

2016

## **Statistical Reservoir Characterization, Simulation, and Optimization Of Field Scale-Gas Assisted Gravity Drainage (GAGD) Process with Uncertainty Assessments**

Watheq Jasim Mohammed Al-Mudhafar  
*Louisiana State University and Agricultural and Mechanical College*

Follow this and additional works at: [https://digitalcommons.lsu.edu/gradschool\\_dissertations](https://digitalcommons.lsu.edu/gradschool_dissertations)



Part of the [Petroleum Engineering Commons](#)

---

### **Recommended Citation**

Al-Mudhafar, Watheq Jasim Mohammed, "Statistical Reservoir Characterization, Simulation, and Optimization Of Field Scale-Gas Assisted Gravity Drainage (GAGD) Process with Uncertainty Assessments" (2016). *LSU Doctoral Dissertations*. 3097.  
[https://digitalcommons.lsu.edu/gradschool\\_dissertations/3097](https://digitalcommons.lsu.edu/gradschool_dissertations/3097)

This Dissertation is brought to you for free and open access by the Graduate School at LSU Digital Commons. It has been accepted for inclusion in LSU Doctoral Dissertations by an authorized graduate school editor of LSU Digital Commons. For more information, please contact [gradetd@lsu.edu](mailto:gradetd@lsu.edu).

STATISTICAL RESERVOIR CHARACTERIZATION, SIMULATION, AND OPTIMIZATION  
OF FIELD SCALE-GAS ASSISTED GRAVITY DRAINAGE (GAGD) PROCESS  
WITH UNCERTAINTY ASSESSMENTS

A Dissertation

Submitted to the Graduate Faculty of the  
Louisiana State University and  
Agricultural and Mechanical College  
in partial fulfillment of the  
requirements for the degree of  
Doctor of Philosophy

in

The Craft and Hawkins Department of Petroleum Engineering

by

Watheq Jasim Mohammed Al-Mudhafar  
B.S., University of Baghdad, 2002  
M.c., University of Baghdad, 2006  
May 2016



*This dissertation is dedicated to the Imam Musa bin Ja'far Al Kadhim (PBUH).*

# Acknowledgments

This dissertation would not be possible without several contributions. It is a pleasure to thank Dr. Dandina Rao for his valuable guidance, encouragement and support. Without his supervision and persistent help this dissertation would not have been possible. It is a pleasure also to thank my ex-advisor, Dr. Christopher D. White, for his guidance, especially in Statistical Reservoir Modeling. I would also like to extend my appreciations to Dr. Stephen Sears and Dr. Mayank Tyagi in Department of Petroleum Engineering for their valuable guidance, support, and for being in my committee.

A special thanks to Dr. Brian Marx and Dr. Bin Li in the Department of Experimental Statistics and Dr. Barbara Dutrow in the Department of Geology and Geophysics for their extensive support. I should not forget to thank Dr. Sanjay Srinivasan in Energy and Mineral Engineering at Penn State University for his valuable guidance in Geostatistical Reservoir Modeling.

I would like to present my appreciation to the Institute of International Education for granting me the Fulbright Science and Technology Awards that has funded three years of my PhD program. Great thanks also for the extra financial support from Craft and Hawkins department of Petroleum Engineering and the research grants from US-DOE, Chevron Innovative Research Fund, and LSU-LIFT fund. Finally, I extend many thanks to all my colleagues and friends in LSU for their more than four years fun company.

Special thanks to my family (my late much-lamented parents, as well as my sisters, brothers, and wife) for their prayer and continuous encouraging and support.

# Table of Contents

ACKNOWLEDGMENTS .....	iii
LIST OF TABLES .....	viii
LIST OF FIGURES .....	x
NOMENCLATURE AND ABBREVIATIONS .....	xxiv
ABSTRACT .....	xxvi
 1 INTRODUCTION .....	 1
1.1 Background .....	1
1.2 Research Motivation and Objectives .....	2
1.3 Problem Statement .....	3
1.4 Software Used .....	8
1.4.1 Petrel .....	8
1.4.2 CMG .....	9
1.4.3 R Statistical Language .....	9
1.5 Review of Chapters .....	10
 2 LITERATURE REVIEW .....	 12
2.1 Applied Multivariate Geostatistics .....	12
2.2 Geostatistical Lithofacies Modeling .....	14
2.2.1 Pixel-Based Modeling .....	15
2.2.2 Object-Based Modeling .....	16
2.2.3 Multiple-Point Geostatistics .....	17
2.3 Geostatistical Petrophysical Modeling .....	18
2.4 Geological Uncertainty .....	20
2.5 Gas Assisted Gravity Drainage (GAGD) Process .....	22
2.5.1 GAGD Process Physical Model .....	22
2.5.2 GAGD Field-Scale Application .....	25
2.6 Field Description .....	25
2.6.1 Introduction .....	25
2.6.2 Geological Description .....	26

2.6.3	Depositional Environment .....	28
2.6.4	Study Sector Description .....	29
2.6.5	Production and Injection Schedules .....	30
2.6.6	Boundary conditions treatment .....	31
2.6.7	Data Available .....	32
<b>3</b>	<b>APPLIED MULTIVARIATE GEOSTATISTICS ..</b>	<b>33</b>
3.1	Univariate and Bivariate Data Analysis .....	33
3.2	Multivariate Lithofacies Modeling .....	35
3.2.1	Linear Discriminant Analysis .....	35
3.2.2	Kernel Support Vector Machines .....	39
3.2.3	Probabilistic Neural Networks .....	42
3.3	Multivariate Petrophysical Modeling .....	43
3.3.1	Generalized Linear Modeling .....	44
3.3.2	Smoothed Generalized Additive Modeling .....	45
3.4	Results and Discussion .....	46
3.4.1	Results of Multivariate Lithofacies Modeling .....	48
3.4.2	Results of Multivariate Petrophysical Modeling .....	53
3.5	Summary and Significance .....	66
<b>4</b>	<b>GEOSTATISTICAL RESERVOIR MODELING .....</b>	<b>67</b>
4.1	Variogram-based Algorithms .....	67
4.2	Geostatistical Lithofacies Modeling .....	69
4.2.1	Sequential Indicator Simulation .....	70
4.2.2	Multiple-Point Geostatistics .....	72
4.3	Geostatistical Petrophysical Modeling .....	75
4.3.1	Sequential Gaussian Simulation .....	76
4.3.2	Model-Based Geostatistical Simulation (Bayesian Kriging) .....	77
4.4	Results and Discussion .....	79
4.4.1	Geostatistical Lithofacies Modeling Results .....	80
4.4.2	Geostatistical Petrophysical Modeling Results .....	86
4.5	Cross-Validation Results .....	93
4.6	Realizations Ranking .....	97
4.7	Summary and Significance .....	100
<b>5</b>	<b>INTEGRATED RESERVOIR MODELING .....</b>	<b>102</b>
5.1	Introduction .....	102
5.2	Workflow of Reservoir Simulation Study .....	103
5.2.1	Defining the Simulation Problem .....	103
5.2.2	Static and Dynamic Reservoir Data .....	104
5.2.3	Geological Modeling .....	105
5.2.4	PVT behavior of fluids .....	105
5.2.5	Property Upscaling and Grid System .....	106

5.2.6	Sensitivity Analysis .....	109
5.2.7	GAGD Process Simulation .....	110
5.3	Results and Discussion .....	114
5.3.1	Initialization Results .....	114
5.3.2	History Matching .....	114
5.3.3	Base-Case GAGD Simulation Results .....	118
5.3.4	CGI, WAG, and GAGD Comparisons .....	131
5.4	Summary and Significance .....	133
6	SENSITIVITY ANALYSIS .....	135
6.1	Introduction .....	135
6.2	Design of Experiments .....	135
6.2.1	Latin Hypercube Sampling Approach .....	137
6.3	Model Selection .....	138
6.3.1	Stepwise Elimination .....	139
6.3.2	Bayesian Model Averaging .....	140
6.4	Results and Discussion .....	144
6.4.1	General Reservoir Sensitivity Case .....	144
6.4.2	Specific Reservoir Sensitivity Case .....	155
6.4.3	Comments on A Previous Sensitivity Analysis Study .....	177
6.5	Summary and Significance .....	178
7	GAGD PROCESS OPTIMIZATION .....	180
7.1	Introduction .....	180
7.2	Optimization Approaches .....	180
7.2.1	Latin Hypercube Plus Proxy Model .....	181
7.2.2	Response Surface Methodology .....	181
7.2.3	DECE Algorithm .....	183
7.3	Optimization under Geological Uncertainty .....	184
7.3.1	Robust Optimization .....	185
7.4	Results and Discussion .....	188
7.4.1	Production Control Optimization .....	188
7.4.2	Injection Pressure Optimization .....	197
7.4.3	Optimization Results of Huff-n-Puff GAGD Process .....	207
7.5	Summary and Significance .....	216
8	UNCERTAINTY ASSESSMENTS .....	219
8.1	Introduction .....	219
8.1.1	Box-Behnken Designs .....	220
8.2	Geological Uncertainty .....	221
8.3	Production Uncertainty .....	222
8.4	Heterogeneity and Anisotropy Effects .....	225
8.5	Results and Discussion .....	225
8.5.1	Geological Uncertainty Quantification .....	225

8.5.2	Production Uncertainty Quantification .....	227
8.5.3	Heterogeneity and Anisotropy Effects .....	235
8.6	Summary and Significance .....	238
9	CONCLUSIONS AND RECOMMENDATIONS .....	240
9.1	Conclusions .....	240
9.1.1	Applied Multivariate Geostatistics .....	240
9.1.2	Geostatistical Reservoir Modeling .....	241
9.1.3	Integrated Reservoir Modeling .....	242
9.1.4	Sensitivity Analysis of the GAGD Process .....	244
9.1.5	GAGD Process Optimization .....	245
9.1.6	Uncertainty Assessments of the GAGD Process .....	247
9.2	Recommendations .....	248
	REFERENCES .....	249
	A BIVARIATE STATISTICAL ANALYSIS.....	264
	B MULTIVARIATE PETROPHYSICAL PROPERTIES.....	268
	C GEOSTATISTICAL PROPERTY SIMULATION .....	277
	D HISTORY MATCHING .....	296
	E INDEX .....	303
	F CONFERENCE PROCEEDINGS.....	309
	VITA .....	313

# List of Tables

3.1	Partial t-test for the Full Generalized Linear Model .....	55
3.2	Partial t-test for the Reduced Generalized Linear Model .....	55
3.3	Partial t-test for Full Permeability sGAM Model .....	57
3.4	Partial t-test for the Reduced Permeability sGAM Model .....	58
3.5	Partial t-Test for Full Porosity GLM Model .....	61
3.6	Partial t-Test for reduced Porosity GLM Model.....	61
3.7	Partial t-Test for Full Porosity sGAM Model .....	63
4.1	Indicator Variogram Parameters for All the Lithofacies in the Four Direction .....	82
4.2	Log Porosity Variogram Parameters Given All Lithofacies in the Four Directions .....	88
4.3	Core Porosity Variogram Parameters Given All Lithofacies in the Four Directions .....	89
4.4	Horizontal Permeability Variogram Parameters Given All the Lithofacies in the Four Directions .....	90
4.5	Porosity, Horizontal and Vertical Permeability Variogram Parameters Given the Lithotypes in MPS Facies Model.....	91
5.1	Component and Fluid properties of main pay/South Rumaila Oil Field .....	104
5.2	Capillary pressure and Relative Permeability for Reservoir Lithofacies .....	105
5.3	Reservoir Properties and Initial and Current fluid Volumes in Place .....	114

5.4	Current Cumulative fluid Production and Injection (Prior to GAGD) .....	119
5.5	Base Case Setting of the GAGD Operational Design Parameters .....	119
5.6	CGI, WAG, and GAGD Operational Design Parameters .....	131
6.1	Partial t-Test of Full Linear Model .....	146
6.2	ANOVA Test of Full Linear Model .....	147
6.3	Partial t-Test of Reduced Linear Model .....	149
6.4	ANOVA Test of Reduced Linear Model.....	149
6.5	GLHT of the Full and Reduced Models.....	150
6.6	The 21 Factors Considered in Specific Sensitivity Analysis Case .....	156
6.7	Partial t-Test of Full Linear Model.....	159
6.8	ANOVA Test of Full Linear Model .....	160
6.9	Partial t-Test of Reduced Linear Model .....	162
6.10	ANOVA Test of Reduced Linear Model.....	163
6.11	GLHT of the Full and Reduced Models.....	163
6.12	The First Best Fitted Models through Bayesian Model Averaging .....	166
7.1	Parameters of the Production Controls Optimization .....	188
7.2	Optimal Levels of the Embedded Factors in the Production Controls Optimization in Comparison with the Base Case .....	197
7.3	Decision Parameters of Nominal Huff and Puff GAGD Optimization .....	209
7.4	Optimal Decision Parameters of Robust Huff and Puff GAGD Optimization .....	216
8.1	Quantiles of Parameters in the Geological Uncertainty Quantification .....	227



# List of Figures

1.1	Schematic Drawing of GAGD process (Rao, 2012) .....	2
1.2	The Dissertation Workflow .....	7
1.3	The Dissertation Structure .....	11
2.1	Geographical Location and Current well Locations-Rumaila Oil Field (Al-Ameri et al., 2009) .....	26
2.2	Geological Lithology Column of the Formations in South Rumaila Oil Field (Mohammed et al, 2010).....	27
2.3	Stratigraphic column with lithology (Al-Muhailan et al., 2013) .....	28
2.4	Field's Sectors, Well Locations, and Isobaric Map of South Rumaila Oil Field (Mohammed et al., 2010) .....	30
2.5	Field Production History of South Rumaila Oil Field .....	32
3.1	Estimation of Box-Cox Transformation Power Parameter (Al-Mudhafar, 2014) .....	34
3.2	Linear Discriminant Analysis between Three Classes (Li et al., 2006) .....	37
3.3	Separating Hyperplanes between Two Classes .....	40
3.4	Structure of Probabilistic Neural Networks Algorithm .....	43
3.5	Vertical Lithofacies and Well Logs Distribution .....	47
3.6	Mean Distribution of Permeability and Porosity given Sand, Shaly Sand, and Shale Facies.....	48
3.7	Histogram of Permeability and Porosity given Sand, Shaly Sand, and Shale Facies .....	48

