratigraphic surface that separates two reservoir units. Typically, a bounding surface would correspond to some fixed geological time.

CV: acronym for coefficient of variation, which is the mean of a distribution divided by the standard deviation. This measure of variation is sometimes preferred since it is dimensionless.

calibration: in the context of geostatistical modeling, calibration is the exercise where the relationship of soft secondary data to the "hard" data being modeled is established. The calibration relationship is almost always probabilistic in the sense that there is no one-to-one relationship.

carbonate: a geological reservoir type consisting of calcareous deposits deposited as the shells of reef-dwelling organisms.

cell-based: a technique to build a lithofacies model where facies codes are assigned to cells such that the statistical relations between the cells reproduce some realistic variations. Compare with "object-based" and "surface-based" techniques.

cell declustering: a method for declustering that weights data such that each sub area (or cell) receives the same weight, see index.

cleaning: a (geo)statistical procedure whereby a categorical variable realization is modified to remove short scale variations that are deemed unreliable.

clinoform: a geological structure formed during deposition. A stratigraphic layer may be made up of clinoforms blurring the identification of the principal directions of continuity.

cloud transform: a probability field (see p-field) technique of simultaneously drawing from conditional distributions. The term "cloud transform" may be used interchangeably with "p-field;" however, there is the connotation with cloud transform that the conditional distributions are coming from a bivariate cross plot.

conditional: made to reproduce certain local data at their locations. Used in the context of geostatistical simulation.

connectivity: the "persistence" or linkage between reservoir units that are good quality and poor quality. Connectivity is a straightforward concept, but more difficult to quantify numerically. The variogram is one

measure of connectivity, but it does not capture multipoint connectivity that would be measured by a flow simulator.

cokriging: Kriging applied with more than one data type. The same basic set of equations are called for; however, the covariance between the different data types is also required.

collocated: at the same location.

coregionalization: the noun used to refer to a random function of two or more variables. A "regionalization" is the probability law of one particular variable.

cosimulation: simulation of one variable using data of multiple types or, more commonly, realizations from a previous variable for conditioning.

covariance: measure of correlation in units of the variance. The covariance is the expected value of the product of the two variables Y and Z, that is, $E\{[Y - \overline{Y}][Z - \overline{Z}]$. Often, the covariance is standardized by the standard deviations of Z and Y, it is then called the correlation coefficient.

cross validation: the procedure of leaving each data (or the entire well) out of the dataset and estimating the reservoir property at that well location. This is repeated for all data and the true values and their estimates can be compared to judge the goodness of the estimation technique.

cross variogram: the variogram (see below) between two different variables. The variogram is defined as $2\gamma(\mathbf{h}) = E\left\{[Z(\mathbf{u}) - Z(\mathbf{u}+\mathbf{h})]^2\right\}$ and the cross variogram is defined as

$$2\gamma(\mathbf{h}) = E\left\{[Z(\mathbf{u}) - Z(\mathbf{u}+\mathbf{h})]\,[Y(\mathbf{u}) - Y(\mathbf{u}+\mathbf{h})]\right\}$$

where Z and Y are two different variables.

cyclicity: a particular variogram signature called the "hole effect" that is induced by geological periodicty that causes repeated patterns.

declustering: assignment of weights whereby redundant data are given less weight in the construction of a representative probability distribution or the calculation of summary statistics.

depositional setting: the geological environment (fluvial river channels, deltaic deposits, deepwater turbidite deposits, carbonate reefs, and so on) in which the sediments were formed. The geostatistical methods we employ are different within different depositional settings.

despiking: a technique to break ties in a data set, for example, certain techniques would require *spikes* of constant values such as 10% zero's to be ranked from smallest to largest for the purposes of transformation and back transformation. The ties would be broken by local averages.

diagenetic: an alteration to the character of reservoir sediments after deposition. Diagenetic alteration could take the form of silica or carbonate cements or remineralization.

direct: adjective used before geostatistical simulation to mean that the values have not been transformed by Gaussian, indicator or any other transformation method. Direct simulation is suited to multiscale data where the variable does not average linearly after transformation.

dispersion dariance: the expected value of the spatial variance of values of support a in larger volume support b, commonly denoted $\sigma^2(a,b)$.

distribution: a word used loosely for many different purposes. A probability distribution is the histogram of cumulative probability values of a random variable. A spatial distribution is the set of values in space that describe a reservoir.

dynamic data: data that is inherently time-varying such as well test or historical production data. The challenge is to relate the time variations to spatial variations.

equally likely to be drawn: the expression applied to simulated realizations that are drawn by (pseudo) random numbers. The result of each realization will enter a histogram of uncertainty with an equal increment.

equiprobable: a term applied to simulated realizations that are considered equally probable given the modeled probability distribution. This is distinct from "equally likely to be drawn." Geostatistical realizations should rightly be referred to as "equally likely to be drawn"

ergodic: A study area is considered "ergodic" when it is large with respect to the range of correlation.

erosion: removal of sediment, which affects the geological correlation style for geostatistical modeling.

exponential Variogram: a common variogram model.

eustasy: large-scale change in structure related to advancing and receding ice during and after ice ages.

facies: categorical classification of reservoir rock into different types based on distinct petrophysical properties. Facies are based on lithology and other considerations. There are many different basis for facies specifications.

faults: structural deformation of the reservoir layers that could lead to a displacement and disconnection between the layers. Faults could also be conduits for fluid flow.

flow unit: a particular portion of the reservoir that is believed (through empirical observation or flow simulation) to be in communication. Historical practice of mapping flow units forces a- priori expectations of reservoir continuity into the model, which is good if that knowledge is based on actual measurements and dangerous if not.

fluvial: river or stream-derived sediments. Fluvial reservoirs may be modeled by object-based procedures where sand-filled channels and other facies are modeled as "clean" objects.

forward modeling: the "forward" problem is the processing of static geostatistical models to calculate flow responses. The "inverse" problem is the determination of the static geostatistical models that lead to known flow responses.

gamma bar: more correctly denoted as $\overline{\gamma}(V, v)$ is the average variogram between volumes V and v, that is, the average variogram value where one end of the lag vector independently defines V and the other independently defines v. V and v could be the same.

Gaussian variable: a variable that follows a Gaussian or normal distribution. Often, data are transformed to follow a Gaussian distribution for ease of calculation.

Gaussian variogram: a common variogram model. This variogram model implies smooth variability and is normally observed with thicknesses or topographic elevations.

GeoEAS: a geostatistical package put together by the EPA for environmental site characterization. The ASCII file format adopted by the GeoEAS software is widely used by other geostatistical software.

geometric mean: exponentiation of the average of logarithms of a data set. The geometric mean can be written in a number of forms, see index.

geostatistics: study of phenomena that vary in space and/or time.

grain size: the size of the rock particles that make up a sedimentary rock. In general, coarser grain size leads to greater permeability. Poorly sorted grains, that is, of many different sizes, often lead to reduced porosity and permeability.

hard data: direct measurements on the petrophysical property of interest. Hard data are considered the reference data for all other potential sources of information. See also *soft data*.

heterogeneities: variations in reservoir properties. These variations often appear as random or unpredictable because of the relatively large data spacing.

horizon markers: the depth or elevation locations of important geological bounding surfaces.

hydrocarbon volume: the volume, expressed in cubic meters or barrels, of reservoir that is hydrocarbon.

indicator: a data transformation method whereby a continuous or categorical data variable is transformed to a probability value. A continuous variable is transformed to the *probability to be less than a certain threshold* and a categorical variable is transformed to the *probability of that category prevailing*. The indicator transform is always between zero and one.

inverse modeling: see *forward modeling*

isochore: thickness of a stratigraphic formation measured in the vertical direction.

isopach: thickness of a stratigraphic formation measured perpendicular to the dip of the formation.

isotropic: showing no directional preference or *anisotropy*.

jackknife: a portion of the data is withheld from the geostatistical analysis until the very end, then, the reservoir properties are estimated at these data locations. The closeness of the withheld true values to the estimates is a measure of the goodness of the estimation procedure.

KT: kriging with a trend model (see also universal kriging).

Kriging: spatial linear regression named in honor of Prof. Krige in South Africa. The kriging equations are derived for the best linear estimator given a stationary covariance model and a particular model for the trend. The "normal equations," developed at about the same time, are the same as kriging.

lag deposits: sediment deposited at the base of a depostional unit, for example, shale clasts, may be found at the base of certain fluvial deposits.

layercake: geological model made up of deterministic layers of constant properties. Although these models are not necessarily realistic, they provide a measure of stratification that typically exists in a reservoir.

lithofacies: reservoir rock type.

lognormal: refers to a distribution where the logarithm of the variable follows a normal or Gaussian probability distribution. This adjective is often wrongly used to refer to any histogram that is lower- bounded by zero and positively skewed.

low pass filter: a term from geophysical data processing, which amounts to filter out high-frequency (short-scale or periodic) variations. Kriging is sometimes called a low-pass filter because of its smoothing.

Markov: a model whereby a limited number of nearest neighbors screen the information of all further away data.

model: a much-abused word in science and engineering typically refering to a numerical replacement for some physical or mathematical quantity. The context is typically required to sort out the exact meaning. A variogram *model* is an equation that gives the variogram for all distances and directions. A reservoir *model* is a gridded set of facies, porosity, and permeability values that represent the entire volume of interest.

modeling cell: a regular volume in transformed stratigraphic coordinate space that is assigned facies, porosity, and permeability by geostatistical means. Modeling cells are typically 1-3 feet thick and 50-500 feet in areal dimension.

Monotonic: consistently increasing or decreasing, for example, a cdf is a non-decreasing or monotonic function.

Monte Carlo: the procedure of drawing or sampling from a specified probability distribution. Monte Carlo simulation typically proceeds in two steps: (i) generating a pseudo-random number that follows a uniform distribution between zero and one, and (ii) reading the quantile, associated to that random number, from the target distribution.

missing scale: the volume scale difference between the data and the grid cells being modeled. This is termed the "missing" scale when no explicit attempt is made to capture the change in statistical properties between the data scale and model scale.

multiphase: normally used to refer to the flow regime in the reservoir, for example, multiphase flow is the flow of multiple phases such as oil, water, or gas.

multple point statistics: statistics related to multiple points taken simultaneously. Variograms are two-point statistics. Three, four, and more points could be simultaneously considered for multiple- point or multi-point statistics.

nested: structures refer to multiple variogram structures that are combined to model the spatial variation in complex settings.

net: a "net" rock type is one that has hydrocarbons that could be produced unless it is bypassed. For example, high-porosity sandstone is a net rock type and shale would be a *non-net* rock type.

net-to-gross: the fraction of net rock type in a reservoir.

normal variable: the word "normal" is interchangeable with "Gaussian." A normally distributed variable is one that has a Gaussian probability distribution. The use of one word or the other is largely dictated by convention, for example, we often write *normal score transform* and *sequential Gaussian simulation*, but rarely do we write *Gaussian score transform* or *sequential normal simulation*.

object-based: a technique to build a lithofacies model where whole geological objects such as fluvial channels, remnant shales, or sand dunes are embedded within a matrix facies. Although the final model may be cellular, it is still called "object-based" because of how it was constructed in the first place.

onlap: a geological correlation style where the "horizontal" lines of continuity truncate against a lower surface that may have represented the present-day topography at time of deposition.

order relations: the required properties of a probability distribution, that is, no negative probabilities and the sum of all probabilities must equal one. The context of this expression is typically indicator kriging that does not enforce order relations be construction, they must be enforced by a posteriori correction.

p-field: the technique of separating the construction of a conditional probability distribution and the Monte Carlo simulation from the distribution. Correlated probability values, or a *p-field*, is used to perform the Monte Carlo simulation.