3.8 Relationship between each Lithotype and the Well Log Variables .....	49
3.9 LDA Discrimination of Lithofacies Given the Well Logs Data .....	50
3.10 Scatter Plot of the Discriminated Lithofacies Given the LDA Components .....	50
3.11 The Predicted Discrete and Probability Distribution of the LDA Classified Lithofacies .....	51
3.12 The Observed and Predicted Discrete and Posterior Distribution of KSVM Classified Lithofacies for the Entire Depth Intervals .....	52
3.13 Box-Plot Comparison between Observed and Predicted Discrete Lithofacies Distribution by PNN .....	52
3.14 ScatterMatrix Plot for Predicted Lithofacies and Petrophysical Properties .....	54
3.15 Scatterplot of Observed and Predicted Permeability-Reduced GLM Model .....	56
3.16 Smooth Term Effect of the Continuous Predictors in Full sGAM model .....	57
3.17 Smooth Term Effect of the Continuous Predictors in Reduced sGAM model .....	58
3.18 Scatterplot of Observed and Predicted Permeability-Reduced sGAM Model .....	59
3.19 Observed and Predicted Permeability-Reduced GLM and sGAM Models .....	59
3.20 Full Depth Permeability Matching-Reduced GLM and sGAM Models .....	60
3.21 Scatterplot of Observed and Predicted Core Porosity-Reduced GLM Model .....	62
3.22 Smooth Term Effect of the Continuous Predictors in Full sGAM model .....	62
3.23 Scatterplot of Observed and Predicted Core Porosity-Reduced sGAM Model .....	63
3.24 Matching of Core Porosity-Reduced GLM and sGAM Models . . .	64

3.25 Full Depth Matching of Core Porosity-Reduced GLM and sGAM Models .....	65
4.1 Variogram Structure .....	68
4.2 Three Different Geological Models with Approximately Same variograms (Caers and Zhang, 2004).....	72
4.3 Three Possible Candidate Training Images (Caers and Zhang, 2004) .....	74
4.4 Well Locations of Available Petrophysical Data .....	80
4.5 Zones Structure Gridded Surface for the Main Pay .....	81
4.6 3D Geostatistical Lithofacies Simulation in Four Different Directions .....	83
4.7 Four Main Prior Steps of Multiple-Point Facies Simulation .....	84
4.8 Top: Conventional Sampled Trend Models. Bottom: Spatial Posterior Lithotypes Trend Models .....	85
4.9 Nominal Multiple-Point Geostatistical Simulation of South Rumaila Oil Field/Main Pay .....	86
4.10 Distribution of Wells in the Two Subset Data Groups.....	86
4.11 Geostatistical Log Porosity Simulation Given the Facies in Four Directions .....	88
4.12 Geostatistical Core Porosity Simulation Given the Facies in Four Directions .....	89
4.13 Geostatistical Permeability Simulation Given the Facies in Four Directions .....	90
4.14 3D Petrophysical and Lithofacies Modeling given MPS Simulation .....	91
4.15 Various Variogram Fitting for the Permeability Data .....	92
4.16 Posterior Distributions of Model's Parameters Coefficients .....	93
4.17 Prior and Posterior Distributions of Sampling in Bayesian Kriging .....	94
4.18 2D Distribution of the Nine Quantiles of the Permeability Simulation .....	94
4.19 Histogram of the Nine Quantiles of the Permeability Simulation .....	95

4.20 2D and Histogram Plots of the Ranked Realizations <i>P</i> 10, <i>P</i> 50, and <i>P</i> 90 of Permeability Distribution. ....	95
4.21 Leave-One-Out Cross Validation of Permeability Modeling- Bayesian Kriging .....	96
4.22 Predicted-Measured Permeability and Porosity Given Lithofacies Cross-Plots .....	96
4.23 Cross-Plot between Porosity and Permeability Given Lithofacies .....	97
4.24 Histogram of the 100 Reservoir Realizations with <i>P</i> 10, <i>P</i> 20,..., and <i>P</i> 90 .....	98
4.25 Histogram of the <i>P</i> 10, <i>P</i> 20,..., and <i>P</i> 90 Reservoir Porosity Realizations .....	98
4.26 3D Distribution Images of the Ranked Porosity Stochastic Models .....	99
4.27 Histogram of the <i>P</i> 10, <i>P</i> 20,..., and <i>P</i> 90 Reservoir Permeability Realizations .....	99
4.28 3D Distribution Images of the Ranked Stochastic Permeability Models .....	100
5.1 Coarse-Scale Geostatistical Models for Facies, Porosity, and Permeability .....	107
5.2 Histogram of Lithofacies Distribution of Coarse and Fine Scale Models .....	108
5.3 Histogram of Porosity Distribution of Coarse and Fine Scale Models .....	108
5.4 Histogram of Permeability <i>I</i> Distribution of Coarse and Fine Scale Models .....	109
5.5 Histogram of Permeability <i>K</i> Distribution of Coarse and Fine Scale Models .....	109
5.6 Production and Injection Well Locations in Sand and Shaly- Sand Zones .....	111
5.7 Locations of Horizontal Producers and Vertical Injectors through the GAGD Process .....	112
5.8 Non-Uniform Grid System for CO <sub>2</sub> -GAGD Injectors in Rock Types Map .....	113

5.9 History Matching of Entire Field Production of South Rumaila Oil Field .....	116
5.10 History Matching of Production Wells .....	117
5.11 History Matching of Entire Field Injection of South Rumaila Oil Field .....	117
5.12 History Matching of Injection Wells .....	118
5.13 Comparison of Field Oil Response between Primary Production and Base Case GAGD Process .....	120
5.14 3D Pressure and Ternary Fluid Distributions for the Primary Production Case at 2016 (top) and 2026 (bottom) .....	121
5.15 Effectiveness Comparison of the 3D Pressure and Fluid Distributions for the Base-Case GAGD Implementation at 2026 .....	122
5.16 Gas Oil Ratio and Average Pressure for the Base-Case GAGD Process .....	122
5.17 Comparison of Field Flow Response between Immiscible and Miscible GAGD Process for 25 Years of Production .....	124
5.18 Gas Oil Ratio and Average Reservoir Pressure for Immiscible and Miscible Special Cases of GAGD Process .....	125
5.19 Side View of Immiscible Injected Gas Advancement in year 2021 .....	125
5.20 Side View of Immiscible Injected Gas Advancement in year 2026 .....	126
5.21 Side View of Immiscible Injected Gas Advancement in year 2035 .....	126
5.22 Side View of Immiscible Injected Gas Advancement in year 2040 .....	126
5.23 Side View of Miscible Injected Gas Advancement in year 2021 .....	127
5.24 Side View of Miscible Injected Gas Advancement in year 2026 .....	127
5.25 Side View of Miscible Injected Gas Advancement in year 2035 .....	127
5.26 Side View of Miscible Injected Gas Advancement in year 2040 .....	128
5.27 Gas and Oil Saturation Distributions in 2016 (Before GAGD) .....	128

5.28 Gas and Oil Saturation Distributions in Immiscible Flooding (2021-2040) .....	129
5.29 Gas and Oil Saturation Distributions in Miscible Flooding (2021-2040).....	130
5.30 Comparison of Cumulative Oil Production between the CGI, WAG, and GAGD processes 10 Years Prediction Period.....	132
6.1 Latin Hypercube Sampling Design given Two Variables(R output) .....	137
6.2 The left vertical axis corresponds to the posterior distribution for $\beta_{13}$ and the right vertical axis represents the posterior distribution for $\beta$ equal to 0. The spike corresponds to $P(\beta = 0 D) = 0.38$ (Hoeting et al., 1999). .....	143
6.3 Sensitivity Analysis of Reservoir Parameters on Field Flow Response .....	145
6.4 Field Cumulative Oil Production of the Designed Experiments .....	146
6.5 Component-Residual Plots of Influential Factors .....	147
6.6 Basic Diagnostic Plots of Full Linear Model .....	148
6.7 Component-Residual Plots of Reduced Influential Factors .....	149
6.8 Basic Diagnostic Plots of Reduced Linear Model.....	150
6.9 Nonparametric Bootstrapping of R-squared in Full Linear Model.....	151
6.10 Nonparametric Bootstrapping of Coefficients in Full Lin- ear Model .....	152
6.11 Nonparametric Bootstrapping of R-squared in Reduced Linear Model .....	152
6.12 Influence of Reservoir Parameters on the GAGD Process Performance .....	153
6.13 Effect of Reservoir Permeability on the Field Flow Response .....	154
6.14 Effect of $K_v/K_h$ on the Field Flow Response .....	154
6.15 Effect of Reservoir Porosity on the Field Flow Response .....	155

6.16 Sensitivity Analysis of Reservoir Parameters on Field Flow Response .....	157
6.17 Field Cumulative Oil Production of the Designed Experiments .....	158
6.18 Effects of Full Reservoir Parameters on the GAGD Process Performance .....	159
6.19 Basic Diagnostic Plots of Full Linear Model .....	161
6.20 Basic Diagnostic Plots of Reduced Linear Model .....	162
6.21 Effects of Reduced Reservoir Parameters on the GAGD Process Performance .....	163
6.22 Normal Transformation for all Factors in the Linear Model .....	165
6.23 Component-Residual Plots for the Full BMA Linear Model .....	167
6.24 Best Selected Linear Model By Bayesian Model Averaging .....	168
6.25 Component-Residual Plots for the Reduced BMA Linear Model .....	169
6.26 Effects of BMA-Based Reservoir Parameters on the GAGD Process .....	169
6.27 Nonparametric Bootstrapping of $R^2$ in Stepwise-Reduced Linear Model .....	170
6.28 Nonparametric Bootstrapping of Four Coefficients in Stepwise-Based Reduced Linear Model .....	171
6.29 Nonparametric Bootstrapping of $R^2$ in BMA-Reduced Linear Model .....	171
6.30 Nonparametric Bootstrapping of Four Coefficients in BMA-Reduced Linear Model .....	172
6.31 Validation Effect of Injection Layer Permeability on the Field Flow Response .....	173
6.32 Validation Effect of Permeability of Transition Layers on the Flow Response .....	173
6.33 Validation Effect of Permeability of Production Layer on the Flow Response .....	174

6.34 Validation Effect of Below-Production Layers' Permeability on the Flow Response .....	174
6.35 Validation Effect of Production Layers' $K_v/K_h$ on the Flow Response .....	175
6.36 Validation Effect of Porosity of Production Layers on the Flow Response .....	176
6.37 Validation Effect of Porosity of Other Layers on the Flow Response .....	176
7.1 Flowchart of Design of Experiments-Proxy Optimization Algorithm .....	182
7.2 Flowchart of Robust Optimization Procedure of Huff-and- Puff GAGD Process .....	187
7.3 General and Optimal Solutions in GAGD Production Control Optimization .....	189
7.4 Iterative Validation of Proxy Model in the Production Control Optimization .....	190
7.5 Validation of the Linear and Polynomial Proxy Models For the Production Controls Optimization Case .....	191
7.6 Most Influential Production Factors from the Full Linear Proxy Model .....	191
7.7 Most Influential Production Factors from the Reduced Linear Proxy Model .....	192
7.8 Most Influential Production Factors from Reduced Quadratic Proxy Model .....	192
7.9 Response Surface Model between MAX _BHP and MIN _BHP .....	193
7.10 Response Surface Model between MAX _BHP and WCUT .....	193
7.11 Response Surface Model between MIN _BHP and WCUT .....	194



7.12 Optimal Field Cumulative Oil Production by the GAGD Production Control Optimization Case .....	195
7.13 Optimal Field Oil Production Rate by the GAGD Production Control Optimization Case.....	195
7.14 Comparison of Oil Production between Primary, Base, and Optimal Cases .....	196
7.15 Frequency Charts of Factor Levels in the Production Control Optimization .....	197
7.16 General and Optimal Solutions in GAGD Injection Pressure Optimization .....	199
7.17 Validation of Proxy Model in the Injection Pressure Optimization .....	200
7.18 Validation of the Linear and Polynomial Proxy Models For the Injection Pressure Optimization Case .....	201
7.19 Most Influential Pressure Factors from Linear Proxy Model .....	201
7.20 Most Influential Pressure Factors from Reduced Linear Proxy Model .....	202
7.21 Most Influential Pressure Factors from Reduced Polynomial Proxy Model .....	202
7.22 Response Surface Model MAX _BHP in 2016 and 2017 .....	203
7.23 Response Surface Model MAX _BHP in 2016 and 2018 .....	203
7.24 Response Surface Model MAX _BHP in 2017 and 2018 .....	204
7.25 Optimal Cumulative Oil Production by the GAGD Injection Pressure Optimization Case .....	204
7.26 Optimal Field Oil Production Rate by the GAGD Injection Pressure Optimization Case .....	205
7.27 Comparison of Cumulative Oil Production between Optimization of Injection Pressure and Production Control, with the Base Case .....	206
7.28 Histograms of Factor Levels in the Injection Pressure Optimization .....	206

7.29 Comparison of Cumulative Oil Production between Huff and Puff Case and Production Base Case .....	208
7.30 General and Optimal Objective Function of Nominal Huff-and-Puff GAGD Optimization .....	209
7.31 Cumulative Oil Production from Nominal Huff-and-Puff GAGD Optimization .....	210
7.32 Field Oil Rate from the Nominal Huff and Puff GAGD Optimization .....	211
7.33 Frequency of Optimization Parameters in Huff-and-Puff GAGD optimization .....	211
7.34 Comparison between Base Case and Nominal Optimal Flow Responses of Huff and Puff GAGD Optimization .....	212
7.35 Robust Optimization of Huff and Puff GAGD Process Simulation .....	213
7.36 Optimal and General Cumulative Oil Production of Robust Huff and Puff GAGD Process Simulation .....	213
7.37 Optimal and General Field Oil Rate of Robust Huff and Puff GAGD Process Simulation .....	214
7.38 Frequency of Factor Levels in the Robust GAGD Optimization .....	215
7.39 Comparison of Reservoir Flow Response between Base-Case and Robust GAGD Optimization .....	215
8.1 Box-Behnken design for Three Factors Example (NIST, 2012).....	220
8.2 Flow Chart of Huff and Puff Geological uncertainty.....	222
8.3 Flowchart of Production Huff and Puff Uncertainty Assessment .....	224
8.4 Effect of Geological Uncertainty on Field Cumulative Oil Production .....	225
8.5 Effect of Geological Uncertainty on Field Oil Production Rate .....	226
8.6 The Geological Uncertainty Quantification of Cumulative Oil production .....	226
8.7 Cumulative Distribution Function of Cumulative Oil production .....	227
8.8 Reservoir Flow Responses Under Production Huff and Puff uncertainty .....	228

8.9	Reservoir Flow Responses Under Production Huff and Puff uncertainty .....	228
8.10	Cumulative Probability Distribution of Cumulative Oil Production under Huff and Puff Uncertainty .....	229
8.11	Cumulative Probability Distribution of Cumulative Oil Production under Huff and Puff Uncertainty .....	230
8.12	Probability Distribution of Flow Response as a Function of Injection Periods .....	231
8.13	Probability Distribution of Flow Response as a Function of Soaking Periods .....	232
8.14	Probability Distribution of Flow Response as a Function of Production Periods .....	233
8.15	Probability Distribution of Flow Response as a Function of Minimum BHP .....	234
8.16	Effect of Permeability Variation on Field Cumulative Oil Production .....	235
8.17	Effect of Permeability Variation on Field Oil Production Rate .....	236
8.18	Effect of Permeability Variation on Field Water Cut .....	236
8.19	Effect of Permeability Anisotropy on Field Cumulative Oil Production .....	237
8.20	Effect of Permeability Anisotropy on Field Oil Production Rate .....	238
8.21	Effect of Permeability Anisotropy on Field Water Cut .....	239
A1	Scatter Plots of the Log and Core Porosity for wells IN-083 & IN-090 .....	264
A2	Scatter Plots of the Log and Core Porosity for wells IN-096 & IN-108 .....	264
A3	Scatter Plots of the Log and Core Porosity for wells IN-110 & IN-118 .....	265
A4	Scatter Plots of the Log and Core Porosity for wells IN-126 & IN-132 .....	265
A5	Scatter Plots of the Log and Core Porosity for wells IN-134 & PR-138 .....	266

A6	Scatter Plots of the Log and Core Porosity for wells PR-041 & PR-044 .....	266
A7	Scatter Plots of the Log and Core Porosity for wells PR-045 & PR-060 .....	267
A8	Scatter Plots of the Log and Core Porosity for wells PR-004, PR-016, & PR-227 .....	267
B1	Well Logs, & Lithofacies and Petrophysical Prediction for Well IN-083 .....	268
B2	Well Logs & Lithofacies and Petrophysical Prediction for Well IN-090 .....	268
B3	Well Logs & Lithofacies and Petrophysical Prediction for Well IN-096 .....	269
B4	Well Logs & Lithofacies and Petrophysical Prediction for Well IN-108 .....	269
B5	Well Logs & Lithofacies and Petrophysical Prediction for Well IN-110.....	270
B6	Well Logs & Lithofacies and Petrophysical Prediction for Well IN-118.....	270
B7	Well Logs & Lithofacies and Petrophysical Prediction for Well IN-126.....	271
B8	Well Logs & Lithofacies and Petrophysical Prediction for Well IN-132 .....	271
B9	Well Logs & Lithofacies and Petrophysical Prediction for Well IN-134 .....	272
B10	Well Logs & Lithofacies and Petrophysical Prediction for Well PR-002 ..	272
B11	Well Logs & Lithofacies and Petrophysical Prediction for Well PR-004. ....	273
B12	Well Logs & Lithofacies and Petrophysical Prediction for Well PR-016..	273
B13	Well Logs & Lithofacies and Petrophysical Prediction for Well PR-041. ....	274
B14	Well Logs & Lithofacies and Petrophysical Prediction for Well PR-045 .....	274
B15	Well Logs & Lithofacies and Petrophysical Prediction for Well PR-060....	275
B16	Well Logs & Lithofacies and Petrophysical Prediction for Well PR-158....	275
B17	Well Logs & Lithofacies and Petrophysical Prediction for Well PR-227.....	276
B18	Well Logs & Lithofacies and Petrophysical Prediction for Well PR-244.....	276
C1	Indicator Variogram for the three Facies in Four Directions.....	277
C2	Porosity Variogram Given the Facies in Four Directions.....	278
C3	Horizontal Permeability Variogram Given the Facies in Four Directions .....	279

C4	Vertical Permeability Variogram Given the Facies in Four Directions. ....	280
C5	Log Porosity Variograms Given Sand (top), Shaly Sand (middle), and Shale (bottom) .....	281
C6	Horizontal Permeability Variograms Given Sand (top), Shaly Sand (middle), and Shale (bottom) .....	282
C7	Vertical Permeability Variograms Given Sand (top), Shaly Sand (middle), and Shale (bottom) .....	283
C8	1st Four Set of MPS Realizations given the Two Subset Well Groups: LHS: 9 Wells & RHS: 10 Wells .....	284
C9	2nd Four Set of MPS Realizations given the Two Subset Well Groups: LHS: 9 Wells & RHS: 10 Wells .....	285
C10	Leave-One-Out Cross Validation of MPS Facies Simulation- Wells Group1 .....	286
C11	Leave-One-Out Cross Validation of MPS Facies Simulation- Wells Group2 .....	287
C12	Leave-One-Out Cross Validation of Log Porosity Simulation- Wells Group1 .....	288
C13	Comparison between Leave-One-Out Log Porosity and All Wells-Wells Group1 .....	289
C14	Leave-One-Out Cross Validation of Log Porosity Simulation- Wells Group2 .....	290
C15	Comparison between Leave-One-Out Log Porosity and All Wells-Wells Group2 .....	291
C16	Leave-One-Out Cross Validation of Horizontal Permeability Simulation-Wells Group1 .....	292
C17	Comparison between Leave-One-Out Horizontal Permeability and All Wells-Wells Group1 .....	293
C18	Leave-One-Out Cross Validation of Horizontal Permeability Simulation-Wells Group2 .....	294
C19	Comparison between Leave-One-Out Horizontal Permeability and All Wells-Wells Group2 .....	295

D1	Wells Production History Matching .....	296
D2	Wells Production History Matching .....	296
D3	Wells Production History Matching .....	297
D4	Wells Production History Matching .....	297
D5	Wells Production History Matching .....	298
D6	Wells Production History Matching .....	298
D7	Wells Production History Matching .....	299
D8	Wells Production History Matching .....	299
D9	Wells Production History Matching .....	300
D10	Wells Injection History Matching .....	300
D11	Wells Injection History Matching .....	301
D12	Wells Injection History Matching .....	301
D13	Wells Injection History Matching .....	302

# Nomenclature and Abbreviations

$\mu$	mean.
$\sigma^2$	variance.
$\tau^2$	nugget error in BK algorithm.
$(\Theta)$	the unknown uncertainty space.
$(\theta_d)$	the ranked nine geostatistical realizations for each property.
$\alpha_{ij}$	the coefficient of the interaction terms between every two variables.
$\alpha_{jj}$	the coefficient of the quadratic terms.
$\alpha_j$	the coefficient of linear term.
$\bar{J}$	the expected value equals the average of all realizations.
$\beta$	coefficients in BK algorithm.
$\epsilon(h)$	the effect of nugget in BK algorithm.
$\gamma(h)$	variance between two points with distance $h$ .
$\lambda$	empirical transformation value.
$\lambda$	the kriging weights in Variogram.
$\omega(h)$	the effect of sill and range in BK algorithm.
$\phi$	range of data in BK algorithm.
$\sum_{\theta_d}$	the standard deviation.
$cov(h)$	covariance or similarity between data in distance $h$ .
$E_{\theta_d}$	the expected value over each realization $\theta_d$ .
$h$	spatial lag distance between two points.
$J_{RO}$	the objective function in robust optimization.

$K$       the covariance matrix between known data in Variogram.  
 $k$       the covariance matrix between known and unknown data in Variogram.  
 $N_r$      the total number of ranked realizations.  
 $p(\theta)$    the prior distribution.  
 $p(x; \theta)$  the spatial predictive distribution of parameter  $x$  in BK algorithm.  
 $Pr(M_k)$  the prior probability of the model.  
 $r$       the risk aversion factor.  
 $var(z)$  sill or variation of parameter  $z$ .  
 $x_1, x_2, \dots, x_M$  the separation distance between the known points.  
 $Y(u)$    variance of parameter at point  $u$ .  
 $Y(u + h)$  variance of parameter at point  $u + h$ .  
 $y(x_i)$    the known points of response factors all locations.  
 $y_i$       the response distribution.  
 $z_0$       the predicted value of parameter  $z$  in simple kriging.  
 $(x \cdot p)$    the estimated response factor.



# Abstract

Unlike the Continuous Gas Injection (CGI) and Water-Alternative Gas (WAG), the Gas Assisted Gravity Drainage (GAGD) process takes advantage of the natural segregation of reservoir fluids to provide gravity-stable oil displacement. The gas is injected through existing vertical wells to formulate a gas cap to allow oil and water drain down to the horizontal producer (s). Therefore, the GAGD process was implemented through immiscible and miscible injection modes to improve oil recovery in a sector of the main pay/upper sandstone member in the South Rumaila oil field, located in Iraq.

A high-resolution geostatistical reservoir characterization model was constructed to model the lithofacies and petrophysical properties in order to provide the most realistic geological environment for the GAGD process evaluation. After upscaling the geostatistical model, a compositional reservoir simulation was built to evaluate the GAGD process and test its effectiveness to improve oil recovery.

Next, sensitivity analysis was performed to determine the most influential reservoir parameters that impact the GAGD process. With no effect of reservoir porosity, it was also investigated that reservoir permeability and anisotropy ratio have the most impact on the reservoir flow response through the GAGD process.

Then, several GAGD optimization and uncertainty assessment approaches were conducted to determine the optimal future reservoir scenario of the largest improvement in oil recovery. These optimal cases include operational decision parameters, injection pressure, and cycling injection. It was concluded that the cycling GAGD process has better performance to improve oil recovery than the continuous injection mode. In addition, all the

successive approaches of the optimization and uncertainty assessments led to increase oil production by more than 500 million barrels over the base-case GAGD process of default operational decision parameters.

Finally, both heterogeneity and anisotropy effects have been identified by showing a significant impact on the reservoir flow responses. Specifically, the effect of permeability anisotropy is higher than reservoir heterogeneity because the main concept of the GAGD process considers vertical fluid movements towards the horizontal producers.

Consequently, the overall reservoir characterization, compositional simulation, and optimization of the immiscible GAGD process have shown its effectiveness to improve oil recovery in a real field-scale evaluation.

# 1. Introduction

## 1.1 Background

The Gas Assisted Gravity Drainage (GAGD) process has been suggested for improved oil recovery in secondary and tertiary modes for both immiscible and miscible gas flooding processes. The GAGD process concept states placing a horizontal producer at the bottom of a payzone above the oil-water contact. Then, the gas is injected either immiscible or miscible in a gravity-stable mode through the vertical wells from the top of formation (Rao et al., 2004). Due to the gravity segregation resulting from the distinct fluid densities at reservoir conditions, the injected gas accumulates at the top of payzone providing gravity stable oil displacement that drains down towards the horizontal producers (Mahmoud and Rao, 2007; Mahmoud and Rao, 2008). The schematic drawing of the GAGD process is shown in Figure 1.1:

In the GAGD process, the formulated gas cap, the fluids gravity segregation and the oil drainage towards the bottom of a payzone lead to better sweep efficiency and higher oil recovery. The  $CO_2$  gas is preferred for injection because it attains high volumetric sweep efficiency with high microscopic displacement efficiency, especially in miscible injection mode. Additionally, the high volumetric sweep efficiency assures delaying  $CO_2$  breakthrough to the producer (Rao et al., 2006). Delaying or eliminating the gas breakthrough results in diminishing concurrent gas-liquid flow, and then leads to increased gas injectivity and maintains the injection pressure.

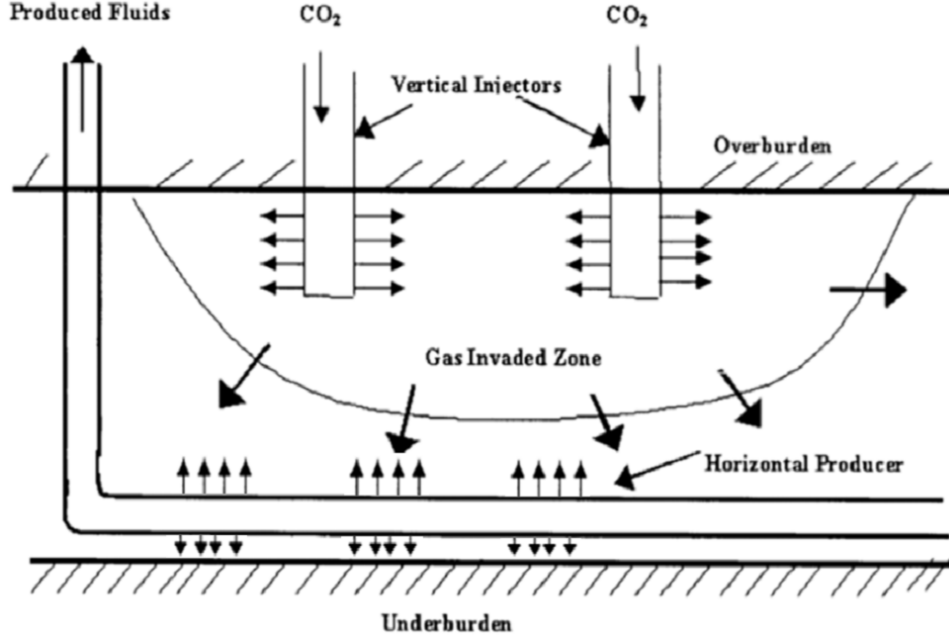


Figure 1.1: Schematic Drawing of GAGD process (Rao, 2012)

Since the GAGD process includes vertical gas injection with vertical oil drain towards the horizontal producers, many factors might be affecting its performance, especially the geological structure in addition to the petrophysical properties and rock facies. Consequently, geostatistical reservoir quantification needs to be conducted to achieve optimal reservoir modeling, recovery optimization, as well as uncertainty quantification.

## 1.2 Research Motivation and Objectives

One of the most crucial problems in petroleum engineering is to determine the optimal future reservoir performance that achieves higher cumulative oil recovery and/or Net Present Value (NPV). Optimization of production performance in a mature oil field requires setting up the most feasible process of IOR/EOR. The Gas Assisted Gravity Drainage (GAGD) process is one of the well-studied EOR techniques that has proved its feasibility and effectiveness to improve oil recovery by handling many different experimental media for fluid flow mechanisms (Rao et al., 2006; Rao, 2012). The GAGD process takes the advantages of the natural tendency of reservoir fluid gravity segregation to recover the bypassed oil from unswept reservoir regions. Therefore, it ensures improved oil recovery

through accomplishing better sweep efficiency and higher microscopic displacement. This process should also result in water-cut reduction to the minimum levels as the horizontal producer should be located above the oil-water contacts at the bottom of the payzone. Additionally, the literature on laboratory and field studies has reported that the GAGD process has led to the improvement of oil recovery to around 100% in laboratory tests and 87-95% in field-scale applications (Kulkarni and Rao, 2006).