paleoslope: the slope (and direction of slope) that existed at the time of deposition. Knowledge of this direction may help in determining directions of continuity for variogram calculation.

parametric: refers to a probability distribution that is characterized by a limited set of numbers or parameters such as the mean and variance. The most common parametric probability distribution would be the normal or Gaussian distribution.

pdf: acronym for probability density function.

periodicity: see *cyclicity*.

permeability: defined as a constant relating fluid flux to pressure gradient in a porous substance.

petroleum: hydrocarbons of (often) commercial value to be extracted from a reservoir.

petrophysical property: some attribute of a given volume of rock, for example, porosity or density.

polygonal declustering: a method for declustering that uses polygonal areas of influence, see index.

pool: a particular reservoir, for example, there are twenty pools in this basin. "Pool" is also used as a verb to mean combining data together, for example, data from all sand types may be pooled together for geostatistical calculations.

population: a statistical or geological group of data that relates to a particular reservoir zone or facies. The interbedded sand facies within a particular reservoir zone would be ab example.

porosity: void space in rock. Effective porosity is the void space that could contribute to fluid flow. Total porosity is the total void space including those pores that could never be connected and drained. We are typically interested in effective porosity.

positive definite: the mathematical property of a covariance model that ensures that the variance of all linear combinations is strictly greater than or equal to zero. The same property of the corresponding variogram is refered to as "non-negative."

power law variogram: a particular type of variogram model without sill that may be interpreted as fractal.

power law scaleup: expressing the average or effective property of a region as a ω or power law average where $\omega = 1$ for arithmetic and $\omega = -1$ for harmonic averaging. The ω is calibrated from flow simulation.

primary variable: the variable currently being modeled. All other data are considered secondary.

proportional correlation: a particular type of stratigraphic correlation style where the "horizontal" continuity is parallel between two grid surfaces. There is no top erosion or bottom onlap.

Q-Q plot: a Q-Q or quantile-quantile plot compares two different histograms or univariate distributions: a scatterplot of matching (according to cumulative probability) quantiles from two distributions.

quantile: any Z-variable value that separates a histogram into two parts. A Z-quantile value is qualified by the cumulative probability associated to it, for example, the 0.5 quantile is the median.

quartile: the first, second, and third quartiles are the 0.25, 0.5, and 0.75 quantiles, respectively. They divide a distribution into four equal parts.

random function: a statistical model of a variable distributed in space (say, porosity) characterized by a multivariate probability distribution. Continuous variable random functions (RF's) are often assumed to me multivariate Gaussian because geostatisticians can handle them easily.

ranking: order a set of realizations from low to high based on some criterion, for example, net-to-gross or average porosity. Selected realizations can then be carried through flow simulation.

realization: specific output of lithofacies, porosity, and permeability drawn by Monte Carlo simulation.

relative permeability: saturation-dependent factor that specifies what fraction of the total or absolute permeability that is available for a particular fluid. There is less relative permeability for fluids with low saturation since they are likely disconnected.

regionalized variable: a variable (such as lithofacies, porosity, or permeability) that may take different values at different spatial locations.

regridding: the process of representing a reservoir at a different grid system, for example, a different stratigraphic and flow simulation grid system are often used and there is a need to go between the two grid systems.

reserves: amount of a resource that can be economically extracted.

reservoir: some interesting finite volume of rock and fluid in the subsurface. Enhancing the economic extraction of important fluids from a reservoir is often the ultimate goal of reservoir characterization.

resource: amount of hydrocarbon that could potentially be extracted.

saturation function: a variable such as relative permeability or capillary pressure that is a function of fluid saturation.

sequential: in geostatistics, the sequential adjective typically refers to the order of simulation, that is, one grid node after another using previously simulated grid nodes as conditioning for the current and future grid nodes.

secondary variable: variable related to the current primary variable being modeled.

siliciclastic: silica-type sand/shale transported by either water or air, deposited, and lithified to sandstone/shale.

simulation: drawing realizations from a probability distribution that may be analytically defined or only observable after the drawing of many realizations. The realizations typically involve multiple variables.

simulated annealing: an optimization technique based on an analogy with the physical process of annealing. Simulated annealing does not require calculating derivatives, but may be difficult to code efficiently.

smoothing: this is used in a number of contexts. Histogram or distribution smoothing refers to a procedure to fit or "fill-in" a sparse distribution with one that removes spikes and adds values that should be there, but are not sampled. Kriging is also said to smooth the data because the kriged values have less variability than the data from which they were calculated.

soft data: indirect measurements of the petrophysical property of interest. These data must be calibrated to available *hard* data, or direct measurements, before being used in geostatistical modeling. See also *soft data*.

space of uncertainty: the range of possible outcomes from a particular probabilistic model. The space of uncertainty is typically sampled by multiple realizations and is inaccessible in practice.

spherical variogram: a common variogram model that takes it name from the geometric interpretation of its mathematical expression; the standardized spherical variogram function for lag distance **h** is 1.0 minus the intersection volume of two unit spheres separated by lag distance **h**.

spread: the deviation of data from the center of the data values. The variance and interquartile range are two common measures of spread.

static: an intrinsic property of the porous media such as porosity and permeability that does not change over the practical time scale of the reservoir. Compare with a *dynamic* measurement that is time dependent.

stationarity: decision that statistics do not depend on location within the study area, for example, the variogram is often assumed "stationary" in the sense that it is applicable over the entire area over which it has been calculated.

stochastic: a numerical procedure is called "stochastic" when it involves the use of a random variable

stratigraphy: geological study of the large scale geometry and continuity of reservoir facies and properties. One typically talks about stratigraphic layers (large scale zones of reservoir rock that belong together), stratigraphic coordinates (conversion of the subsea depth or elevation coordinates to one convenient to model), or sequence stratigraphy (relates to correlating layers according to the chronography or time of deposition).

structural analysis: typically the analysis and modeling of geometric architecture of the reservoir, that is, faults, folds and deformation. Sometimes used in place of *variography*.

support: the measurement volume of a particular type of data. For example, the support of a core plug is a few cubic centimeters, which is less than a well log measurement, which is less than a well test measurement.

surface-based modeling: the procedure of constructing a reservoir model where the reservoir layers are built up by a stochastic succession of surfaces that define depositional events at a scale smaller than the well spacing.

systems tract: a particular geological setting for deposition leading to particular lithofacies and trends in porosity and permeability. Some systems tracts include fluvial, deltaic, shoreface systems, deepwater depositional systems, platform carbonates, and so on.

trend: vertical or areal variability that is predictable. For example, there is often a fining-upwards trend in alluvial sand bodies since coarse grains settle quicker than fine grains. The variability of most natural phenomena is both deterministic or predictable (the trend) and stochastic. The stochastic part is modeled by geostatistics.

transformation: there are different types of transformation: *data* transformation that converts the data to a different histogram, and *coordinate* transformation used to convert real reservoir depth or elevation coordinates to a system that is more convenient for modeling.

truncated Gaussian: a technique to model a categorical lithofacies variable. A continuous "Gaussian" variable is simulated first and then divided (or truncated) into classes.

uncertainty: lack of complete information about some quantity. Uncertainty is often represented by a probability distribution, whose variability or "spread" tells us the level of uncertainty.

universal kriging: kriging where the mean is specified to follow a particular form of trend. The coefficients of the trend are estimated implicitly by the kriging procedure. More commonly referred to as "kriging with a trend."

variogram: measure of spatial variability for a separation lag distance $\mathbf{h}$. Defined as the average squared difference in the variable value for that distance, that is, $2\gamma(\mathbf{h}) = E\left\{[Z(\mathbf{u}) - Z(\mathbf{u}+\mathbf{h})]^2\right\}$. The variogram is $2\gamma(\mathbf{h})$ and the *semi*variogram is $\gamma(\mathbf{h})$.

variography: the iterative procedure of calculating, interpreting, and modeling variograms.

volumetrics: a reporting of the volume of hydrocarbon in a particular reservoir interval (over some limited areal extent). Volumetics typically refer to a static resource before consideration of a particular recovery scheme or development plan.

Walther's Law: a geological *law* that (very roughly) states that facies associations encountered in the horizontal direction are also encountered in the vertical direction. This is often brought up as an arguement to use the same type of horizontal variogram as vertical variogram, just with a different range of correlation.

water cut: fraction of water to total flow rate. This is an important response variable in reservoir management.

zonal anisotropy: directional variability induced by stratigraphic or areal zonation. The variogram in one direction would not encounter the full variability because of this stratification, see index.

A.2 Notation

$\forall$: whatever

a: range parameter

$B(z)$: Markov-Bayes calibration parameter

$C(0)$: covariance value at separation vector $\mathbf{h} = 0$. It is also the stationary variance of random variable $Z(\mathbf{u})$

$C(\mathbf{h})$: stationary covariance between any two random variables $Z(\mathbf{u})$, $Z(\mathbf{u}+\mathbf{h})$ separated by vector $\mathbf{h}$

$C(\mathbf{u}, \mathbf{u}')$: non-stationary covariance of two random variables $Z(\mathbf{u})$, $Z(\mathbf{u}')$

$C_l(\mathbf{h})$: nested covariance structure in the linear covariance model $C(\mathbf{h}) = \sum_{l=1}^{L} C_l(\mathbf{h})$

$C_I(\mathbf{h}; z_k, z_{k'})$: stationary indicator cross covariance for cutoffs z_k, $z_{k'}$; it is the cross covariance between the two indicator random variables $I(\mathbf{u}; z_k)$ and $I(\mathbf{u}+\mathbf{h}; z_{k'})$

$C_I(\mathbf{h}; z_k)$: stationary indicator covariance for cutoff z_k

$C_I(\mathbf{h}; s_k)$: stationary indicator covariance for category s_k

$C_{ZY}(\mathbf{h})$: stationary cross covariance between the two random variables $Z(\mathbf{u})$ and $Y(\mathbf{u}+\mathbf{h})$ separated by lag vector $\mathbf{h}$

$d\mathbf{u}$: denotes an infinitesimal area (volume) centered at location $\mathbf{u}$

$E\{\cdot\}$: expected value

$E\{Z(\mathbf{u})|(n)\}$: conditional expectation of the random variable $Z(\mathbf{u})$ given the realizations of n other neighboring random variables (called data)

$Exp(\mathbf{h})$: exponential semivariogram model, a function of vector $\mathbf{h}$

$F(\mathbf{u}; z)$: non-stationary cumulative distribution function of random variable $Z(\mathbf{u})$

$F(\mathbf{u}; z|(n))$: non-stationary conditional cumulative distribution function of the continuous random variable $Z(\mathbf{u})$ conditioned by the realizations of n other neighboring random variables (called data)

$F(\mathbf{u}; k|(n))$: nonstationary conditional probability distribution function of the categorical variable $Z(\mathbf{u})$

$F(\mathbf{u}_1, \ldots, \mathbf{u}_K; z_1, \ldots, z_K)$: K variate cumulative distribution function of the K random variables $Z(\mathbf{u}_1), \ldots, Z(\mathbf{u}_K)$

$F(z)$: cumulative distribution function of a random variable Z, or stationary cumulative distribution function of a random function $Z(\mathbf{u})$

$F^{-1}(p)$: inverse cumulative distribution function or quantile function for the probability value $p \in [0, 1]$

$f_k(\cdot)$: function of the coordinates used in a trend model

$\gamma(\mathbf{h})$: stationary semivariogram between any two random variables $Z(\mathbf{u})$, $Z(\mathbf{u}+\mathbf{h})$ separated by lag vector $\mathbf{h}$

$\gamma_I(\mathbf{h}; z)$: stationary indicator semivariogram for lag vector $\mathbf{h}$ and cutoff z; it is the semivariogram of the binary indicator random function $I(\mathbf{u}; z)$