The GAGD process was suggested for an integrated high quality geostatistical and compositional reservoir simulation study to improve oil recovery in the main pay of South Rumaila Oil Field, located in Iraq. It was proposed for the application on this field to test its feasibility and effectiveness to improve oil recovery and reduce the high water cut values that reached to more than 95% in many production wells. Since the GAGD process concept is established by installing a horizontal producer at the bottom of the payzone, there is a good opportunity to decrease the water cut values to lower levels. Field-scale application requires wide investigation for all the uncertain geological and reservoir properties. In addition, the controllable operational decision factors may have influential impacts on the flow response. Therefore, the first task was to build a full field geostatistical reservoir characterization model in order to capture the most realistic heterogeneity model. This also provides a solid basis to evaluate and optimize the reservoir performance and to investigate the most influential geological and production factors affecting the GAGD process performance. Evaluation of all these factors can be accomplished by incorporating the geostatistical reservoir characterization models into the compositional reservoir simulation and Design of Experiments.

### **1.3 Problem Statement**

The objective of this dissertation is to build an efficient multi-task geostatistical model for all the reservoir zones and layers of the Upper Sandstone formation in the South Rumaila Oil Field. This encompasses modeling of the 3D facies and petrophysical properties. Both are extremely necessary for accurately representing the geological process towards attaining

accurate reservoir evaluation, and a precise prediction of the future reservoir performance. Prior to petrophysical property modeling, the facies should be modeled to capture the real depositional environment. That step is essential because the reservoir properties are affected by the different rock types, which lead to distinct capillary pressure and relative permeability curves that directly affect the reservoir performance, field development, and economic evaluation (Kalla and White, 2007).

Indeed, many steps are required for reconstructing a full geostatistical reservoir model. The first step was to collect all the static well and reservoir data. The data has been gathered for 20 wells in the reservoir and includes well log interpretations of: Neutron Porosity, Shale Volume, and Water Saturation. Unfortunately, there was only one well with a core analysis report that had permeability, porosity, and Lithofacies for some intervals. This increases the uncertainty in data required for the geostatistical modeling.

The geostatistical modeling mainly includes multivariate and geospatial distribution of lithofacies, permeability, and porosity. For multivariate lithofacies analysis, several algorithms were compared to model and predict the discrete and posterior lithofacies distributions along the entire depths. These algorithms included Linear Discriminant Analysis, Boosting Logistic Regression, Boosting Logistic Regression (LogitBoost), Kernel Support Vector Machine, and Probabilistic Neural Networks (PNN). The most accurate model found was the Probabilistic Neural Networks, and was adopted for predicting the lithofacies for all wells. Later, Generalized Linear and Smooth Generalized Additive (sGAM) were adopted to predict the core permeability and porosity as a function of well log interpretations and PNN-predicted Lithofacies for the non-cored wells. Based on the cross-validation, the sGAM is selected for its precise permeability prediction with regard to the least variance between the measured and predicted measurements.

For geospatial modeling, multiple-point geostatistics (MPS) was conducted first for the 3D spatial lithofacies distribution for all the reservoir layers. Later, sequential Gaussian simulation was considered the petrophysical data distribution for each Lithofacies to create

distinct variogram fits in order to capture the most realistic geological heterogeneity model. Reconstructing the geological process cannot be perfectly implemented without analyzing the uncertainty assessment, especially when there is spatial complexity, anisotropy, and rock heterogeneity. Therefore, this process requires generating multiple equiprobable reservoir images (realizations) within the geostatistical modeling. Creating these realizations helps quantify the uncertainty through the reservoir simulation of the GAGD process performance.

After upscaling the full-resolution geostatistical model, the compositional reservoir model is built to evaluate and optimize the GAGD process performance. Sensitivity analysis was then conducted to investigate the reservoir properties effect on the process as a prior basis for history matching and optimization. In the optimization processes, the Latin Hypercube Sampling (LHS) and DECE design algorithms were adopted as efficient global search approaches that reduce computational time and achieve the optimality criteria by producing low-discrepancy and more uniform space-filling design of experiments (Jin, Chen, and Sudjianto, 2005). The optimization schemes include continuous and cycling CO<sub>2</sub>-GAGD injection modes followed by optimization under geological uncertainty and uncertainty quantification with regards to geological and production parameters. Further, The anisotropy ratio and horizontal reservoir permeability are likely to most affect the GAGD process performance according to the vertical fluid movement concepts in the GAGD process. The main steps for reservoir characterization, simulation, optimization and geological uncertainty quantification are outlined below:

1. Collect the available core and well log data for the wells and conduct the univariate and bivariate statistical analysis in order to estimate the missing values.
2. Implement Exploratory Data analysis to predict the vertical facies sequences by Probabilistic Neural Networks (PNN). Then, Smooth Generalized Additive Model (sGAM) is conducted to predict the core permeability in non-cored intervals and other wells in the reservoir.

3. Build the horizon model by subdividing the current reservoir zones into many layers in order to capture the geological reservoir reality.
4. Construct the structural gridding model to be considered for the spatial data distribution and the reservoir modeling.
5. Generate a detailed geostatistical lithofacies model that includes the 3D distribution of sand, shaly-sand, and shale by using multiple-point facies geostatistics.
6. Distribute the petrophysical properties spatially over all the layers given each type of Lithofacies through Sequential Gaussian Simulation (SGSIM).
7. Compositional reservoir simulation is implemented to evaluate the reservoir performance through the GAGD process.
8. Consider an efficient new statistical algorithm, Bayesian Model Averaging, to fit the experimental design model and to eliminate the non-influential factors on the GAGD process performance.
9. Optimize the GAGD process performance in immiscible CO<sub>2</sub> flooding mode through production control factors and Huff-and-Puff scenarios under geological uncertainty.
10. Assess the uncertainty in terms of geological and production factors.
11. Quantify the effect of heterogeneity and anisotropy on the GAGD process performance through simulation of all realizations.

These steps are shown predominantly in Figure 1.2.



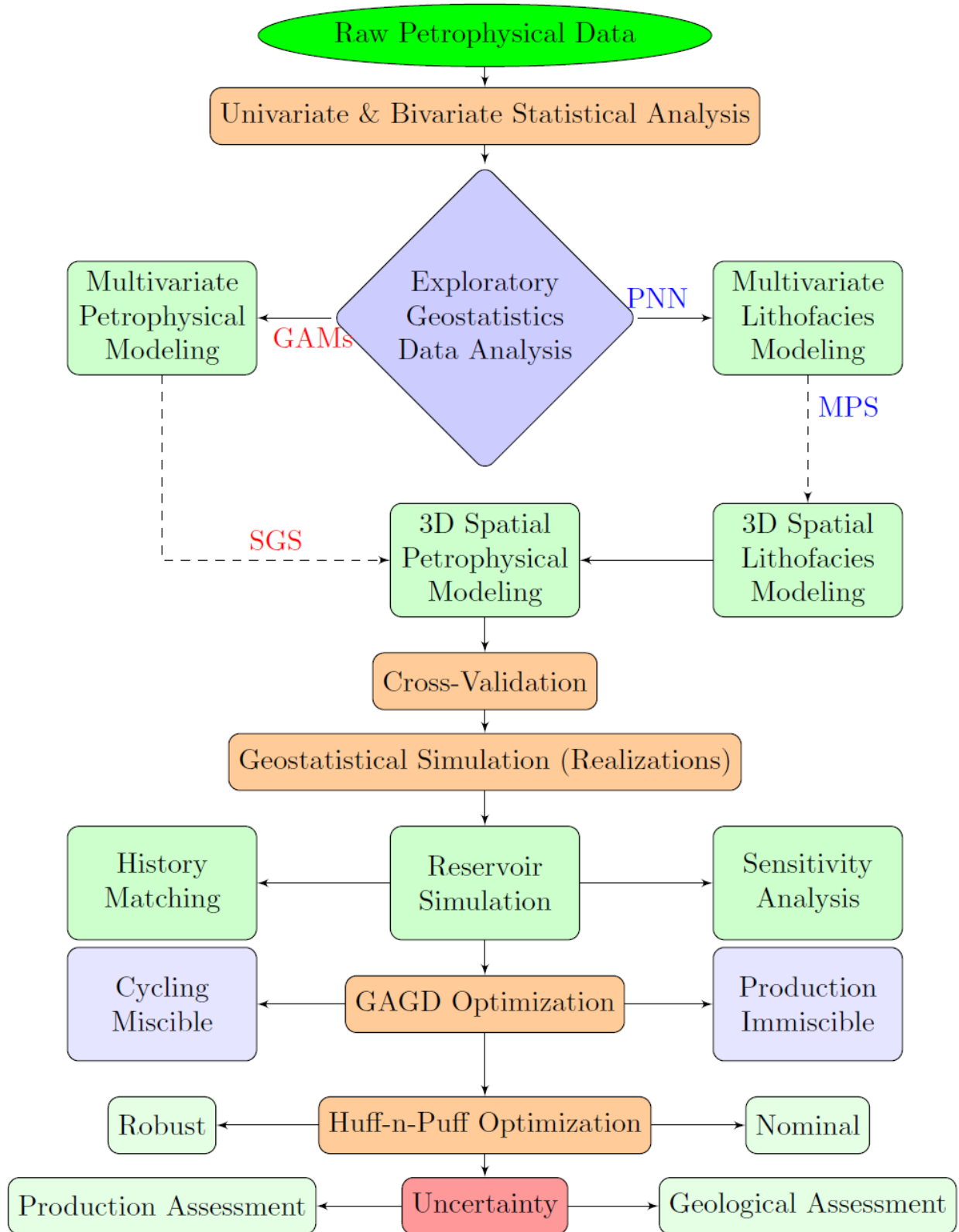


Figure 1.2: The Dissertation Workflow

## 1.4 Software Used

Two main software and one programming language were adopted in this dissertation: Petrel, CMG, and R. Petrel was used to build the geological model, CMG-WinProp for PVT modeling, CMG-GEM for compositional reservoir simulation of the GAGD process, CMG-CMOST for running the simulation jobs in sensitivity analysis, optimization, and uncertainty assessments. On the other hand, R as the most powerful statistical language was used for conducting the multivariate statistics and analyzing the results of sensitivity analysis, optimization, and uncertainty assessments, which are conducted by CMG-CMOST.

### 1.4.1 Petrel

Petrel is a software platform used in the exploration and production sector of the petroleum industry and it is developed and built by Schlumberger. It allows the user to interpret seismic data, well log analysis, perform well correlation, build reservoir models, visualize reservoir simulation results, calculate volumes, in addition to produce 2D and 3D maps. Petrel also has the ability to design development strategies to maximize reservoir exploitation.

For geological modeling, Petrel has specifically several tools for geostatistical reservoir characterization for facies and petrophysical properties modeling. In facies modeling, the user can build the lithofacies model through indicator kriging, pixel-based modeling, object-based modeling, and multiple-point geostatistics. On the other hand, kriging tools, sequential gaussian simulation, truncated gaussian simulation, and random field simulation can be conducted for petrophysical property distribution.

On the other hand, full reservoir simulation model can be constructed by Petrel including gridding system, upscaling, aquifer setting, history matching, as well as implementing uncertainty assessments. In this dissertation, Petrel was used only for building the 3D geological model of facies, porosity, and permeability through Sequential Indicator Simulation, Multiple-Point Geostatistics, and Sequential Gaussian Simulation, respectively.

### 1.4.2 CMG

The CMG is a Computer Modelling Group’s simulator that was employed to model the CO<sub>2</sub> flooding through the GAGD process in the main pay of South Rumaila oil field. CMG is a windows-based computer software for finite-difference reservoir flow simulation. Finite-difference solves the governing differential equations of fluid flow phenomena in a reservoir formation. The differential equations are derived from combining material balance, Darcy’s law, and mass conservation equations (Aziz and Settari, 1979).

In CMG, package GEM is an advanced equation of state compositional reservoir simulator (CMG, 2011). GEM was adopted for implementing the GAGD process simulation, sensitivity analysis, optimization, and uncertainty quantification.

In compositional reservoir simulation, the fluid flow in porous media is tracked in terms of the compositions of each existing phase. Based on the overall reservoir fluid compositions, mass fractions of each component in fluid phases are determined with the equation of state (EOS) based on phase equilibrium calculations. The compositions of each fluid phase and fluid phase properties such as viscosity and density are determined from the equilibrium calculations. While, the equation of state parameters are tuned to match phase behavior parameters with laboratory PVT data at surface (separator) and reservoir conditions. More specifically, Peng-Robinson EOS was employed for phase equilibrium calculations in the CMG software (CMG, 2011).

### 1.4.3 R Statistical Language

R is the most open-source software environment for statistical computing and graphics. It compiles and runs on a wide variety of UNIX platforms, Windows and MacOS. Several statistical tools can be conducted through R. In this dissertation, R was used for multivariate and spatial statistical analyses of petrophysical properties and rock facies. It was also adopted to analyze the results of sensitivity analysis, optimization, and uncertainty assessments.

More specifically, R was adopted for multivariate statistical analyses of facies and petrophysical properties. The algorithms that were implemented through R: linear discriminant analysis, kernel support vector machine, and probabilistic neural networks for facies classification along with generalized linear modeling and smooth generalized additive models for porosity and permeability. In addition, R was used for Bayesian Kriging Simulation of petrophysical properties. In advance simulation, R was again fully employed for analyzing the results of sensitivity analysis, optimization, and uncertainty assessments.

## **1.5 Review of Chapters**

Chapter 1 introduced the motivation of doing research about the GAGD process and showed the main objectives of the problem statement. Chapter 2 includes literature review of the current approaches and field description. Applied multivariate statistics was outlined in Chapter 3, and spatial Lithofacies and petrophysical modeling follows in Chapter 4. Chapter 5 encompasses the GAGD process reservoir simulation and history matching. Sensitivity Analysis of the GAGD process was included in Chapter 6, that also encompasses the new Bayesian selection of influential factors affecting the GAGD process. Chapter 7 presents both the production control and injection pressure optimizations of the GAGD process through immiscible and miscible modes, respectively. It also includes the Huff-and-Puff optimization under geological uncertainty. Chapter 8 addresses the uncertainty assessments for the geological and production parameters, as well as the effect of reservoir heterogeneity and anisotropy. Lastly, Chapter 9 discusses the main findings of the entire research and recommendations for future work. The workflow of dissertation structure is illustrated in Figure 1.3.