$\gamma_I(\mathbf{h}; p)$: same as above, but the cutoff z is expressed in terms of p quantile with $p = F(z)$

$\gamma_{ZY}(\mathbf{h})$: stationary cross semivariogram between the two random variables $Z(\mathbf{u})$ and $Y(\mathbf{u}+\mathbf{h})$ separated by lag vector $\mathbf{h}$

$G(y)$: standard normal cumulative distribution function

$G^{-1}(p)$: standard normal quantile function: $G(G^{-1}(p)) = p \in [0, 1]$

$\mathbf{h}$: separation vector

$I(\mathbf{u}; z)$: binary indicator random function at location $\mathbf{u}$ and for cutoff z
$i(\mathbf{u}; z)$: binary indicator value at location $\mathbf{u}$ and for cutoff z
$I(\mathbf{u}; s_k)$: binary indicator random function at location $\mathbf{u}$ for category s_k
$i(\mathbf{u}; s_k)$: binary indicator value at location $\mathbf{u}$ and for category s_k
$j(\mathbf{u}; z)$: binary indicator transform arising from constraint interval
λ_α, $\lambda_\alpha(\mathbf{u})$: kriging weight associated to datum α for estimation at location $\mathbf{u}$. The superscripts (OK), (KT), (m) are used when necessary to differentiate between the various types of kriging
M: stationary median of the distribution function $F(z)$
m: stationary mean of the random variable $Z(\mathbf{u})$
$m(\mathbf{u})$: mean at location $\mathbf{u}$, expected value of random variable $Z(\mathbf{u})$; trend component model in the decomposition $Z(\mathbf{u}) = m(\mathbf{u}) + R(\mathbf{u})$, where $R(\mathbf{u})$ represents the residual component model
$m^*_{KT}(\mathbf{u})$: estimate of the trend component at location $\mathbf{u}$
$\mu, \mu(\mathbf{u})$: Lagrange parameter for kriging at location $\mathbf{u}$
$N(\mathbf{h})$: number of pairs of data values available at lag vector $\mathbf{h}$
$\prod_{i=1}^n y_i = y_1 \cdot y_2 \cdots y_n$: product
$p_k = E\{I(\mathbf{u}; s_k)\}$: stationary mean of the indicator of category k
$q(p) = F^{-1}(p)$: quantile function, that is, inverse cumulative distribution function for the probability value, $p \in [0, 1]$
$R(\mathbf{u})$: residual random function model in the decomposition $Z(\mathbf{u}) = m(\mathbf{u}) + R(\mathbf{u})$, where $m(\mathbf{u})$ represents the trend component model
ρ: correlation coefficient $\in [-1, +1]$
$\rho(\mathbf{h})$: stationary correlogram function $\in [-1, +1]$
$\sum_{i=1}^n y_i = y_1 + y_2 + \ldots + y_n$: summation
$\Sigma(\mathbf{h})$: matrix of stationary covariances and cross covariances
σ^2: variance
$\sigma^2_{OK}(\mathbf{u})$: ordinary kriging variance of $Z(\mathbf{u})$
$\sigma^2_{SK}(\mathbf{u})$: simple kriging variance of $Z(\mathbf{u})$
$Sph(\mathbf{h})$: spherical semivariogram function of separation vector $\mathbf{h}$
s_k: kth category
$\mathbf{u}$: coordinates vector
$Var\{\cdot\}$: variance
$Y = \varphi(Z)$: transform function $\varphi(\cdot)$ relating two random variables Y and Z
$Z = \varphi^{-1}(Y)$: inverse transform function $\varphi(\cdot)$ relating random variables Z and Y
Z: generic random variable
$Z(\mathbf{u})$: generic random variable at location $\mathbf{u}$, or a generic random function of location $\mathbf{u}$
$Z^*_{COK}(\mathbf{u})$: cokriging estimator of $Z(\mathbf{u})$
$Z^*_{KT}(\mathbf{u})$: estimator of $Z(\mathbf{u})$ using some form of prior trend model
$Z^*_{OK}(\mathbf{u})$: ordinary kriging estimator of $Z(\mathbf{u})$
$Z^*_{SK}(\mathbf{u})$: simple kriging estimator of $Z(\mathbf{u})$

$\{Z(\mathbf{u}), \mathbf{u} \in A\}$: set of random variables $Z(\mathbf{u})$ defined at each location $\mathbf{u}$ of a zone A

$z(\mathbf{u})$: generic variable function of location $\mathbf{u}$

$z(\mathbf{u}_\alpha)$: z-datum value at location $\mathbf{u}$

z_k: kth cutoff value

$z^{(l)}(\mathbf{u})$: lth realization of the random function $Z(\mathbf{u})$ at location $\mathbf{u}$

$z_c^{(l)}(\mathbf{u})$: lth realization conditional to some neighboring data

$z^*(\mathbf{u})$: an estimate of value $z(\mathbf{u})$

$[z(\mathbf{u})]_E^*$: E-type estimate of value $z(\mathbf{u})$, obtained as an arithmetic average of multiple simulated realizations $z^{(l)}(\mathbf{u})$ of the random function $Z(\mathbf{u})$

Bibliography

[1] S. I. Aanonsen, A. L. Eide, and L. Holden. Optimizing reservoir performance under uncertainty with application to well location. In *1995 SPE Annual Technical Conference and Exhibition*. Society of Petroleum Engineers, October 1995. SPE paper # 30710.

[2] E. Aarts and J. Korst. *Simulated Annealing and Boltzmann Machines*. John Wiley & Sons, New York, 1989.

[3] F. G. Alabert. Stochastic imaging of spatial distributions using hard and soft information. Master's thesis, Stanford University, Stanford, CA, 1987.

[4] F. G. Alabert. Constraining description of randomly heterogeneous reservoirs to pressure test data: a Monte Carlo study. In *SPE Annual Conference and Exhibition, San Antonio*. Society of Petroleum Engineers, October 1989.

[5] F. G. Alabert and B. Corre. Heterogeneity in a complex turbiditic reservoir: Impact on field development. In *66th Annual Technical Conference and Exhibition*, pages 971–984, Dallas, TX, October 1991. Society of Petroleum Engineers. SPE Paper # 22902.

[6] F. G. Alabert and G. J. Massonnat. Heterogeneity in a complex turbiditic reservoir: Stochastic modelling of facies and petrophysical variability. In *65th Annual Technical Conference and Exhibition*, pages 775–790. Society of Petroleum Engineers, September 1990. SPE Paper Number 20604.

[7] J. R. L. Allen. A review of the origin and characteristics of recent alluvial sediments. *Sedimentology*, 5:89–191, 1965.

[8] J. R. L. Allen. Studies in fluviatile sedimentation: an exploratory quantitative model for the architecture of avulsion-controlled alluviate suites. *Sedim. Geol.*, 21:129–147, 1978.

[9] A. S. Almeida and A. G. Journel. Joint simulation of multiple variables with a Markov-type coregionalization model. *Math. Geology*, 26:565–588, 1994.

[10] A. S. D. Almeida. *Joint Simulation of Multiple Variables with a Markov-Type Coregionalization Model*. PhD thesis, Stanford University, Stanford, CA, 1993.

[11] M. P. Anderson. Comment on universal scaling of hydraulic conductivities and dispersivities. *Water Resources Research*, 27(6):1381–1382, 1991.

[12] M. P. Anderson and W. W. Woessner. *Applied Groundwater Modeling*. Academic Press, San Diego, 1992.

[13] T. Anderson. *An Introduction to Multivariate Statistical Analysis*. John Wiley & Sons, New York, 1958.

[14] H. C. Andrews and B. R. Hunt. *Digital Image Restoration*. Prentice Hall, Englewood Cliffs, NJ, 1989.

[15] M. Armstrong. Improving the estimation and modeling of the variogram. In G. Verly et al., editors, *Geostatistics for natural resources characterization*, pages 1–20. Reidel, Dordrecht, Holland, 1984.

[16] K. Aziz, S. Arbabi, and C. V. Deutsch. Why it is so difficult to predict the performance of non-conventional wells? SPE paper # 37048, 1996.

[17] K. Aziz and A. Settari. *Petroleum Reservoir Simulation*. Elsevier Applied Science, New York, 1979.

[18] P. R. Ballin, A. G. Journel, and K. A. Aziz. Prediction of uncertainty in reservoir performance forecasting. *JCPT*, 31(4), April 1992.

[19] W. M. Bashore and U. G. Araktingi. Importance of a geological framework and seismic data integration. In J. M. Yarus and R. L. Chambers, editors, *Stochastic Modeling and Geostatistics: Principles, Methods, and Case Studies*, pages 159–176. AAPG Computer Applications in Geology, No. 3, 1995.

[20] R. P. Batycky, M. J. Blunt, and M. R. Thiele. A 3d field-scale streamline-based reservoir simulator. *SPE Reservoir Engineering*, pages 246–254, November 1997.

[21] R. A. Behrens, M. K. Macleod, T. T. Tran, and A. O. Alimi. Incorporating seismic attribute maps in 3D reservoir models. *SPE Reservoir Evaluation & Engineering*, 1(2):122–126, 1998.

[22] J. P. Benkendorfer, C. V. Deutsch, P. D. LaCroix, L. H. Landis, Y. A. Al-Askar, A. A. Al-AbdulKarim, and J. Cole. Integrated reservoir modeling of a major arabian carbonate reservoir. In *SPE Middle East Oil Show*. Bahrain, March 1995. Society of Petroleum Engineers. SPE paper # 29869.

[23] V. Berteig, K. B. Halvorsen, H. Omre, A. K. Hoff, K. Jorde, and O. A. Steinlein. Prediction of hydrocarbon pore volume with uncertainties. In *63rd Annual Technical Conference and Exhibition*, pages 633–643. Society of Petroleum Engineers, October 1988. SPE Paper Number 18325.

[24] J. Besag. On the statistical analysis of dirty pictures. *J. R. Statistical Society B*, 48(3):259–302, 1986.

[25] H. Beucher, A. Galli, G. Le Loc'h, and C. Ravenne. Including a regional trend in reservoir modelling using the truncated Gaussian method. In A. Soares, editor, *Geostatistics-Troia*, volume 1, pages 555–566. Kluwer, Dordrecht, Holland, 1993.

[26] A. C. Bittencourt and R. N. Horne. Reservoir development and design optmization. In *1997 SPE Annual Technical Conference and Exhibition Formation Evaluation and Reservoir Geology*. Society of Petroleum Engineers, October 1997. SPE paper # 38895.

[27] L. J. Bortoli, F. Alabert, A. Haas, and A. G. Journel. Constraining stochastic images to seismic data. In A. Soares, editor, *Geostatistics Troia 1992*, volume 1, pages 325–334. Kluwer, 1993.

[28] R. B. Bratvold, L. Holden, T. Svanes, and K. Tyler. STORM: Integrated 3D stochastic reservoir modeling tool for geologists and reservoir engineers. SPE paper # 27563, 1994.

[29] J. S. Bridge. A FORTRAN IV program to simulate alluvial stratigraphy. *Computers & Geosciences*, 1979.

[30] J. S. Bridge and M. R. Leeder. A simulation model of alluvial stratigraphy. *Sedimentology*, 26:617–644, 1979.

[31] E. Byers and D. B. Stephens. Statistical and stochastic analyses of hydraulic conductivity and particle-size in a fluvial sand. *Soil Science Society of America Journal*, 47:1072–1081, 1983.

[32] J. Caers. Adding local accuracy to direct sequential simulation. *Mathematical Geology*, 32(7):815–850, 2000.

[33] W. T. Cardwell and R. L. Parsons. Average permeabilities of heterogeneous oil sands. *Trans. AIME*, 160:34–42, 1945.

[34] J. Carr and D. E. Myers. COSIM: A Fortran IV program for coconditional simulation. *Computers & Geosciences*, 11(6):675–705, 1985.

[35] J. Carrera and S. P. Neuman. Estimation of aquifer parameters under transient and steady state conditions: 1. maximum likelihood method incorporating prior information. *Water Resources Research*, 22(2):199–210, 1986.

[36] J. Chu and A. G. Journel. Conditional fBm simulation with dual kriging. In R. Dimitrakopoulos, editor, *Geostatistics for the Next Century*, pages 407–421. Kluwer, 1994.

[37] J. Chu and a. A. G. J. W. Xu. 3-d implementation of geostatistical analyses - the amoco case study. In J. M. Yarus and R. L. Chambers, editors, *Stochastic Modeling and Geostatistics: Principles, Methods, and Case Studies*, pages 201–216. AAPG Computer Applications in Geology, No. 3, 1995.

[38] R. Clemensten, A. R. Hurst, R. Knarud, and H. Omre. A computer program for evaluation of fluvial reservoirs. In Buller et al., editors, *North Sea Oil and Gas Reservoirs II*. Graham and Trotman, London, 1990.

[39] D. L. Cox, S. J. Lindquist, C. L. Bargas, K. G. Havholm, and R. M. Srivastava. Integrated modeling for optimum management of a giant gas condensate reservoir, Jurassic eolian nugget sandstone, Anschutz Ranch East Field, Utah overthrust (USA)'. In J. M. Yarus and R. L. Chambers, editors, *Stochastic Modeling and Geostatistics: Principles, Methods, and Case Studies*, pages 287–321. AAPG Computer Applications in Geology, No. 3, 1995.

[40] N. Cressie. *Statistics for Spatial Data*. John Wiley & Sons, New York, 1991.

[41] N. Cressie and D. Hawkins. Robust estimation of the variogram. *Mathematical Geology*, 12(2):115–126, 1980.

[42] P. Cruz, R. Horne, and C. V. Deutsch. The quality map: A tool for reservoir uncertainty quantification and decision making. In *1999 SPE Annual Technical Conference and Exhibition*. Society of Petroleum Engineers, October 1999.

[43] P. S. da Cruz. *Reservoir Management Decision-Making in the Presence of Geological Uncertainty*. PhD thesis, Stanford University, Stanford, CA, 2000.

[44] M. Dagbert, M. David, D. Crozel, and A. Desbarats. Computing variograms in folded strata-controlled deposits. In G. Verly et al., editors, *Geostatistics for natural resources characterization*, pages 71–89. Reidel, Dordrecht, Holland, 1984.

[45] E. Damsleth, A. Hauge, and R. Volden. Maximum information at minimum cost: A north sea field development study with an experimental design. *Journal of Petroleum Technology*, pages 1350–1356, December 1992.

[46] E. Damsleth and C. B. Tjølsen. Scale consistency from cores to geologic description. *SPE Formation Evaluation*, 9(4):295–299, 1994.

[47] E. Damsleth, C. B. Tjølsen, H. Omre, and H. H. Haldorsen. A two-stage stochastic model applied to a north sea reservoir. *Journal of Petroleum Technology*, pages 402–408, April 1992.

[48] A. Datta-Gupta, L. W. Lake, and G. A. Pope. Characterizing heterogeneous permeability media with spatial statistics and tracer data using sequential simulation annealing. *Math. Geology*, 27(6):763–787, 1995.

[49] M. David. *Geostatistical Ore Reserve Estimation*. Elsevier, Amsterdam, 1977.

[50] J. M. Davis, F. M. Phillips, J. L. Wilson, R. C. Lohman, and D. W. Love. A sedimentological-geostatistical model of aquifer heterogeneity based on outcrop studies. *EOS Trans*, 73(14):122, 1992.

[51] J. P. Dejean and G. Blanc. Managing uncertainties on production predictions using integrated statistical methods. In *1999 SPE Annual Technical Conference and Exhibition*. Society of Petroleum Engineers, October 1999. SPE paper # 56696.

[52] A. J. Desbarats. Numerical estimation of effective permeability in sand-shale formations. *Water Resources Research*, 23(2):273–286, 1987.

[53] A. J. Desbarats. *Stochastic Modeling of Flow in Sand-Shale Sequences.* PhD thesis, Stanford University, Stanford, CA, 1987.

[54] C. V. Deutsch. Calculating effective absolute permeability in sand-stone/shale sequences. *SPE Formation Evaluation*, pages 343–348, September 1989.

[55] C. V. Deutsch. DECLUS: A Fortran 77 program for determining optimum spatial declustering weights. *Computers & Geosciences*, 15(3):325–332, 1989.

[56] C. V. Deutsch. *Annealing Techniques Applied to Reservoir Modeling and the Integration of Geological and Engineering (Well Test) Data.* PhD thesis, Stanford University, Stanford, CA, 1992.

[57] C. V. Deutsch. Algorithmically defined random function models. In Dimitrakopoulos, editor, *Geostatistics for the Next Century*, pages 422–435. Kluwer, 1994.

[58] C. V. Deutsch. Constrained modeling of histograms and cross plots with simulated annealing. *Technometrics*, 38(3):266–274, August 1996.

[59] C. V. Deutsch and P. W. Cockerham. Geostatistical modeling of permeability with annealing cosimulation (acs). In *69th Annual Technical Conference and Exhibition*, pages 523–532, New Orleans, LA, September 1994. Society of Petroleum Engineers. SPE paper # 28413.

[60] C. V. Deutsch and P. W. Cockerham. Practical considerations in the application of simulated annealing to stochastic simulation. *Mathematical Geology*, 26(1):67–82, 1994.

[61] C. V. Deutsch and A. G. Journel. Integrating well test-derived absolute permeabilities. In J. M. Yarus and R. L. Chambers, editors, *Stochastic Modeling and Geostatistics: Principles, Methods, and Case Studies*, pages 131–142. AAPG Computer Applications in Geology, No. 3, 1995.

[62] C. V. Deutsch and A. G. Journel. *GSLIB: Geostatistical Software Library and User's Guide.* Oxford University Press, New York, 2nd edition, 1997.

[63] C. V. Deutsch and H. Kupfersberger. Geostatistical simulation with large-scale soft data. In V. Pawlowsky-Glahn, editor, *Proceedings of IAMG'97*, volume 1, pages 73–87. CIMNE, 1993.

[64] C. V. Deutsch and R. W. Lewis. Advances in the practical implementation of indicator geostatistics. In *Proceedings of the 23rd International APCOM Symposium*, pages 133–148, Tucson, AZ, April 1992. Society of Mining Engineers.

[65] C. V. Deutsch and S. Srinivasan. Improved reservoir management through ranking stochastic reservoir models. In *SPE/DOE Tenth Symposium on Improved Oil Recovery, Tulsa, OK*, pages 105–113, Washington, DC, April 1996. Society of Petroleum Engineers. SPE Paper Number 35411.

[66] C. V. Deutsch and L. Wang. Hierarchical object-based stochastic modeling of fluvial reservoirs. *Math Geology*, 28(7):857–880, 1996.

[67] R. Dimitrakopoulos and A. J. Desbarats. Geostatistical modeling of gridblock permeabilities for 3D reservoir simulators. *SPE Reservoir Engineering*, pages 13–18, February 1993.

[68] P. M. Doyen. Porosity from seismic data: A geostatistical approach. *Geophysics*, 53(10):1263–1275, 1988.

[69] P. M. Doyen, L. D. den Boer, and W. R. Pillet. Seismic porosity mapping in the ekofisk field using a new form of collocated cokriging. In *1996 SPE Annual Technical Conference and Exhibition Formation Evaluation and Reservoir Geology*, pages 21–30, Denver, CO, October 1996. Society of Petroleum Engineers. SPE Paper Number 36498.

[70] P. M. Doyen and T. M. Guidish. Seismic discrimination of lithology in sand/shale reservoirs: A Bayesian approach. Expanded Abstract, SEG 59th Annual Meeting, 1989, Dallas, TX., 1989.

[71] P. M. Doyen, T. M. Guidish, and M. de Buyl. Monte Carlo simulation of lithology from seismic data in a channel sand reservoir. SPE paper # 19588, 1989.

[72] P. M. Doyen, D. E. Psaila, and L. D. den Boer. Reconciling data at seismic and well log scales in 3-d earth modelling. In *1997 SPE Annual Technical Conference and Exhibition Formation Evaluation and Reservoir Geology*, pages 465–474, San Antonio, TX, October 1997. Society of Petroleum Engineers. SPE Paper Number 38698.

[73] P. M. Doyen, D. E. Psaila, and S. Strandenes. Bayesian sequential indicator simulation of channel sands from 3-d seismic data in the oseberg field, norwegian north sea. In *69th Annual Technical Conference and Exhibition*, pages 197–211, New Orleans, LA, September 1994. Society of Petroleum Engineers. SPE paper # 28382.

[74] O. Dubrule. A review of stochastic models for petroleum reservoirs. In Armstrong, editor, *Geostatistics*, volume 2, pages 493–506. Kluwer, 1989.

[75] O. Dubrule. Introducing more geology in stochastic reservoir modeling. In A. Soares, editor, *Geostatistics-Troia*, volume 1, pages 351–370. Kluwer, 1993.

[76] O. Dubrule, M. Thibaut, P. Lamy, and A. Haas. Geostatistical reservoir characterization constrained by seismic data. *Petroleum Geoscience*, 2(2), 1998.

[77] G. Dueck and T. Scheuer. Threshold accepting: A general purpose optimization algorithm appearing superior to simulated annealing. *Journal of Computational Physics*, 90:161–175, 1990.

[78] L. J. Durlofsky. Numerical calculation of equivalent grid block permeability tensors for heterogeneous porous media. *Water Resources Research*, 27(5):699–708, 1991.

[79] L. J. Durlofsky, R. A. Behrens, R. C. Jones, and A. Bernath. Scale up of heterogeneous three dimensional reservoir descriptions. *SPE Journal*, pages 313–326, September 1996.

[80] R. C. Earlougher. *Advances in Well Test Analysis*. Society of Petroleum Engineers, New York, 1977.

[81] B. Efron. *The Jackknife, the Bootstrap, and other Resampling Plans*. Society for Industrial and Applied Math, Philadelphia, 1982.

[82] B. Efron and R. J. Tibshirani. *An Introduction to the Bootstrap*. Chapman & Hall, New York, 1993.

[83] T. Egeland, E. Hatlebakk, L. Holden, and E. A. Larsen. Designing better decisions. In *European Petroleum Conference*, Stavanger, Norway, 1992. SPE paper # 24275.

[84] J. R. Eggleston, S. A. Rojstaczer, and J. J. Peirce. Identification of hydraulic conductivity structure in sand and gravel aquifers: Cape cod data set. *Water Resources Research*, 32(5):1209–1222, 1996.

[85] D. Emery and K. J. Mayers. *Sequence Stratigraphy*. Blackwell Science, London, 1996.

[86] L. M. Fælt, A. Henriquez, L. Holden, and H. Tjelmeland. MOHERES, a program system for simulation of reservoir architecture and properties. In *European Symposium on Improved Oil Recovery*, pages 27–39, 1991.

[87] C. L. Farmer. Numerical rocks. In P. R. King, editor, *The Mathematical Generation of Reservoir Geology*, Oxford, 1992. Clarendon Press. (Proceedings of a conference held at Robinson College, Cambridge, 1989).

[88] F. J. Fayers and T. A. Hewett. A review of current trends in petroleum reservoir description and assessing the impacts on oil recovery. In *Ninth International Conference on Computational Methods in Water Resources*, June 1992.