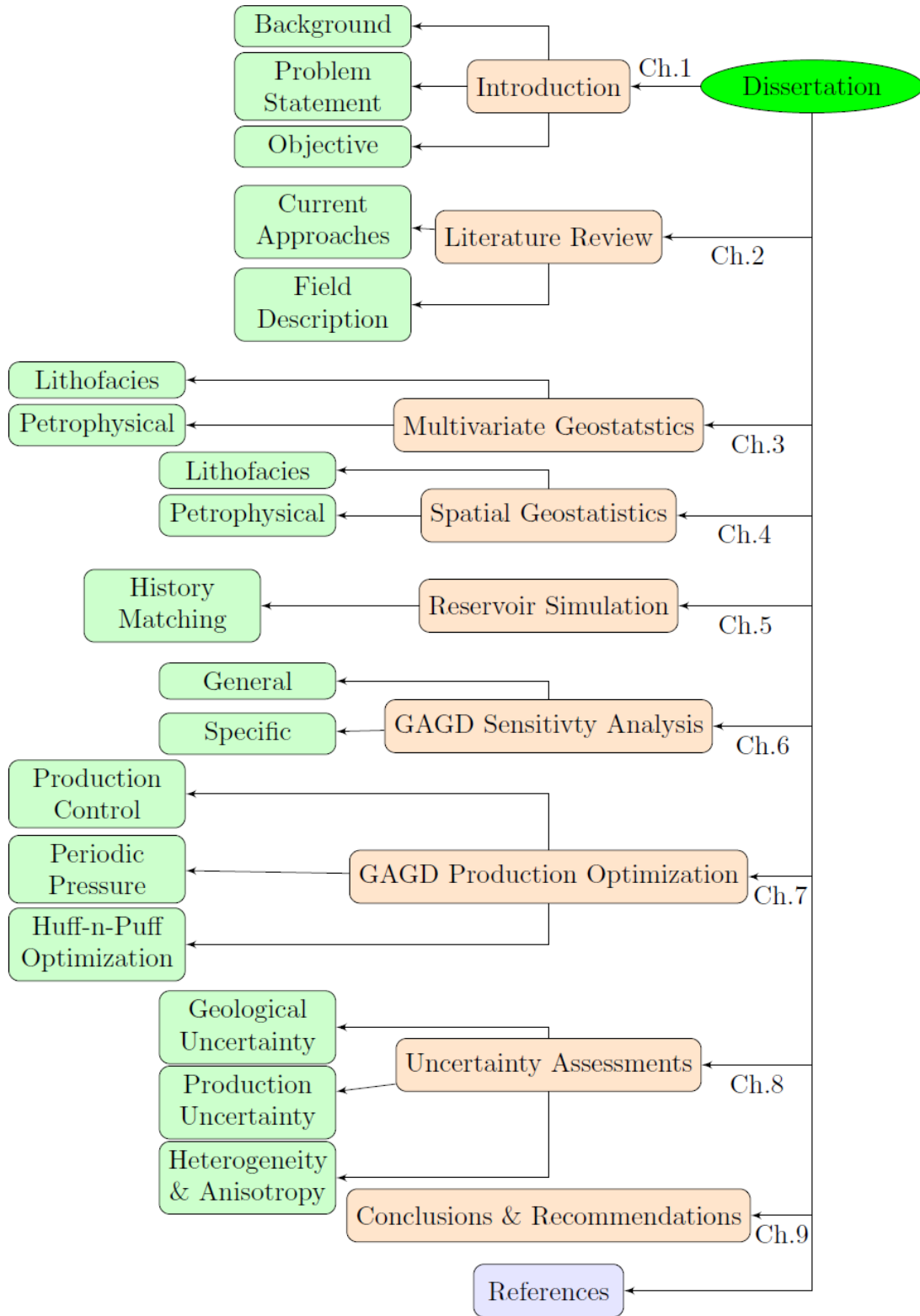


Figure 1.3: The Dissertation Structure

## 2. Literature Review

### 2.1 Applied Multivariate Geostatistics

The applied multivariate geostatistics include petrophysical properties estimations and lithofacies/electrofacies classification given well log data, core measurements, and DST data. Since well log data and core measurements come from different sources with various scales, it is required to efficiently combine both of them through statistical modeling. The most conventional way is regression analysis. However, it is not efficient for sparse data with high variance estimation of how far sets of numbers given each parameter are spread out (Al-Mudhafar and Mohamed, 2015). Many statistical learning algorithms have been studied in the literature that have addressed petrophysical property modeling.

Several studies have been done to estimate the Electrofacies sequences in the well log data. A combination of principal component analysis, cluster analysis, and multivariate statistical algorithms, such as discriminant analysis, are used to characterize the Electrofacies distribution of the well log records (Lee and Gupta, 1999). The most efficient cluster analysis methodologies are K-means, model-based classification, and the Ward Hierarchical Model (Lee and Gupta, 1999; Mathisen and Gupta, 2003; Taware et al., 2008; Nashawi and Malallah, 2009, Al-Mudhafar and Bondarenko 2015).

Some other studies consider the depositional environments, diagnostic overprints, and pore throat size distribution in the facies modeling. These studies also account for porosity and permeability in the facies classification. The facies characterization mainly depends on the vertical changes in depositional sequences. The discriminant analysis has been widely

adopted to predict the rock types in non-cored intervals (Al-Mudhafer, 2014, Al-Mudhafar, 2015). Support Vector Machine has also been used for multivariate discrete and posterior lithofacies prediction (Al-Mudhafar, 2015). Additionally, facies could be characterized in uncored wells based on the classification tree analysis methodology, which provides flexible procedures to determine the most influential factors affecting the response (Perez and Gupta, 2005).

For permeability modeling and prediction, many different algorithms have been adopted to predict the core permeability given core measurements and/or well log records. These algorithms include multiple linear regressions (Silva et al., 2002), generalized linear modeling (Al-Mudhafer, 2014), generalized additive modeling (Lee and Gupta, 1999; Al-Mudhafar and Bondarenko, 2015), neural networks (Lee and Gupta, 1999), Least Angle Regression (Al-Mudhafar and Rostami, 2014), and Multiple Addaptive Regression Splines (Al-Mudhafar and Al-Khazraji, 2016). Bayesian Model Averaging has also been used as a stochastic algorithm to generate multiple models with different sets of coefficients in permeability modeling (Al-Mudhafar, 2015). Neural networks are being used in many studies for applicability to construct a nonlinear relationship between the petrophysical properties (Lee and Datta-Gupta, 1999; Nashawi and Malallah, 2009). Other procedures of artificial neural networks consider characterization of hydraulic flow units and flow zone indicators prior to permeability estimation (Lacentre et al., 2003). Flow zone indicator (FZI) is created based on Carmen-Kozeny's equation that relates permeability and porosity. Hydraulic flow units cannot be efficiently characterized unless there is enough core and well log data. The resulted permeability has a specific link to porosity at each characterized flow zone unit (Lacentre et al., 2003; Perez and Gupta, 2005; Teh and Willhite, 2012; Yerramilli et al., 2013). Some other studies used the fuzzy logic algorithm for well log-based permeability modeling. Selection of the influential well log data on permeability depends on the correlation coefficient to prepare the final suitable input data for the fuzzy training procedure. This model proved its efficiency to capture the high level true nonlinear rela-

tionship between permeability and well log data, which leads to accurate prediction with excellent matching with the observed measurements (Abdulraheem et al., 2007; Nashawi and Malallah, 2010).

These methods were used for different lithology formations such as sandstone, limestone, and carbonate (Lee and Gupta, 1999; Fernando et al., 2002; Lacentre et al., 2003; Mathisen and gupta, 2003; Perez and Gupta, 2005; Nashawi and Malallah, 2010 Teh and Willhite, 2012; Yerramilli et al., 2013; Al-Mudhafer and Gebrai, 2014, Al-Mudhafar and Al Isawi, 2015).

## **2.2 Geostatistical Lithofacies Modeling**

The type and architecture of the geological facies and the related petrophysical data distribution are crucial issues that have direct influence on the reservoir heterogeneities, fluid flow performance prediction, field development, and economic evaluation (White and Royer, 2003). Consequently, the process of integrating all the available geological information into the numerical reservoir simulation in terms of 3D facies distribution is called facies modeling. For accurate facies distribution, it is necessary to efficiently integrate the different scales of data to capture the reservoir heterogeneity and transfer it into the reservoir model. The most recent methods that have been used to quantify the facies and petrophysical properties are the deterministic (estimation) and stochastic (simulation) approaches (Journel, 1989; Journel, 1990; Journel and Alabert, 1990; Journel and Gomez-Hernandez, 1993; Overeem, 2008).

There are several methods for facies modeling such as variogram-based methods, object-based modeling, and multiple-point geostatistics (Zhang, 2008, Liu et al., 2004). The most commonly performed geostatistical models are pixel-based methods. Pixel-based models are used for stochastic modeling of either discrete distribution variables such as facies or continuous distribution variables such as porosity and permeability. More specifically, pixel-based models encompass sequential indicator simulation for facies and sequential Gaussian simulation for petrophysical properties.



### 2.2.1 Pixel-Based Modeling

The pixel-based model is used to build the categorical image, pixel by pixel, considering nonparametric condition distribution. The pixel-based model, or sometimes referred to as two-point statistics, depends on the variogram to address and model the variation between any two spatial points from the available data (Caers and Zhang, 2004; Zhang, 2008; Liu et al., 2004; Deutsch and Journel, 1998). Although it is simple and requires few parameters, the pixel (variogram)-based modeling has some limitations regarding the inability to capture complex geological structures such as channel shapes, thickness, and sinuosity. For instance, this inability would include scenarios where there are many different geological formations that have the same variogram, but with various architectures (Caers and Zhang, 2004).

Sequential Indicator Simulation (SISIM) is the most common pixel-based algorithm that has been widely accepted in many different reservoir characterization studies. SISIM has been successfully used for lithofacies modeling to characterize flow barriers and/or flow paths and fractures in sandstone formations (Journel and Alabert, 1990), complex turbiditic reservoir (Alabert and Massonnat, 1990), and sand-shale reservoirs (Massonnat, Alabert, and Giudicelli, 1992). The different data sources and scales such as well data, seismic, and geological interpretation can be integrated to build the facies model through SISIM and Co-SISIM (Journel and Alabert, 1990).

More specifically, Sequential Indicator Simulation is used to build a conditional probability density function of nonparametric properties, such as rock types, considering binary simulation (Deutsch and Journel, 1998; Caers, 2000). Once the continuous variable is transformed into a number of indicator variables, one for each class, then each of the indicator variables may be spatially modeled using any of the variogram or covariance functions. In SISIM, the facies data are encoded into binary values (0,1) controlled by giving threshold values. Then, spatial distribution is characterized by the indicator variograms in order to generate alternative equiprobable stochastic reservoir images. The binary encoding could

also be adopted for the main facies characterization such as sand and shale (Journel and Gomez-Hernandez, 1993). Moreover, the facies can be encoded as binary variables given a specific location to represent different rock types such as one for sand channels and zero for lobe in a complex turbiditic reservoir (Alabert and Massonnat, 1990). Additionally, SISIM can predict the facies distribution in inter-well regions through the log signatures to characterize major log associations in uncored wells and provide conditional data for the stochastic facies distribution. In addition, to capture the marine to fluvial transition, the SISIM has been modified to incorporate a vertical trend in the mean proportion of each of the major facies associations (Begg et al., 1996). Sequential Indicator Simulation has also been widely described in many other studies (Alabert and Modot, 1992; Srivastava, 1994; Goovaerts, 1997; Seifert and Jensen, 1999).

### **2.2.2 Object-Based Modeling**

Unlike the variogram-based geostatistics method that has some limitations with regard to capturing the reality of complex subsurface heterogeneous formations (Caers and Zhang, 2004), Object-based modeling (OBM) was first introduced by Haldorsen and Lake (1984), Haldosen and Chang (1986) and Stoyan et al. (1987) for more reliable spatial distributions of geological environments. Much more data is required to efficiently characterize the reservoir through OBM than when using pixel-based modeling. In addition to well logging, core measurements, depth to horizons, and lithofacies information, seismic data is also an important source for characterizing geological horizons and large faults, especially in high-resolution records. Furthermore, well testing and production history data might be useful for limited interpretations within the object-based modeling procedure, such as measuring the distance to a barrier. Additionally, outcrops may be beneficial for complex characterizations such as channel width, distance between channels, fault properties, and smoothness of horizons (Damsleth, 1994).

The main characteristics of the object-based technique is to provoke multiple training images, which consider the repetition and interactions between patterns, to optimally cap-

ture the geological features realistically in well-defined geometry such as sedimentary units or faults (Alabert and Modot, 1992). Deutsch and Journel (1998) have employed many parameters that control the 3D sand bodies such as: channel geometry and dimensions, spatial distribution of channels, interaction between channels, channel sinuosity, channel direction, and grouping the different channels in channel belts. Object-based methods have been successfully applied in many case studies, especially to capture large-scale geological heterogeneities. Examples include: the North Sea reservoir (Damsleth et. al., 1992; Hove et. al., 1992), the North Sea coastal-deltaic reservoir (MacDonald et.al, 1991), Ness formation of a field in North Sea (Kaas et. al, 1994), and Fluvial reservoirs (Deutsch and Wang, 1996).

### **2.2.3 Multiple-Point Geostatistics**

The multiple-point geostatistics (MPS) has been introduced to incorporate both geologic and 3D seismic data for reproduction of curvilinear facies structures in order to overcome the limitations of variogram-based and object-based stochastic modeling (Strebelle and Journel, 2001). The MPS procedure is mainly based on utilizing the training images as a numerical geological model that contains three-dimensional information about a geological architecture (Liu et. al., 2005). Outcrops, analogues, well log interpretations, core analysis, and seismic records are required for training images that are used for facies architecture reconstruction (Liu et al., 2004). The tools used to create training images are: sequence stratigraphy studies, object-based algorithms, user-defined images, and the process-based models, which can be considered as a prior model to constrain statistics, such as facies transition, architecture, etc, and to construct the facies model (Zhang, 2008). The training images can be either for discrete variables (facies) or continuous variables such as porosity and permeability (Zhang, 2008).