[89] G. S. Feitosa, L. Chu, L. G. Thompson, and A. C. Reynolds. Determination of reservoir permeability distributions from pressure buildup data. In *1993 SPE Annual Technical Conference and Exhibition Formation Evaluation and Reservoir Geology*, pages 417–429, Houston, TX, October 1993. Society of Petroleum Engineers. SPE Paper Number 26457.

[90] G. S. Feitosa, L. Chu, L. G. Thompson, and A. C. Reynolds. Determination of reservoir permeability distributions from well test pressure data. In *1993 SPE Western Regional Meeting*, pages 189–204, Anchorage, Alaska, May 1993. Society of Petroleum Engineers. SPE Paper Number 26047.

[91] G. E. Fogg, F. J. Lucia, and R. K. Sengen. Stochastic simulation of inter-well scale heterogeneity for improved prediction of sweep efficiency in a carbonate reservoir. In L. W. Lake, H. B. Carroll, and T. C. Wesson, editors, *Reservoir Characterization II*, pages 355–381. Academic Press, 1991.

[92] F. Fournier and J.-F. Derain. A statistical methodology for deriving reservoir properties from seismic data. *Geophysics*, pages 1437–1450, September-October 1995.

[93] R. Froidevaux. Probability field simulation. In A. Soares, editor, *Geostatistics Troia 1992*, volume 1, pages 73–84. Kluwer, 1993.

[94] W. E. Galloway. Clastic depositional systems and sequences: Applications to reservoir prediction, delineation, and characterization. *The Leading Edge*, pages 173–180, 1998.

[95] L. S. Gandin. *Objective Analysis of Meteorological Fields*. Gidrometeorologicheskoe Izdatel'stvo (GIMEZ), Leningrad, 1963. Reprinted by Israel Program for Scientific Translations, Jerusalem, 1965.

[96] G. R. Gavalas, P. C. Shah, and J. H. Seinfeld. Reservoir history matching by bayesian estimation. *SPE Journal*, pages 337–349, December 1976.

[97] D. E. Gawith, P. A. Gutteridge, and Z. Tang. Integrating geoscience and engineering for improved field management and appraisal. In *1995 SPE Annual Technical Conference and Exhibition Formation Evaluation and Reservoir Geology*, pages 11–23, Dallas, TX, October 1995. Society of Petroleum Engineers. SPE Paper Number 29928.

[98] L. W. Gelhar. Stochastic subsurface hydrology from theory to applications. *Water Resources Research*, 22(9):135S–145S, 1986.

[99] L. W. Gelhar, C. Welty, and K. R. Rehfeldt. A critical review on field-scale dispersion in aquifers. *Water Resources Research*, 28(7):1955–1974, 1992.

[100] S. Geman and D. Geman. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-6(6):721–741, November 1984.

[101] F. Georgsen and H. Omre. Combining fibre processes and gaussian random functions for modeling fluvial reservoirs. In A. Soares, editor, *Geostatistics Troia 1992*, volume 2, pages 425–440. Kluwer, 1993.

[102] D. J. Goggin, M. A. Chandler, G. Kocurek, and L. W. Lake. Patterns of permeability in eolian deposits: Page sandstone (jurassic), northeastern arizona. *SPE Formation Evaluation*, 3(2):207–306, 1988.

[103] A. Goldberger. Best linear unbiased prediction in the generalized linear regression model. *JASA*, 57:369–375, 1962.

[104] J. J. Gómez-Hernández. *A Stochastic Approach to the Simulation of Block Conductivity Fields Conditioned upon Data Measured at a Smaller Scale*. PhD thesis, Stanford University, Stanford, CA, 1991.

[105] J. J. Gómez-Hernández, A. Sahuquillo, and J. E. Capilla. Stochastic simulation of transmissivity fields conditional to both transmissivity and piezometric data, 1. the theory. *Journal of Hydrology*, 203(1-4):162–174, 1998.

[106] J. J. Gómez-Hernández and R. M. Srivastava. ISIM3D: An ANSI-C three dimensional multiple indicator conditional simulation program. *Computers & Geosciences*, 16(4):395–410, 1990.

[107] P. Goovaerts. Comparative performance of indicator algorithms for modeling conditional probability distribution functions. *Math. Geology*, 26(3):385–410, 1994.

[108] P. Goovaerts. Comparative performance of indicator algorithms for modeling conditional probability distribution functions. *Math Geology*, 26:389–411, 1994.

[109] P. Goovaerts. Comparison of CoIK, IK and mIK performances for modeling conditional probabilities of categorical variables. In R. Dimitrakopoulos, editor, *Geostatistics for the Next Century*, pages 18–29. Kluwer, 1994.

[110] P. Goovaerts. *Geostatistics for Natural Resources Evaluation*. Oxford University Press, New York, 1997.

[111] E. Gringarten and C. Deutsch. Methodology for variogram interpretation and modeling for improved petroleum reservoir characterization. In *1999 SPE Annual Technical Conference and Exhibition Formation Evaluation and Reservoir Geology*, Houston, TX, October 1999. Society of Petroleum Engineers. SPE paper # 56654.

[112] S. F. Gull and J. Skilling. The entropy of an image. In C. R. Smith and W. T. Gandy Jr., editors, *Maximum Entropy and Bayesian Methods in Inverse Problems*, pages 287–301. Reidel, Dordrecht, Holland, 1985.

[113] R. Gundesø and O. Egeland. SESIMIRA - a new geologic tool for 3-d modeling of heterogeneous reservoirs. In Buller et al., editors, *North Sea Oil and Gas Reservoirs II*. Graham and Trotman, London, 1990.

[114] A. L. Gutjahr. Fast Fourier transforms for random fields. Technical Report No. 4-R58-2690R, Los Alamos, NM, 1989.

[115] A. Haas and O. Dubrule. Geostatistical inversion - a sequential method of stochastic reservoir modeling constrained by seismic data. *First Break*, 12(11):561–569, 1994.

[116] H. H. Haldorsen, P. J. Brand, and C. J. Macdonald. Review of the stochastic nature of reservoirs. In S. Edwards and P. R. King, editors, *Mathematics in Oil Production*, pages 109–209. Clarendon Press, Oxford, 1988.

[117] H. H. Haldorsen and D. M. Chang. Notes on stochastic shales: from outcrop to simulation model. In L. W. Lake and H. B. Caroll, editors, *Reservoir Characterization*, pages 445–485. Academic Press, 1986.

[118] H. H. Haldorsen and E. Damsleth. Stochastic modeling. *Journal of Petroleum Technology*, pages 404–412, April 1990.

[119] H. H. Haldorsen and E. Damsleth. Challenges in reservoir characterization. *Bulliten of the American Association of Petroleum Geologists*, 77(4):541–551, 1993.

[120] H. H. Haldorsen and L. W. Lake. A new approach to shale management in field-scale models. *SPE Journal*, pages 447–457, April 1984.

[121] D. E. Hamilton and T. A. J. (editors). Computer modeling of geologic surfaces and volumes. American Association of Petroleum Geologists, Computer Applications in Geology No. 1, 1992.

[122] J. M. Hammersley and D. C. Handscomb. *Monte Carlo Methods*. John Wiley & Sons, New York, 1964.

[123] J. W. Harbaugh. *Computing Risk for Oil Prospects: Principles and Programs*. Pergamon Press, New York, 1995.

[124] A. S. Hatløy. Numerical facies modeling combining deterministic and stochastic method. In J. M. Yarus and R. L. Chambers, editors. *Stochastic Modeling and Geostatistics: Principles, Methods, and Case Studies*, pages 109–120. AAPG Computer Applications in Geology, No. 3, 1995.

[125] J. Helgesen, I. Magnus, S. Prosser, G. Saigal, G. Aamodt, D. Dolberg. and S. Busman. Comparison of constrained sparse spike and stochastic inversion for porosity prediction at kristin field. *The Leading Edge*, 40. April 2000.

[126] A. Henriquez, K. Tyler, and A. Hurst. Characterization of fluvial sedimentology for reservoir simulation modeling. *SPEFEJ*, pages 211–216. September 1990.

[127] K. M. Hess, S. H. Wolf, and M. A. Celia. Large-scale natural gradient tracer test in sand and gravel, Cape Cod, Massachusetts 3. Hydraulic conductivity variability and calculated macrodispersivities. *Water Resources Research*, 28(8):2011–2027, 1992.

[128] T. A. Hewett. Fractal distributions of reservoir heterogeneity and their influence on fluid transport. SPE paper # 15386, 1986.

[129] T. A. Hewett. Modelling reservoir heterogeneity with fractals. In A. Soares, editor, *Geostatistics-Troia*, volume 1, pages 455–466. Kluwer. 1993.

[130] T. A. Hewett. Fractal methods for fracture characterization. In J. M Yarus and R. L. Chambers, editors, *Stochastic Modeling and Geostatistics: Principles, Methods, and Case Studies*, pages 249–260. AAPG Computer Applications in Geology, No. 3, 1995.

[131] T. A. Hewett and C. V. Deutsch. Challenges in reservoir forecasting *Math. Geology*, 28(7):829–842, 1996.

[132] R. J. Hoeksema and P. K. Kitanidis. Analysis of the spatial structure of properties of selected aquifers. *Water Resources Research*, 21(4):563-572, 1985.

[133] R. N. Horne. *Modern Well Test Analysis. A Computer-Aided Approach* Petroway Inc, 926 Bautista Court, Palo Alto, CA, 94303, 1995. Second Edition.

[134] K. Hove, G. Olsen, S. Nilsson, M. Tonnesen, and A. Hatloy. From stochastic geological description to production forecasting in heterogeneous layered reservoirs. In *SPE Annual Conference and Exhibition, Washington, DC*, Washington, DC, October 1992. Society of Petroleum Engineers. SPE Paper Number 24890.

[135] L. Y. Hu, G. Blanc, and B. Noetinger. Gradual deformation and iterative calibration of sequential stochastic simulations. In S. Lippard, A. Nass, and R. Sinding-Larsen, editors, *Proceedings of IAMG'99*. Trondheim, Norway, 1999.

[136] Intera ECL Petroleum Technologies, Highlands Farm, Greys Road, Henley-on-Thames, Oxon, UK. *ECLIPSE 100 Reference Manual*, 1991.

[137] E. H. Isaaks. *The Application of Monte Carlo Methods to the Analysis of Spatially Correlated Data*. PhD thesis, Stanford University, Stanford, CA, 1990.

[138] E. H. Isaaks and R. M. Srivastava. *An Introduction to Applied Geostatistics*. Oxford University Press, New York, 1989.

[139] P. Jacquard and C. Jain. Permeability distribution from field pressure data. *SPE Journal*, pages 281–294, December 1965.

[140] J. L. Jensen, P. W. M. Corbett, G. E. Pickup, and P. S. Ringrose. Permeability semivariograms, geological structure, and flow performance. *Math. Geology*, 28(4):419–435, 1996.

[141] N. L. Johnson and S. Kotz. *Continuous Univariate Distributions - 1*. John Wiley & Sons, New York, 1970.

[142] A. D. W. Jones, S. Al-Qabandi, C. E. Reddick, and S. A. Anderson. Rapid assessment of pattern waterflooding uncertainty in a giant oil reservoir. In *1997 SPE Annual Technical Conference and Exhibition Formation Evaluation and Reservoir Geology*. Society of Petroleum Engineers, October 1997. SPE paper # 38890.

[143] T. A. Jones, D. E. Hamilton, and C. R. Johnson. *Contouring Geologic Surfaces with the Computer*. Van Nostrand Reinhold, New York, 1986.

[144] A. G. Journel. Non-parametric estimation of spatial distributions. *Mathematical Geology*, 15(3):445–468, 1983.

[145] A. G. Journel. Constrained interpolation and qualitative information. *Mathematical Geology*, 18(3):269–286, 1986.

[146] A. G. Journel. Geostatistics: Models and tools for the earth sciences. *Math. Geology*, 18(1):119–140, 1986.

[147] A. G. Journel. *Fundamentals of Geostatistics in Five Lessons*. Volume 8 Short Course in Geology. American Geophysical Union, Washington, DC, 1989.

[148] A. G. Journel. Resampling from stochastic simulations. *Environmental and Ecological Statistics*, 1:63–84, 1994.

[149] A. G. Journel. The abuse of principles in model building and the quest for objectivity: Opening keynote address. In *Fifth International Geostatistics Congress*, Wollongong, Australia, September 1996.

[150] A. G. Journel. Markov models for cross-covariances. *Math. Geology*, 31(8):955–964, 1999.

[151] A. G. Journel and F. G. Alabert. Focusing on spatial connectivity of extreme valued attributes: stochastic indicator models of reservoir heterogeneities. SPE Paper # 18324, 1988.

[152] A. G. Journel and F. G. Alabert. Non-Gaussian data expansion in the earth sciences. *Terra Nova*, 1:123–134, 1989.

[153] A. G. Journel and F. G. Alabert. New method for reservoir mapping. *Journal of Petroleum Technology*, pages 212–218, February 1990.

[154] A. G. Journel and C. V. Deutsch. Entropy and spatial disorder. *Mathematical Geology*, 25(3):329–355, April 1993.

[155] A. G. Journel, C. V. Deutsch, and A. J. Desbarats. Power averaging for block effective permeability. In *56th California Regional Meeting*, pages 329–334. Society of Petroleum Engineers, April 1986. SPE paper # 15128.

[156] A. G. Journel and J. J. Gómez-Hernández. Stochastic imaging of the Wilmington clastic sequence. *SPEFE*, pages 33–40, March 1993. SPE paper # 19857.

[157] A. G. Journel and C. J. Huijbregts. *Mining Geostatistics*. Academic Press, New York, 1978.

[158] A. G. Journel and E. H. Isaaks. Conditional indicator simulation: Application to a Saskatchewan uranium deposit. *Mathematical Geology*, 16(7):685–718, 1984.

[159] A. G. Journel and W. Xu. Posterior identification of histograms conditional to local data. *Math. Geology*, 26:323–359, 1994.

[160] W. J. Kennedy Jr. and J. E. Gentle. *Statistical Computing*. Marcel Dekker, Inc, New York, 1980.

[161] A. Khan, D. Horowitz, A. Liesch, and K. Schepel. Semi-amalgamated thinly-bedded deepwater gom turbidite reservoir performance modeling using object-based technology and bouma lithofacies. In *1996 SPE Annual Technical Conference and Exhibition Formation Evaluation and Reservoir Geology*, pages 443–455, Denver, CO, October 1996. Society of Petroleum Engineers. SPE paper # 36724.

[162] S. Kirkpatrick, C. D. Gelatt Jr., and M. P. Vecchi. Optimization by simulated annealing. *Science*, 220(4598):671–680, May 1983.

[163] M. G. Kittridge, L. W. Lake, F. J. Lucia, and G. E. Fogg. Outcrop/subsurface comparisons of heterogeneity in the san andres formation. SPE paper # 19596, 1990.

[164] C. E. Koltermann and S. M. Gorelick. Heterogeneity in sedimentary deposits: A review of structure-imitating, process-imitating, and descriptive approaches. *Water Resources Research*, 32(9):2617–58, September 1996.

[165] G. Korvin. Axiomatic characterization of the general mixture rule. *Geoexploration*, 19:267–276, 1981.

[166] D. G. Krige. A statistical approach to some mine valuations and allied problems at the Witwatersrand. Master's thesis, University of Witwatersrand, South Africa, 1951.

[167] H. Kupfersberger and C. V. Deutsch. Methodology for integrating analogue geologic data in 3-d variogram modeling. *AAPG Bulliten*, 83(8):1262–1278, 1999.

[168] Landmark Graphics Corporation, 15150 Memorial Drive, Houston, Texas 77079. *Landmark Graphics Suite of Software*, 1999. www.lgc.com.

[169] V. Langlais and J. Doyle. Comparison of several methods of lithofacies simulation on the fluvial Gypsy Sandstone of Oklahoma. In A. Soares, editor, *Geostatistics-Troia*, volume 1, pages 299–310. Kluwer, 1993.

[170] R. J. LeBlanc. Distribution and continuity of sandstone reservoirs - part 1. *JPT*, pages 776–92, July 1977.

[171] R. J. LeBlanc. Distribution and continuity of sandstone reservoirs - part 2. *JPT*, pages 793–804, July 1977.

[172] M. R. Leeder. A quantitative stratigraphic model for alluvium with special reference to channel deposit density and interconnectedness. In A. D. Miall, editor, *Fluvial Sedimentology*, pages 587–596. 1978.

[173] T. Lo and W. M. Bashore. Seismic constrained facies modeling using stochastic seismic inversion and indicator simulation, a north-sea example. Expanded Abstract, SEG 69th Annual Meeting, 1999.

[174] G. L. Loc'h and A. Galli. Truncated pluriGaussian method: Theoretical and practical points of view. In E. Baafi and N. Schofield, editors, *Fifth International Geostatistics Congress*, volume 1, pages 211–222, Wollongong, Australia, 1996. Kluwer.

[175] P. Lowry and A. Raheim. Characterization of delta front sandstones from a fluvial-dominated delta system. In L. W. Lake, H. B. Carroll, and T. C. Wesson, editors, *Reservoir Characterization II*, pages 665–676. Academic Press, 1991.

[176] D. G. Luenberger. *Optimization by Vector Space Methods*. John Wiley & Sons, New York, 1969.

[177] H. J. Lund, H. Ates, E. Kasap, and R. W. Tillman. Comparison of single and multi-facies variograms of newcastle sandstone: Measures for the distribution of barriers to flow. SPE paper # 29596, 1995.

[178] G. R. Luster. *Raw Materials for Portland Cement: Applications of Conditional Simulation of Coregionalization*. PhD thesis, Stanford University, Stanford, CA, 1985.

[179] S. D. Mackey and J. S. Bridge. A revised FORTRAN program to simulate alluvial stratigraphy. *Computers & Geosciences*, 18(2):119–181, 1992.

[180] J. L. Mallet. Structural unfolding. http:www.ensg.u-nancy.frGOCAD, 1999.

[181] J. Mansoori. A review of basic upscaling procedures :advantages and disadvantages. In J. M. Yarus and R. L. Chambers, editors, *Stochastic Modeling and Geostatistics: Principles, Methods, and Case Studies*, pages 65–74. AAPG Computer Applications in Geology, No. 3, 1995.

[182] G. Marsaglia. The structure of linear congruential sequences. In S. K. Zaremba, editor, *Applications of Number Theory to Numerical Analysis*, pages 249–285. Academic Press, London, 1972.

[183] B. Matern. *Spatial Variation*, volume 36 of *Lecture Notes in Statistics*. Springer-Verlag, New York, second edition, 1980. First edition published by Meddelanden fran Statens Skogsforskningsinstitut, Band 49, No. 5, 1960.

[184] G. Matheron. Traité de géostatistique appliquée. Vol. 1 (1962), Vol. 2 (1963). ed. Technip, Paris, 1962.

[185] G. Matheron. Structure et composition des perméabilités. *Revue de l'IFP Rueuil*, 21(4):564–580, 1966.

[186] G. Matheron. La théorie des variables régionalisées et ses applications. Fasc. 5. École Nat. Sup. des Mines, Paris, 1971.

[187] G. Matheron, H. Beucher, H. de Fouquet, A. Galli, D. Guerillot, and C. Ravenne. Conditional simulation of the geometry of fluvio-deltaic reservoirs. SPE paper # 16753, 1987.

[188] N. Metropolis, A. Rosenbluth, M. Rosenbluth, A. Teller, and E. Teller. Equation of state calculations by fast computing machines. *J. Chem. Phys.*, 21(6):1087–1092, June 1953.

[189] A. D. Miall. *The Geology of Fluvial Deposits*. Springer-Verlag, New York, 1996.

[190] J. W. Mulholland. Sequential stratigraphy: Basic elements, concepts, and terminology. *The Leading Edge*, pages 37–40, 1998.

[191] C. Murray. *Geostatistical applications in petroleum geology and sedimentary geology*. PhD thesis, Stanford University, Stanford, CA, 1992.

[192] C. Murray. Indicator simulation of petrophysical rock types. In A. Soares, editor, *Geostatistics-Troia*, volume 1, pages 399–411. Kluwer, 1993.

[193] C. J. Murray. Identification and 3-D modeling of petrophysical rock types. In J. M. Yarus and R. L. Chambers, editors, *Stochastic Modeling and Geostatistics: Principles, Methods, and Case Studies*, pages 323–338. AAPG Computer Applications in Geology, No. 3, 1995.

[194] D. E. Myers. Matrix formulation of co-kriging. *Mathematical Geology*, 14(3):249–257, 1982.

[195] D. E. Myers. Cokriging - new developments. In G. Verly et al., editors, *Geostatistics for natural resources characterization*, pages 295–305. Reidel, Dordrecht, Holland, 1984.

[196] D. E. Myers. Pseudo-cross variograms, positive-definiteness, and cokriging. *Math. Geology*, 23(6):805–816, 1991.

[197] R. A. Olea, editor. *Geostatistical Glossary and Multilingual Dictionary*. Oxford University Press, New York, 1991.

[198] R. A. Olea. Fundamentals of semivariogram estimation, modeling, and usage. In J. M. Yarus and R. L. Chambers, editors, *Stochastic Modeling and Geostatistics: Principles, Methods, and Case Studies*, pages 27–36. AAPG Computer Applications in Geology, No. 3, 1995.

[199] D. S. Oliver. The averaging process in permeability estimation from well test data. *SPE Formation Evaluation*, pages 319–324, September 1990.

[200] D. S. Oliver. Estimation of radial permeability distribution from well test data. In *SPE Annual Conference and Exhibition, New Orleans, LA*, pages 243–250, New Orleans, LA, September 1990. Society of Petroleum Engineers. SPE Paper Number 20555.

[201] D. S. Oliver. Incorporation of transient pressure data into reservoir characterization. *In Situ*, 18(3):243–275, 1994.

[202] H. Omre. *Alternative Variogram Estimators in Geostatistics*. PhD thesis, Stanford University, Stanford, CA, 1985.

[203] H. Omre. Heterogeneity models. In *SPOR Monograph: Recent Advances in Improved Oil Recovery Methods for North Sea Sandstone Reservoirs*, Norway, 1992. Norwegian Petroleum Directorate.

[204] N. Oreskes, K. Shrader-Frechette, and K. Belitz. Verification, validation, and confirmation of numerical models in the earth sciences. *Science*, 263:641–646, February 1994.

[205] G. E. Pickup, J. L. Jensen, P. S. Ringrose, and K. S. Sorbie. A method for calculating permeability tensors using perturbed boundary conditions. In *3rd European Conference on the Mathematics of Oil Recovery*, pages 225–237, Delft, June 1992.

[206] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling. *Numerical Recipes*. Cambridge University Press, New York, 1986.

[207] R. Raghavan. *Well Test Analysis*. PTR Prentice-Hall, Inc., New Jersey, 1993.

[208] P. Renard and G. de Marsily. Calculating equivalent permeability: A review. *Advances in Water Resources*, 20:253–278, 1997.