In a particular, the training images should be created in 3D to capture all the facies patterns and reservoir heterogeneity (Caers and Zhang, 2004). Pixel-based models can then be utilized to include the training images into numerical modeling through the

single normal equation simulation (SNESIM) algorithm (Isaaks, 1990; Gomez and Srivastava, 1990; Strebel, 2003). Multiple-point geostatistics has been proved a more powerful tool for incorporating the geological information and capturing the facies architecture than variogram-based and object-based geostatistical modeling (Liu and Journel, 2005). In addition to the reliability and stationarity of training images, the MPS algorithm requires a lot of data to efficiently capture the reservoir reality. Alternative to SNESIM, there are other algorithm for MPS facies modeling such as FILTERSIM, DIRECT SAMPLING, GROWTHSIM, PATCHSIM ETC (Zhang, Bombarde, Strebel, and Oatney, 2006; Lallier, Vignau, and Kombrink, 2013).

### **2.3 Geostatistical Petrophysical Modeling**

Geostatistical models are used for spatial simulation of petrophysical properties to build a 3D reservoir model (Gringarten and Deustch, 1999). Since the spatial petrophysical property distribution has a direct affect on the fluid flow through porous media, it should be efficiently implemented through the variogram analysis and kriging algorithms. Specifically, the variogram allows understanding the geometry and continuity of the reservoir (Gringarten and Deustch, 1999), moreover, kriging represents the Best Linear Unbiased Estimator (BLUE). Initially, kriging was introduced by Daniel G. Krige to estimate the spatial heterogeneity and then it was developed by Georges Matheron (David, 1986; Rivoirard, 2005). Kriging uses the spatial covariance model between the known and unknown data to estimate the missing variable values at new locations based on their neighbors (Deutsch and Journel, 1998). The kriging of two variables at one time is called co-kriging and can be applied, for instance, using the relationship between the porosity and permeability to estimate the primary parameter as a function of the secondary one (Pyrzcz and Deutsch, 2014). When there is full secondary variable distributed such as acoustic impedance from the seismic attributes, porosity in some cases can be estimated as a primary variable within the collocated co-kriging (Xu et al., 1992). The collocated co-kriging does not require the covariance function of the secondary data. However, it requires cross-covariance function

calculations. On the other hand, the covariance of the secondary data is needed in Markov model II of collocated co-kriging (Doyen et al., 1996).

On the other hand, the stochastic simulation has been successfully adopted for generating multiple equiprobable realizations of petrophysical property modeling rather than just one estimation by kriging interpolation. The main algorithm for continuous property distribution is the sequential Gaussian simulation. The procedure of conditional simulation includes transforming the spatial data into a normal distribution, then compute and fit variograms, and finally it generates properties by sequentially sampling from the conditional distributions by representing the variable at each grid as a random following Gaussian distribution (Cai and Hicks, 1999; Deutsch and Journel, 1998; Goovearts, 1997; Pyrcz and Deutsch, 2014).

Additionally, the seismic attribute maps can be incorporated through sequential Gaussian simulation with block kriging to build 3D reservoir property models. This method considers the seismic map as a soft estimate of average reservoir property over all layers. As a result, it was considered to be secondary data for controlling the primary variables at each of the flow zones (Behrens et.al, 1998). The main advantage of block kriging-sequential Gaussian simulation is to integrate the spatial heterogeneity obtained from seismic attribute maps with well log data variation.

Furthermore, the simulated annealing algorithm has been hybridized with sequential Gaussian simulation to be pointed out as Annealing Cosimulation (ACS) in order to integrate different types of data. For instance, the acquired core permeability measurements along with its relevant statistical analysis can be integrated together with permeability measurements from well test data (Deutsch and Cockerham, 1994). In the annealing implementation in Deutsch and Cokerham, it is pure annealing where data integration is in the form of minimization of a global objective function. ACS adopts an objective function to be a typical example of difference between variogram of the realization and original variogram that has been created from the hard data.

Since the petrophysical properties have many different sources such as well log, core, well test, and seismic records, it definitely has a degree of uncertainty. Consequently, Bayesian kriging has been used to capture some prior knowledge about the observation to be considered as a qualified guess in the estimation procedure (Omre, 1987). The qualified guess gives the opportunity to decrease the degree of uncertainty in the estimation, and obtain more realistic observations under the assessment of uncertainty (Omre and Halvorsen, 1989). The uncertainty is quantified as a posterior distribution to avoid unrealistic small regions within the observation spatial distribution in order to attain optimal linear interpolation. This process could be conducted by multivariate statistics (Le and Zidek, 1992). Bayesian kriging has been applied in a sandstone reservoir to test its effectiveness relative to other conventional kriging algorithms (Al-Mudhafar and da Silva, 2015).

## 2.4 Geological Uncertainty

Since there is limited geological information required to build efficient reservoir models, there is the inability to precisely evaluate the reservoir performance. That increases the uncertainty associated with forecasting in the future reservoir performance and negatively impacts the economic returns (Zhang, 2003). The limited knowledge comes from lack of observations, distinct scales of petrophysical data, measurement error, limited understanding of physical problems, and non-linear data behavior. Therefore, the uncertainty assessment is necessary to construct a solid basis in the reservoir management for development and optimization in order to improve the decision making (Cruz, 2000).

The most common methods for uncertainty assessment encompass: Monte Carlo simulation, experimental design-response surface methodology, multiple realization tree, relative variation method, and Bayesian methods. An empirical response surface model (RSM) has been developed with design of experiments (DoE) to diagnose and investigate several geologic factors (White et al., 2001). RSM is a collection of mathematical and statistical techniques used for modeling and analysis of physical problems, in which a response of interest is influenced by several variables and the objective is to optimize this response.

The response surface model is not only used only in identifying the most sensitive factors that affect a response, but also to obtain the most likely scenario that achieves the optimal response through a process. RSM also identifies the possibility of estimating interactions between the factors (White and Royer, 2003). IN RSM, a surface is fitted through the responses corresponding to a few sampled set of parameters. This surface could be a polynomial function or established using ordinary kriging and the surface is subsequently used for uncertainty quantification. Comparative applications of polynomial and kriged response surface models have been conducted considering Orthogonal Arrays sampling, Latin Hypercube Sampling, and Hammersley Sequence design as an efficient approach for uncertainty quantification (Kalla and White, 2007).

Many other uncertainty quantification studies have been carried out through Experimental Design concepts with different methods of handling either variable sampling or sub-selections. Some of these methods represent coupling of streamline reservoir simulation and the Neighborhood Approximation Algorithm along with the conventional reservoir model to assemble the model parameters separately from less likely misfit regions of parameter space (Subbey et al., 2003). Another approach integrates the Neighborhood Approximation Algorithm with Markov Chain Monte Carlo as a stochastic sampling method to surmount the poor convergence (Okano et. al., 2006). There are many lessons learned from real oil and gas fields using experimental design and response surface models in uncertainty quantification (Amudo et al., 2009). They have shown how efficient using the design of experiments method to minimize the reservoir simulation runs during the formulation of the proxy models. The experimental design has proved efficient in conducting uncertainty framing, screening parameters, Bayesian updating of parameters, and risk analysis (Amudo et al., 2009).

Many statistical validation tools such as ANOVA table and partial t-test are required to validate the experimental design approach and assess the factors uncertainty. Analysis of variance (ANOVA) is used in modeling the relationship between response and uncertainty

factors. Moreover, it is used to investigate the accuracy of the regression model and likewise to measure the effect of each factor on the response by computing the variance and mean square error (White and Royer, 2003). Due to computer limitations, the uncertainty assessment of 3D Lithofacies, porosity, and permeability distributions should be accomplished considering the limited number of stochastic reservoir images (realizations). Many techniques have been adopted for ranking the geostatistical realizations to obtain the less-likely (pessimistic), median, and more-likely (optimistic) levels within a large number of realizations that capture the various reservoir heterogeneities. The main concepts of ranking is to calculate a reservoir flow response, which represents the objective function for comparison between the realizations. Examples of the objective functions include: hydrocarbon production over a cross section (Ballin et al., 1993), the minimized error between calculated and observed field production through history matching (Kalogerakis, 1994; Oliver et al., 1996), cumulative field hydrocarbon production, net present value (NPV), and the production response through infill drilling (Deutsch and Journal, 1998; Hegstad and Henning, 2001; Landa and Horne, 1997; Xian-Huan et al., 1997), and streamline reservoir simulation (Ates et. al., 2003).

## **2.5 Gas Assisted Gravity Drainage (GAGD) Process**

### **2.5.1 GAGD Process Physical Model**

The main GAGD process, which was introduced by Rao et al. in 2004, encompasses placing a horizontal production well at the bottom of the payzone exactly above the oil-water contact. Then, gas is injected through existing vertical injectors to accumulate at the top of the formation. The GAGD process has been put forward for application into two applicable modes either immiscible or miscible gas flooding. In immiscible mode, the gas accumulates at the top of the reservoir forming a gas cap. The higher density fluids, i.e. oils, drain down towards the horizontal producer (Rao et al., 2004). The authors have constructed a scaled experimental physical model in order to study the GAGD process effectiveness and to understand the mutual action and reaction of capillary, gravitational,



and viscous forces. The scaled physical model has been utilized to investigate the most influential factors affecting the GAGD process such as: miscibility/immiscibility flooding, wettability, heterogeneity, and etc. Additionally, the scaled physical model has demonstrated the feasibility of the GAGD process in comparison with Continuous Gas Injection (CGI) and Water-Alternative Gas (WAG) in both secondary and tertiary recovery. From the scaled physical model, the incremental recovery through miscible WAG process is higher than the miscible CGI in tertiary mode. However, miscible CGI has better incremental recovery than miscible WAG in secondary mode. Moreover, the miscible WAG has more distinguished incremental recovery than its immiscible case in both secondary and tertiary modes. On the other hand, the GAGD process has demonstrated a better incremental recovery through immiscible mode than both immiscible WAG and CGI processes (Rao et al., 2004).

The effective comparison between WAG and GAGD processes has been efficiently investigated through dimensional analysis, which was employed on nine gravity stable along with eight miscible and immiscible WAG field applications in the world (Kulkarni and Rao, 2006). The petrophysical dimensionless numbers include Bond(NB), Dombrowski-Brownell(ND), and Grattoni et al.'s (N), with the operational numbers of Capillary (NC) and Gravity (NG), Capillary(NC) and Bond (NB) have been calculated for the aforementioned field projects. These groups' numbers have been coupled with Microscopic Bond Number to characterize the flow regimes in addition to the controlling forces of gravity, viscous, and capillary in both field and laboratory displacements. On the basis of these dimensionless numbers, the gravity stable gas injection has demonstrated much better production rates and recovery factors than both miscible and immiscible WAG processes (Kulkarni and Rao, 2006).

A new gravity-drainage mechanism has been proposed and validated with two analytical and empirically free gravity drainage models. This approach captures the multiphase mechanisms and fluid dynamics of the gravity drainage forces through the GAGD pro-

cess that could not be characterized by Buckley-Leverett and the gravity-drainage theories (Kulkarni and Rao, 2006). More specifically, the analytical model, which was introduced by Richardson and Blackwell, was employed to predict oil recoveries for 1D and 2D laboratory GAGD floods within 6.4% error. In addition, the Richardson and Blackwell model has been conducted on different GAGD experiments of secondary mode such as gravity and non-gravity stable displacement through immiscible and miscible GAGD flood. Their model has been validated against the Hawkins Dexter field data and the ultimate oil recovery was under-predicted by 5.2% IOIP. On the other hand, empirical model of Li and Horne was utilized to over-predict the GAGD oil recoveries and predict oil production rates by incorporating capillary pressure data to improve gravity drainage recovery predictions. The predicted oil pressure through that empirical model was tested against experimental, numerical and field data and it was proved to be more suitable than others (Li and Horne, 2003). Overcoming the limitation of the Richardson and Blackwell model, it has been investigated to conclude that the empirical model of Li and Horne has led to over-prediction of oil production rates through 1D GAGD core floods and 2D physical models in only free gravity drainage floods.

A physical model of two parallel glass plates confining a sand-pack has been constructed for implementing envisioned experiments of the Gas Assisted Gravity Drainage process. The purpose was to identify the optimal operational factors of the GAGD process with regard to the feasibility to improve oil recovery in secondary and tertiary modes (Mahmoud and Rao, 2007). The oil recovery in immiscible secondary has been investigated to be 83% IOIP. However, oil recovery in immiscible tertiary mode was 54% ROIP. Moreover, the microscopic sweep efficiency has reached up to 100% in miscible secondary and tertiary modes. Furthermore, the free gravity drainage and the hydrostatic head of fluid in place represent the dominant mechanisms prior to the gas breakthrough.

### 2.5.2 GAGD Field-Scale Application

The  $CO_2$  GAGD process has been applied in a North Louisiana field to achieve the optimal field development prediction performance through an economical analysis (Paidin, Mwangi, and Rao, 2010). Optimizing the net present value (NPV) and rate of return has been conducted through coupling the compositional reservoir simulator with the economic model considering the approximate oil price, capital (CAPEX), and operational (OPEX) expenditures for that field. The operational decision factors that were investigated for their possible influential role on the flow response are the following: number of injector/producer well pairs, well pair locations, location of the horizontal producer above the oil-water contact,  $CO_2$  injection rates, production rates, and the lag time between the start of injection until production. In conclusion, the gas injection rate has a direct effect on the GAGD process performance as a recovery factor increases with high injection rates. Since the cost per foot for drilling horizontal wells is significantly low, the GAGD process performance has proven its efficiency to attain high economic values of Net Present Value (NPV). The most influential economic factor affecting the NPV is oil price. However, the capital expenditure (CAPEX) and other factors have negligible effects on the economic response. Thus, the higher incremental of oil recovery through the GAGD process has led to a higher NPV.

## 2.6 Field Description

### 2.6.1 Introduction

The giant South Rumaila oil field was discovered by Basrah Petroleum Company in October 1953. The Rumaila oil field is located in southern Iraq about 50 km west of Basrah city and about 30 km to the west of the Zubair oil field (Al-Ansari, 1993). The Rumaila field is about 100 km long with a depth to more than 3 km below sea level and ranges between 12 to 14 km width. Dip angles on the flanks do not exceed  $3^\circ$ , whereas in the crest is about  $1^\circ$ . Faults are not indicated in all layers of the Rumaila field (Al-Ansari, 1993). The geographical location of Rumaila field with the current well locations is shown in Figure 2.1.