[209] Reservoir Characterization Research and Consulting, Inc., 9200 E. Mineral Ave., Ste. 101 Englewood, Colorado, 80112 USA. *RC Squared Suite of Software*, 1999. www.rc2.com.

[210] B. D. Ripley. *Spatial Statistics*. John Wiley & Sons, New York, 1981.

[211] B. D. Ripley. *Stochastic Simulation*. John Wiley & Sons, New York, 1987.

[212] B. D. Ripley. *Statistical Inference for Spatial Processes*. Cambridge University Press, New York, 1988.

[213] R. W. Ritzi and D. F. Dominic. Evaluating uncertainty in flow and transport in heterogeneous buried-valley aquifers. In *Groundwater Modeling Conference*, Golden, Colorado, 1993.

[214] N. M. S. Rock. Numerical geology. In S. Bhattacharji et al., editors, *Lecture Notes in Earth Sciences*, volume 18. Springer Verlag, New York, 1988.

[215] C. Rossini, F. Brega, L. Piro, M. Rovellini, and G. Spotti. Combined geostatistical and dynamic simulations for developing a reservoir management strategy: A case history. *Journal of Petroleum Technology*, pages 979–985, November 1994.

[216] D. H. Rothman. Nonlinear inversion, statistical mechanics, and residual statics estimation. *Geophysics*, 50:2784–2796, 1985.

[217] M. A. Sabet. *Well Test Analysis*, volume 8 of *Contributions in Petroleum Geology and Engineering*. Gulf Publishing Company, Houston, 1991.

[218] M. Sams, D. Atkins, N. Siad, E. Parwito, and P. van Riel. Stochastic inversion for high resolution reservoir characterization in the central sumatra basin. In *1999 SPE Asia Pacific Improved Oil Recovery Conference*, Kuala Lumpur, Malaysia, October 1999. Society of Petroleum Engineers. SPE paper # 57260.

[219] E. Schnetzler. Visualization and cleaning of pixel-based images. Master's thesis, Stanford University, Stanford, CA, 1994.

[220] D. W. Scott. *Multivariate Density Estimation: Theory, Practice, and Visualization*. John Wiley & Sons, New York, 1992.

[221] R. K. Senger, F. J. Lucia, C. Kerans, C. Ferris, and G. E. Fogg. Geostatistical/geological permeability characterization of carbonate ramp deposits in San Andres Outcrop, Algerita Escarpment, New Mexico. SPE paper # 23967, 1992.

[222] L. E. Shmaryan and A. G. Journel. Two markov models and their application. *Math. Geology*, 31(8):965–988, 1999.

[223] B. W. Silverman. *Density Estimation for Statistics and Data Analysis*. Chapman and Hall, New York, 1986.

[224] Smedvig Technologies, Inc., 2925 Briarpark Drive, Suite 1000, Houston, Texas, 77042 USA. *Smedvig Suite of Software*, 1999. www.smedtech.com.

[225] L. Smith. Spatial variability of flow parameters in a stratified sand. *Math. Geology*, 13(1):1–21, 1981.

[226] R. M. Srivastava. Minimum variance or maximum profitability? *CIM Bulletin*, 80(901):63–68, 1987.

[227] R. M. Srivastava. A non-ergodic framework for variogram and covariance functions. Master's thesis, Stanford University, Stanford, CA, 1987.

[228] R. M. Srivastava. An application of geostatistical methods for risk analysis in reservoir management. SPE Paper Number 20608, 1990.

[229] R. M. Srivastava. *INDSIM2D: An FSS International Training Tool.* FSS International, Vancouver, Canada, 1990.

[230] R. M. Srivastava. Reservoir characterization with probability field simulation. In *SPE Annual Conference and Exhibition, Washington, DC*, pages 927–938, Washington, DC, October 1992. Society of Petroleum Engineers. SPE Paper Number 24753.

[231] R. M. Srivastava. An annealing procedure for honouring change of support statistics in conditional simulation. In Dimitrakopoulos, editor, *Geostatistics for the Next Century.* Kluwer, 1993.

[232] R. M. Srivastava. An overview of stochastic methods for reservoir characterizat ion. In J. M. Yarus and R. L. Chambers, editors, *Stochastic Modeling and Geostatistics: Principles, Methods, and Case Studies*, pages 3–16. AAPG Computer Applications in Geology, No. 3, 1995.

[233] R. M. Srivastava and H. M. Parker. Robust measures of spatial continuity. In Armstrong, editor, *Geostatistics*, pages 295–308. Reidel, 1989.

[234] F. I. Stalkup. Permeability variations observed at the faces of cross-bedded sandstone outcrops. In L. W. Lake and H. B. Carroll, editors, *Reservoir Characterization*, pages 141–175. Academic Press, 1986.

[235] K. O. Stanley, K. Jorde, N. Raestad, and C. P. Stockbridge. Stochastic modeling of reservoir sand bodies for input to reservoir simulation, Snorre Field, northern North Sea. In Buller et al., editors, *North Sea Oil and Gas Reservoirs II.* Graham and Trotman, London, 1990.

[236] D. Stoyan, W. S. Kendall, and J. Mecke. *Stochastic Geometry and its Applications.* John Wiley & Sons, New York, 1987.

[237] S. Strebelle. *Sequential Simulation: Drawing Structures from Training Images.* PhD thesis, Stanford University, Stanford, CA, 2000.

[238] J. A. Stripe, A. Kazuyoshi, and M. Durandeau. Integrated field development planning using risk and and decision analysis to minimize the impact of reservoir and other uncertainties: A case study. In *Middle East Oil Technical Conference and Exhibition*, pages 155–167. Society of Petroleum Engineers, April 1993. SPE Paper Number 25529.

[239] E. A. Sudicky. A natural gradient experiment on solute transport in a sand aquifer: Spatial variability of hydraulic conductivity and its role in the dispersion process. *Water Resources Research*, 22(13):2069–2082, 1986.

[240] J. A. Sullivan. *Non-parametric Estimation of Spatial Distributions.* PhD thesis, Stanford University, Stanford, CA, 1985.

[241] T-Surf Corporation, 11011 Richmond Avenue, Suite #350, Houston, Texas, 77042 USA. *gOcad suite of software*, 2000. www.t-surf.com.

[242] H. Tjelmeland and H. Omre. Semi-markov random fields. In A. Soares, editor, *Geostatistics Troia 1992*, volume 2, pages 493–504. Kluwer, 1993.

[243] C. B. Tjølsen and G. Jonhsen. Seismic data can improve the stochastic facies modelling significantly. In *1995 SPE Annual Technical Conference and Exhibition Formation Evaluation and Reservoir Geology*, pages 373–380, Dallas, TX, October 1995. Society of Petroleum Engineers. SPE paper # 30567.

[244] L. Tomutsa, S. R. Jackson, and M. Szpakiewicz. Geostatistical characterization and comparison of outcrop and subsurface facies: Shannon shelf sand ridges. SPE paper # 15127, 1986.

[245] C. Torres-Verdin, M. Victoria, G. Merletti, and J. Pendrel. Trace-based and geostatistical inversion of 3-d seismic data for thin sand delineation: An application to san jorge basin, Argentina. *The Leading Edge*, 40, September 1999.

[246] T. T. Tran. *Stochastic Simulation of Permeability Fields and Their Scale-Up for Flow Modeling*. PhD thesis, Stanford University, Stanford, CA, 1995.

[247] T. T. B. Tran. *Stochastic Simulation of Permeability Fields and Their Scale-Up for Flow Modeling*. PhD thesis, Stanford University, Stanford, CA, 1995.

[248] T. T. B. Tran. The missing scale and direct simulation of block effective properties. *Journal of Hydrology*, 182:37–56, 1996.

[249] T. T. B. Tran, X.-H. Wen, and R. A. Behrens. Efficient conditioning of 3d fine-scale reservoir model to multiphase production data using streamline-based coarse-scale inversion and geostatistical downscaling. In *1999 SPE Annual Technical Conference and Exhibition*. Society of Petroleum Engineers, October 1999.

[250] K. Tyler, A. Henriquez, F. Georgsen, L. Holden, and H. Tjelmeland. A program for 3d modeling of heterogeneities in a fluvial reservoir. In *3rd European Conference on the Mathematics of Oil Recovery*, pages 31–40, Delft, June 1992.

[251] K. Tyler, A. Henriquez, A. MacDonald, T. Svanes, and A. L. Hektoen. MOHERES - a collection of stochastic models for describing heterogeneities in clastic reservoirs. In *3rd International Conference on North Sea Oil and Gas Reservoirs III*, pages 213–221, 1992.

[252] K. Tyler, A. Henriquez, and T. Svanes. Modeling heterogeneities in fluvial domains: A review on the influence on production profile. In J. M. Yarus and R. L. Chambers, editors, *Stochastic Modeling and Geostatistics: Principles, Methods, and Case Studies*, pages 77–89. AAPG Computer Applications in Geology, No. 3, 1995.

[253] K. Tyler, T. Svanes, and A. Henriquez. Heterogeneity modelling used for production simulation of fluvial reservoir. *SPE Formation Evaluation*, pages 85–92, June 1992.

[254] N. Tyler, M. D. Barton, and R. J. Finley. Outcrop characterization of flow unit and seal properties and geometries, Ferron sandstone, Utah. SPE paper # 22670, 1991.

[255] S. Vasantharajan and A. S. Cullick. Well site selection using integer programming optimization. In *Third Annual Conference*, volume 1, pages 421–426, International Association for Mathematical Geology. Vera Pawlowsky Glahn (ed.), CIMNE press, Barcelona, Spain, September 1997.

[256] V. Černy. Thermodynamical approach to the travelling salesman problem: an efficient simulation algorithm. *Journal of Optimization Theory and Applications*, 45:41–51, 1985.

[257] G. Verly. *Estimation of Spatial Point and Block Distributions: The MultiGaussian Model.* PhD thesis, Stanford University, Stanford, CA, 1984.

[258] G. Vincent, B. Corre, and P. Thore. Managing structural uncertainty in a mature field for optimal well placement. In *1990 SPE Annual Technical Conference and Exhibition.* Society of Petroleum Engineers, October 1990. SPE paper # 48953.

[259] H. Wackernagel. Geostatistical techniques for interpreting multivariate spatial information. In C. Chung et al., editors, *Quantitative Analysis of Mineral and Energy Resources*, pages 393–409. Reidel, 1988.

[260] J. C. V. Wagoner, R. M. Mitchum, K. M. Campion, and V. D. Rahmanian. *Siliclastic Sequence Stratigraphy in Well Logs, Cores, and Outcrops: Concepts for High-Resolution Correlation of Time Facies.* The American Association of Petroleum Geologists, Tulsa, Oklahoma, 1990.

[261] A. T. Watson, J. H. Seinfelf, G. R. Gavalas, and P. T. Woo. History matching in two-phase petroleum reservoirs. *SPE Journal*, pages 521–532, December 1980.

[262] E. K. Webb. *Simulating the spatial heterogeneity of sedimentological and hydrogeological characteristics for braided stream deposits.* PhD thesis, University of Wisconsin, Madison, Madison, WI, 1992.

[263] K. J. Weber. Influence of common sedimentary structures on fluid flow in reservoir models. *JPT*, pages 665–672, March 1982.

[264] K. J. Weber and dL. C. Van Geuns. Framework for constructing clastic reservoir simulation models. *JPT*, pages 1248–1253,1296–1297, October 1990.

[265] X.-H. Wen, J. E. Capilla, C. V. Deutsch, J. J. Gómez-Hernández, and A. S. Cullick. Ssc: A fortran program to create permeability fields that honor single-phase flow rate and pressure data. *Computers & Geosciences*, 25:217–230, 1999.

[266] X.-H. Wen, C. V. Deutsch, and A. S. Cullick. High-resolution reservoir models integrating multiple well pressure transient data. *SPE Journal*, pages 344–355, 1998.

[267] X. H. Wen and J. J. Gómez-Hernández. Upscaling of hydraulic conductivity in heterogeneous media: An overview. *Journal of Hydrology*, 183:ix–xxxii, 1996.

[268] C. D. White and R. N. Horne. Computing absolute transmissibility in the presence of fine-scale heterogeneity. In *Ninth SPE Symposium on Reservoir Simulation*, pages 209–220, San Antonio, TX, February 1987. Society of Petroleum Engineers. SPE paper # 16011.

[269] R. S. Wikramaratna. ACORN - a new method for generating sequences of uniformly distributed pseudo-random numbers. *Journal of Computational Physics*, 83:16–31, 1989.

[270] A. D. Woodbury and E. A. Sudicky. The geostatistical characteristics of the borden aquifer. *Water Resources Research*, 27(4):533–546, 1991.

[271] W. Xu. *Stochastic Modeling of Reservoir Lithofacies and Petrophysical Properties*. PhD thesis, Stanford University, Stanford, CA, 1995.

[272] W. Xu and A. G. Journel. Gtsim: Gaussian truncated simulations of reservoir units in a West Texas carbonate field. SPE paper # 27412, 1993.

[273] W. Xu and A. G. Journel. Histogram and scattergram smoothing using convex quadratic programming. *Math Geology*, 27:83–103, 1995.

[274] W. Xu, T. T. Tran, R. M. Srivastava, and A. G. Journel. Integrating seismic data in reservoir modeling: the collocated cokriging alternative. In *67th Annual Technical Conference and Exhibition*, pages 833–842, Washington, DC, October 1992. Society of Petroleum Engineers. SPE paper # 24742.

[275] C. T. Yang, A. K. Chopra, and J. Chu. Integrated geostatistical reservoir description using petrophysical, geological, and seismic data for Yacheng 13-1 gas field. In *1995 SPE Annual Technical Conference and Exhibition Formation Evaluation and Reservoir Geology*, pages 357–372, Dallas, TX, October 1995. Society of Petroleum Engineers. SPE paper # 30566.

[276] T. Yao and A. G. Journel. Automatic modeling of (cross) covariance tables using fast fourier transform. *Mathematical Geology*, 30(6):589–615, 1998.

[277] J. M. Yarus. Selected readings on geostatistics. In J. M. Yarus and R. L. Chambers, editors, *Stochastic Modeling and Geostatistics: Principles, Methods, and Case Studies*, pages 369–370. AAPG Computer Applications in Geology, No. 3, 1995.

[278] W. W.-G. Yeh. Review of parameter identification procedures in groundwater hydrology: The inverse problem. *Water Resources Research*, 22(2):95–108, 1986.

[279] H. Zhu. *Modeling Mixture of Spatial Distributions with Integration of Soft Data*. PhD thesis, Stanford University, Stanford, CA, 1991.

[280] H. Zhu and A. G. Journel. Formatting and integrating soft data: Stochastic imaging via the Markov-Bayes algorithm. In A. Soares, editor, *Geostatistics Troia 1992*, volume 1, pages 1–12. Kluwer, 1993.

Index

accuracy, 300
additivity of variance, 326
anisotropy,
    geometric, 109, 116
    horizontal to vertical anisotropy, 128,131
    zonal, 119
analogue data, 9, 130
annealing, *see* simulated annealing
areal trends, 119, 182
artifacts in annealing, 291
autocovariance, *see* covariance
azimuth angle, 108

back transformation, *see* transformation
Bayes law, 296
Bayesian updating, 85, 201
benefits of geostatistics, 12
bias, *see* unbiasedness
binary variable, 102
bivariate distribution, 41
block cokriging, 202, 253
Boolean simulation, *see* simulation
bootstrap, 65

calibration
    for declustering, 61
    seismic to facies, 198
    seismic to porosity, 256
Cartesian grid, 80
categorical variable statistics, 41
cdf, *see* cumulative distribution function
cell based modeling, 193
cell declustering, *see* declustering
cell size, 82
change of support, 326
checking distributions of uncertainty, 289
cleaning cell-based facies models, 210
clustering, *see* declustering
coefficient of variation, 40
cokriging, 160, 207
collocated cokriging, 161
collocated cosimulation, 252
combination gridding, 88
component objective functions, 279
conditional distributions, 259
conditional expectation, 259
conditional simulation, *see* simulation
continuous variable, 39
constraint interval, 170
coordinate rotation, 84, 236
coordinate system, 80
coordinate transformation, 88, 233
correlation, 42
coregionalization, 147
correlogram, 121
covariance
    cross, 143
    definition, 42, 104
    matrix, 158
    rank, 43
    reproduction, 163
crevasse, 234
cross correlation, 143
cross validation, 298
cross variogram, 139
cumulative distribution function, 34
cyclicity, *see* variogram

data
    consistency, 14
    integration, 9
    types, 12
dampened hole effect variogram model, 134
declustering
    cell, 51
    concept, 50
    multiple variables, 59
    polygonal, 50
    soft data, 59
decision making, 316
despiking, 47
deterministic, *see* trend
direct sequential simulation, 167
directional average, 333
discrete cdf, 40
dispersion variance, 326
distribution, *see* cumulative distribution or probability
drift, *see* trend
dynamic data, 331
dynamic reservoir changes, 10

entropy, 165
ergodicity, *see* fluctuations
estimation, *see* kriging
estimation variance, 155
expected value, *see* trend, conditional expectation
exponential variogram model, 133
external drift, 160

facies, 6, 32, 193
faults, 93
flow simulation, 332
fluctuations (of the simulated statistics), 165, 291
fluvial modeling, 230
fractal, 122
fuzzy closeness, 308

Gaussian
    distribution, 36
    transformation, *see* transformation
    variogram model, 133
geological populations, 32
geological shapes, 225
geometric anisotropy, *see* anisotropy
geometric average, 41, 333
gOcad, 84, 108
graphical transformation, *see* transformation
grids
    correlation, 88
    stratigraphic, 85
    triangulated, 84
gridding, 79

h-scatterplot, 105
hard data, 7
harmonic average, 333
histogram
    definition, 33
    smoothing, 64
hole effect variogram model, 133
horizontal variogram, 124
how many realizations, 309

indicator
    for categorical variable, 171
    choice of thresholds, 169
    for continuous variable, 168
    variogram, 170
    facies modeling, 204
    statistics, 41
inequality data, 170
invariance, *see* stationarity
inversion, 14
isopach, 80
isochore, 80
isotropy, *see* anisotropy

jackknife, 298
joint distribution, 102

kernel function, 64
krige's relation, 326

kriging
cokriging, *see* cokriging
external drift, *see* external drift
ordinary, *see* ordinary kriging
simple, *see* simple kriging
with soft data, 160
with a trend model, 160
variance, 159

LS, *see* least squares
layercake model, 16
lag
distance, 108,110
direction, 108
histogram, 113
tolerance, 111
spacing, 111
least squares, 154, 257
levee, 234
linear model of coregionalization, 144
linear regression, *see* cokriging, kriging
lithofacies modeling, 193, 223
lobe simulation, 238
locally varying mean, 182, 200, 251
locally varying directions, 178
lognormal distribution, 36
loss (function), 318

MAP variant in annealing, 290
MAPS cleaning, 211
marked point processes, *see* object-based modeling
Markov-Bayes, 201, 266
Markov model of coregionalization, 147, 161
mean, 39, 104
median, 38
metropolis algorithm, 276
missing scale, 5, 246, 326
modeling flow chart, 21
modeling scale, 7
Monte Carlo methods, 66
motivation for reservoir models, 10
multiGaussian model, *see* Gaussian
multinormal, *see* multiGaussian model
multiple point statistics, 281, 330
multivariate distribution, 102

net facies or reservoir, 194, 314
nested indicators, 203
nested structures, 134
non-centered covaraince, 103
non-conditional simulation, *see* simulation
non-ergodic correlogram, *see* correlogram
non-ergodic covariance, *see* covariance
non-parametric, 37
non-representative data, *see* declustering
non-stationary, *see* trend
normal
distribution, 36
equations, *see* kriging
probability plot, 36
score transform, *see* transformation, Gaussian
notation, 33
number of realizations, 249, 309
numerical modeling, 8
nugget effect, 116, 133

OK, *see* ordinary kriging
objective function, 279
object-based modeling, 223
offlap, *see* combination gridding
optimal estimate, 154
onlap, 88
order relations, 171
ordinary kriging, 160
outlier resistance in variograms, 107

P-P plot, 44
p-field simulation, 175
parametric distribution, 36
periodicty, *see* variogram cyclicity
permeability, 6, 246
perturbation mechanism, 284
petrophysical properties, 6

polygonal declustering, *see* declustering
porosity, 6, 256
positive definiteness, 131
power average, 334
precision, 300
probability distribution, 35
probability
  density function, 35
  discrete variable, 39
  distribution, *see* cumulative distribution function
  interval, 38
proportion curves, 183, 207
proportional gridding, 88

Q-Q plot, 44
qualitative information, *see* soft data
quantiles, 38
quantile transformation, 45

RF, *see* random function
RV, *see* random variable
random function, 102
random number generation, 66
random path, 164
random variable, 102
ranking, 313
range, 114
rank, 43, 313
realization, *see* simulation
redundancy, 156
regression
  see kriging, cokriging
  porosity/permeability, 257
regridding, 20
representative data, 59
residuals from mean, 186
rotation of coordinates, *see* coordinate rotation

sGs, *see* sequential Gaussian simulation
SK, *see* simple kriging
scale up, 326, 333
scatterplot smoothing, 64
scenario-based modeling, 296
seismic
  calibration to facies, 197
  calibration to porosity, 250
sequence, 85
semivariogram, *see* variogram
semirodogram, *see* rodogram
sequential Gaussian simulation, 162
sequential indicator simulation, 174, 196
sequential simulation, *see* sequential indicator or Gaussian simulation
sill, *see* variogram
simple kriging, 157
simulated annealing
  artifacts, 291
  decision rule, 288
  perturbation, 284
  schedule, 277
  updating, 286
  weighting, 282
simulation,
  annealing, 275
  Boolean, 227
  categorical, *see* indicator
  direct, 167
  Gaussian, 162
  indicator, 174
  p-field, 175
smoothing, 9, 159
smoothing histograms and cross plots, 64
spatial variability, *see* variogram
spherical variogram model, 133
soft data, 7
stationarity, 32, *see* trend and nonstationarity
stochastic image, *see* simulation
stochastic inversion, 253
stochastic shales, 229
stochastic simulation, *see* simulation
straightening functions, 90
stratigraphic coordinates, 88

structural analysis, *see* variogram
support effect, 327
surface-based modeling, 328
systems tract changes, 195

tetrahydra, 84
threshold accepting, 290
tolerance angle, *see* variogram
transformation
    Gaussian, 47
    general (quantile), 45
training image, 130
trend
    building a model, 182
    categorical variable, 187
    directional, 178
    kriging, 181
    variogram, 117, 122
triangulated grids, 84
truncated Gaussian simulation, 204
truncation, 88
two-point statistics, 104
unbiasedness, 154
uncertainty, 8
updating distributions, 85, 201

variable, 33, 39
variance, 39, 104
variogram
    anisotropy, 109, 128
    calculation, 105
    cross, 139
    cyclicity, 116
    definition, 104
    directional, 108
    horizontal, 124
    interpretation, 114
    map, 109
    modeling, 132
    nested structures, 134
    sill, 114

well site selection, 317
work flow diagrams, 26

zonal anisotropy, *see* anisotropy