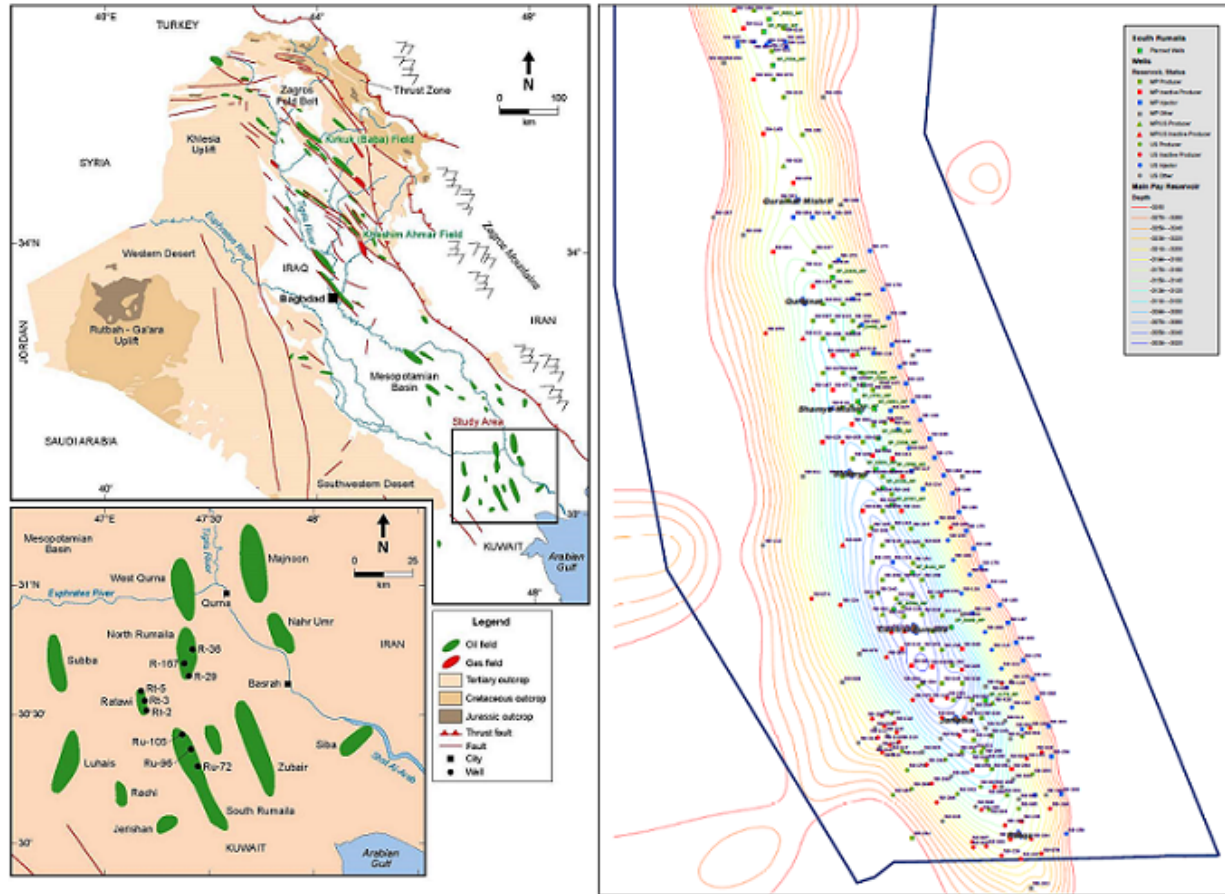


Figure 2.1: Geographical Location and Current well Locations-Rumaila Oil Field (Al-Ameri et al., 2009)

### 2.6.2 Geological Description

The South Rumaila oil field is composed of many oil-producing reservoirs. Zubair is one of the oil reservoirs that is represented by the Late Berriasian-Albian cycle and its sediments, which belongs to the Lower Cretaceous age. The Zubair formation is rich in organic deposition and accumulation of sedimentary matter (Al-Obaidi, 2010). The thickness of the Zubair formation ranges between 280-400 m with levels increasing towards north-east end of the field (Al-Obaidi, 2009). Based on sand to shale ratio, the Zubair formation encompasses five members. These members named from top to bottom are as follows: upper shale member, upper sandstone member, middle shale member, lower sand member, and lower shale member. The upper sandstone member of the Zubair formation

is the main pay zone of South Rumaila Oil Field (Mohammed et al., 2010). The main pay comprises of five dominated sandstone units, separated by two shale units (C and K). The shale units act as barriers impeding vertical migration of the reservoir fluids, however in some areas they are not as prominent. The reservoir and shale units have been denoted from top to bottom as AB, C, DJ1, DJ2, K, LN1, and LN2 (Mohammed et. al., 2010). The geological column for the entire reservoir's zones in the Rumaila oil field is shown below in Figure 2.2.

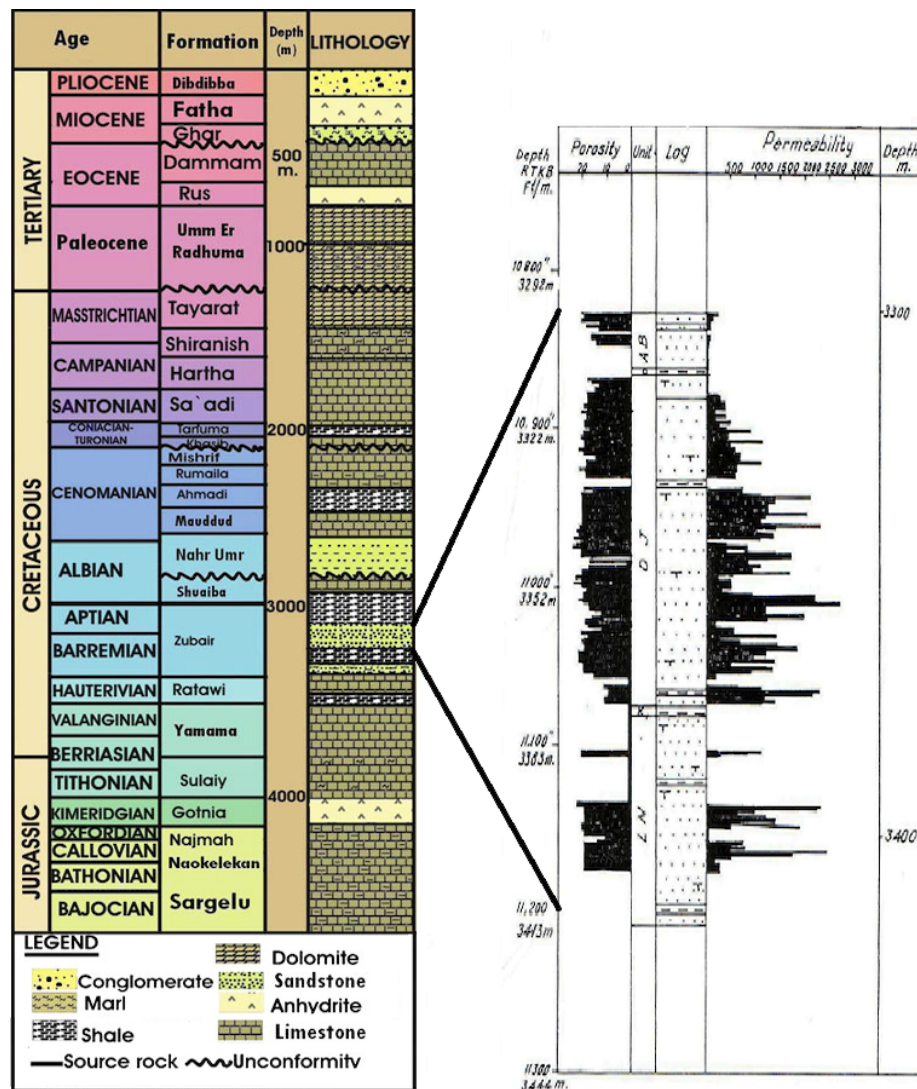


Figure 2.2: Geological Lithology Column of the Formations in South Rumaila Oil Field (Mohammed et al, 2010)

### 2.6.3 Depositional Environment

The Zubair formation is part of the lower Cretaceous sequence age. It is located between two dolomite-limestone and shale formations that are named Shuaiba and Ratawi, respectively. The Zubair formation encompasses mainly sandstone with some inter-bedded shale zones as mentioned previously (Al-Muhailan et al., 2013). Figure 2.3 shows the stratigraphic column with vertical lithology description and well log indications for all the zones in the South Rumaila oil field.

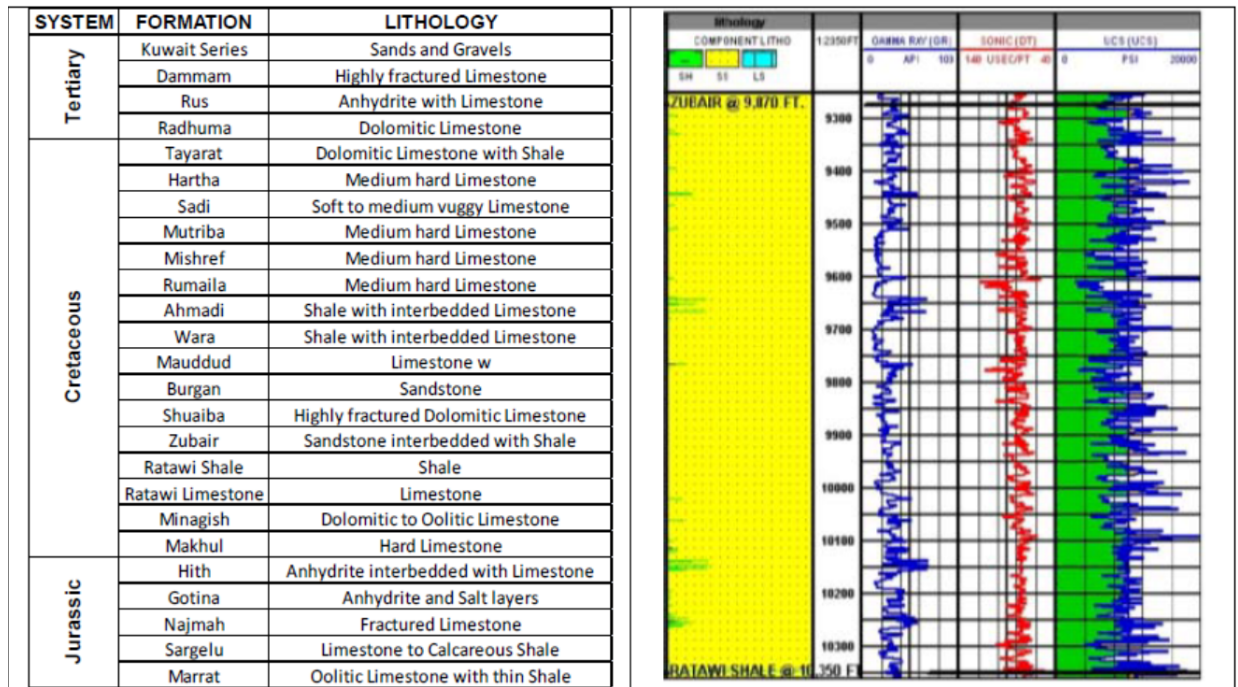


Figure 2.3: Stratigraphic column with lithology (Al-Muhailan et al., 2013)

Overall, the Zubair formation was deposited through deltaic, estuarine and fluvial environments (Harris et al., 2012). The upper shale member, which is located above upper sandstone member, was deposited in a wave/shoreface and offshore-dominated environment resulting in sand channels that extend East-West across the anticline (Al-Muhailan et al., 2013). The upper and middle sandstone members were deposited in a tidal/estuarine environment in which the sand channels are stacked and continuous everywhere across the formation. The lower sandstone member was deposited in a fluvial/mouthbar dominated

environment. The analyses of core samples have indicated coarsening upward black clay-stones at the base of lower part of the upper sandstone member overlain by bioturbated very fine grained sandstones, which in turn are overlain by fine-grained, trough cross-bedded sandstones. The same events took place at the middle sandstone member (Harris et al., 2012). The core sedimentology, palynofacies observations, and biostratigraphic analysis indicate both marine and terrestrial microflora confirm a marginal marine gross depositional environment. On the other hand, the Palynology studies have proven the succession of the very fine and fine grain sandstones (Kitching et al., 2013).

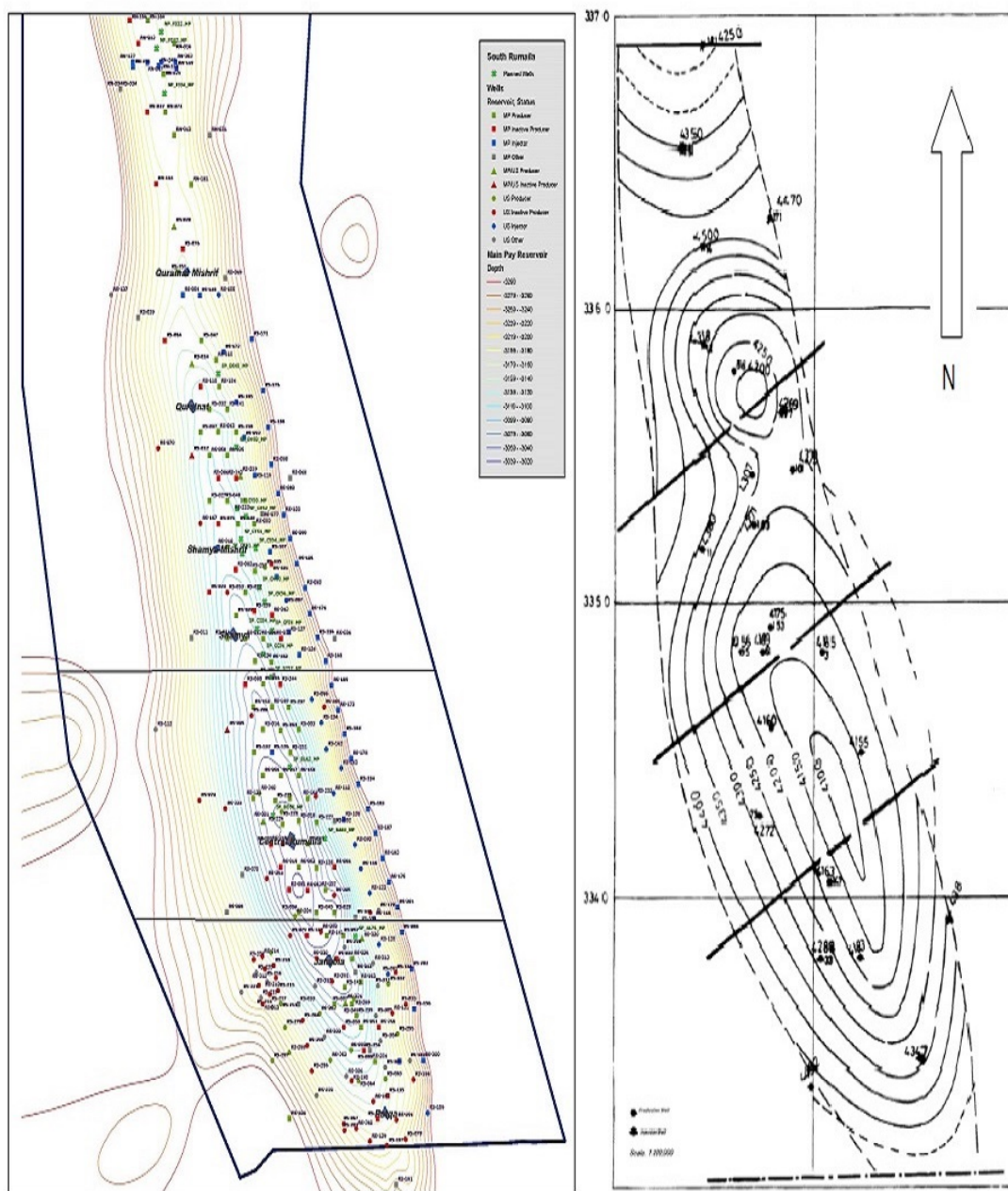
Another study stated that the main pay was a fluvially dominated, sand-rich deltaic environment (Well et al., 2013). The study likewise highlights the phased advance and retreat of a river-dominated and tidally influenced delta system. Additionally, the observation of cyclic bundles of the thin forest shale laminae confirms the influence of tides. Marine flooding surfaces were also observed and usually succeeded by prodelta shales that affect the vertical connectivity in the formation. Nevertheless, the majority of the effective reservoir rocks are found as high quality in lateral and vertical amalgamates. Excluding the shale distribution, the entire sand distribution tends to be heterogeneous through rapid lateral and vertical variations (Al Naqib, 1985).

In the southern part of the field, core analysis indicates fine to medium grained cross-bedded sandstones, which were mainly deposited in lower delta plain distributed channels. However, the grain sizes tend to decrease in mixed sand and shale deposits that were preserved in the northern area of the field (Well et al., 2013).

#### **2.6.4 Study Sector Description**

The South Rumaila field is divided into four sectors. Listed from north to south, the sectors are: Qurainat, Shamiya, Rumaila, and Janubia. The area investigated focused on the Rumaila sector and small parts from Shamiya and Janubia sectors. The choice of this area was made based on the availability of data and the ability to represent the largest part, where the production and injection operations are carried out, shown in Figure 2.4.







the boundary of the main pay reservoir also existed for more than two decades. In the late 1970s, natural depletion and water drive were the only production mechanisms (Al-Mudhafer, 2013). Later, 20 injection wells were drilled at the east flank to maintain this huge aquifer support from west flank, which accumulates up to 20 times the influx from east one (Kabir, 2004; Al-Mudhafer et. al, 2010). The production of some layers were ceased because the high water cut values exceeded 98%. The 20 injectors were primarily arranged in two parallel downdip rows. Specifically, the inner row was completed only in the bottom two layers, while the others were completed only in the second two layer intervals. By the year 2004, the cumulative water injection was approximately 1.1 billion barrels. The injection rates have varied widely with a maximum of nearly 426,000 BPD for two months in 1988. Artificial lift has been recently installed in the main pay wells in order to handle the wells incapable of flowing to the surface after the water cuts reach approximately 80%. Furthermore, the estimated original oil in place (OOIP) for the main pay is 19.5 billion barrels and for the sector is around 6.123 billions barrel. Moreover, the approximate current recovery factor is 55%. The peak oil production was 1.35 MMBPD in May 1979. The current oil production in July 2013 was approximately 1,250,000 MMBPD. Figure 2.5 shows the production history for the field.

### **2.6.6 Boundary conditions treatment**

There are two types of boundary conditions encountered in the main pay reservoir, a no-flow boundary and an aquifer. The northern and southern most boundaries are assumed to be a no-flow region. This assumption was considered realistic because the balanced production and injection rates were adopted in the reservoir. Furthermore, the isobaric lines crossing the northern streamlines is perpendicular to these boundaries. Thus, the direction of flow is parallel to the northern and southern boundaries, as shown in Figure 2.4. The flow boundaries at the east and west flanks represents the natural water drive due to the infinite aquifer (Al-Mudhafer et.al., 2010).

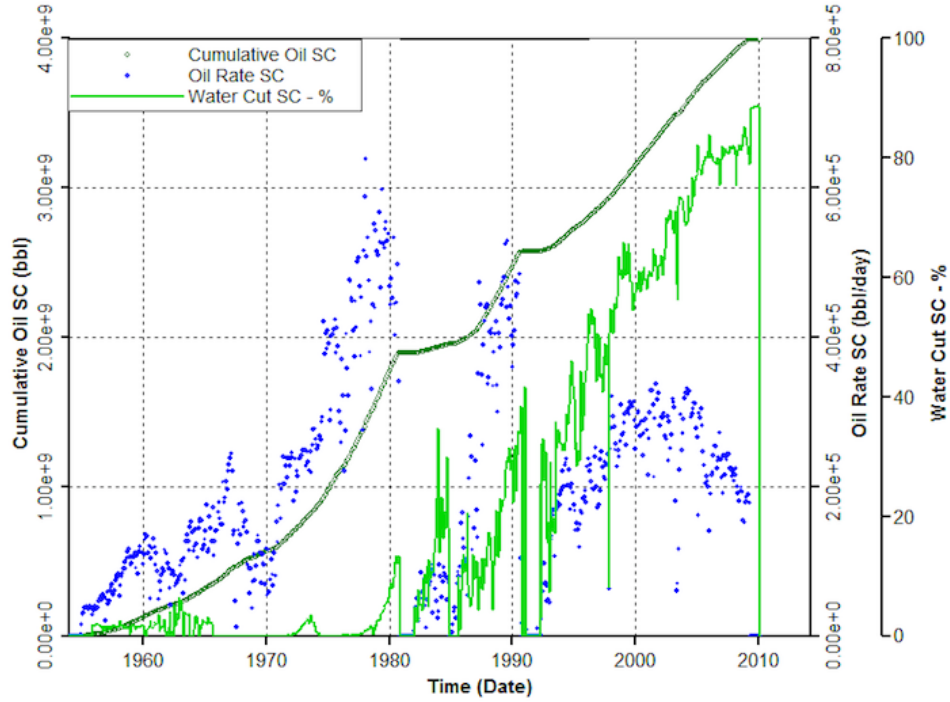


Figure 2.5: Field Production History of South Rumaila Oil Field

### 2.6.7 Data Available

The formation depth for the main pay is 3154 m sub-sea. The average bubble point pressure is 2660 psia, the average initial reservoir pressure was 5186 psi and the current reservoir pressure is 2500 psia. The other available data is outlined below:

1. Well sketches with layer depth, thickness, and coordination.
2. Neutron porosity, shale volume, and water saturation logs for 20 wells.
3. Core data measurements and lithofacies sequence for 2 wells.
4. PVT data, Capillary Pressure, and Relative Permeabilities data.
5. Production and injection history from 1954 until 2010 with current well status and perforation intervals.
6. Thickness and depth maps for all five reservoir zones.

### 3. Applied Multivariate Geostatistics

The applied multivariate Geostatistics or Exploratory Data Analysis includes petrophysical properties of permeability and porosity, as well as rock facies modeling. The petrophysical response can be modeled from well log, core measurements, and DST data in order to predict the petrophysical properties at non-cored intervals and other wells. Since the well log and core measurements have different scales from discrete sources, statistical modeling could be an efficient way to link between these variables for accurate modeling and prediction.

In this chapter, the rock types, porosity, permeability are analyzed, modeled and predicted given the core measurements and well log interpretations of neutron porosity, shale volume, and water saturation. The predicted rock types, porosity, and permeability are then incorporated for spatial discrete and continuous reservoir property modeling through various geostatistical simulation approaches.

#### 3.1 Univariate and Bivariate Data Analysis

The early steps in the multivariate statistics are to perform univariate and bivariate statistical analyses. Univariate statistics includes testing whether or not the response factors such as porosity and permeability have normal distributions prior to modeling. One of the statistical tools that efficiently transform the data into normal or Gaussian distribution is the Box-Cox transformation. In geostatistics, the normal distribution is necessary to handle the sparseness of geological data in order to obtain the best regression fit.

Box-Cox transformation achieves approximate normal distribution by estimating the appropriate value of  $\lambda$  to obtain normal transformation (Box and Cox, 1964) as shown in Equation 3.1.

$$y_i^{(\lambda)} = \begin{cases} \frac{y_i^\lambda - 1}{\lambda} & \text{for } \lambda \neq 0 \\ \ln(y_i) & \text{for } \lambda = 0 \end{cases} \quad (3.1)$$

The  $\lambda$  has been empirically estimated based on a likelihood function of data mean and variance through fitting the response as a function of the explanatory variables in linear modeling as shown in Figure 3.1. The interval between the 1st and 3rd columns in Figure 3.1 represent the probable values of  $\lambda$  that lead to obtain normal distribution of the response variable. The bivariate analysis also includes the possible correlation between different

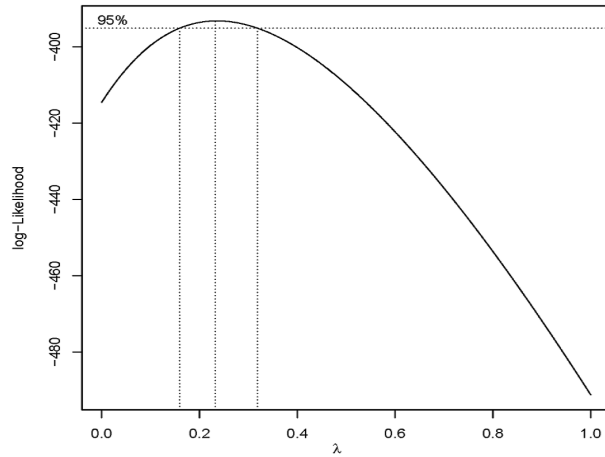


Figure 3.1: Estimation of Box-Cox Transformation Power Parameter (Al-Mudhafar, 2014)

factors. For example, there is a correlation between core porosity and core permeability or between log porosity and core permeability. Correlation is quantified by constructing a scatter plot between two variables. For instance, a scatter plot between log porosity and core permeability given lithofacies usually shows a low value area for shale and a high value area for sand. In addition, bivariate analysis includes correction of core permeability given corrected core porosity, which is corrected given log porosity in a prior step (Woodhouse, 2003).

## 3.2 Multivariate Lithofacies Modeling

Prediction of rock facies leads to adequate reservoir characterization by improving the porosity-permeability relationships to estimate the properties in non-cored intervals. Facies classification also lead to accurately identify the spatial facies distribution required to perform an accurate reservoir model for optimal future reservoir performance prediction. However, this procedure is difficult to implement with the conventional statistical approaches because of the lack of data. Moreover, the petrophysical properties are inferred from different sources and not directly measured in the reservoir, so they have levels of uncertainty (Pasternack, 2009; Moore et al., 2011). Therefore, advance statistical learning algorithms could be efficiently suitable for these sensitive cases.

In this work, comparative discrete and conditional posterior probabilities of continuous well facies distribution were estimated through Linear Discriminate Analysis (LDA), Kernel Support Vector Machine (KSVM), and Probabilistic Neural Networks (PNN) given explanatory variables. The core-measured lithofacies are sand, shaly sand, and shale. The explanatory variables represent well log data that includes neutron porosity, water saturation, and shale volume as function of depth. The multinomial response factor is the vertical lithofacies sequences of sand, shaly sand, and shale. The lithofacies were modeled given well log data through the aforementioned three algorithms to be compared based on matching between measured and predicted lithofacies distribution. The classification approach of highest modeling accuracy is considered for lithofacies prediction at other wells in the reservoir and for further spatial modeling. The entire process of lithofacies classification were implemented through R statistical language.

### 3.2.1 Linear Discriminant Analysis

Linear Discriminate Analysis (LDA) is a statistical technique that seeks a linear transformation (discriminate function) of both the dependent and explanatory variables in order to produce a new set of features that separates two or more classes of objects or events into different classes with reduction in dimensionality. The produced new features provides

easier classification because they are quantified by Fisher’s criterion that enables gaussian distribution for the resulting classes (Fisher, 1936). Excellent example of LDA description can be found in (Minka, 2005; Pires and Branco, 2010).

LDA is not be able to preserve any complex structure of data with non-Gaussian distribution as it is difficult to linearly discriminate between classes. Despite having many independent variables, the number of the produced classes depends on the number of discrete classes in the specific categorical variable or response factor (Li et al., 2006). Discrimination between the resulted classes requires a measure definition of separation between them, such as the distance between the projected means, in order to attain good projection vector. Fisher’s criterion is usually considered to maximize a function that represents the difference between the means (Guyon and Elisseeff, 2003). For more than two classes in the response categorical variable, Fisher’s criterion is used to generate multiple projections by means of projection vectors which can be arranged by columns into a projection matrix representing the number of original classes. In order to estimate the generalization error of a given model from many models, cross-validation can be used to rank and select the optimal model with the smallest estimated generalization error. It can be used to choose a subset of the independent variables that are relevant to all other variables.

LDA was chosen to estimate the maximum likelihood and minimize standard error for the nonlinear relationships between facies, core and well log data. The new set of linear combination separates the multinomial outcomes in order to reduce the dimensionality. Beta distribution of facies, as a conjugate prior for the multinomial facies distribution, was considered as prior knowledge and the resulted predicted probability was estimated from LDA based on Bayes’ theorem that represents the relationship between predicted probability (posterior) with the conditional probability and the prior knowledge. For two independent variables  $X_1$  and  $X_2$  in a decision boundary of the problem, which is orthogonal

to the direction of projection  $W$ , the linear combination can be represented as:

$$Y = W^T X + W_0. \quad (3.2)$$

$W$ : direction of projection in Linear Discriminant Analysis.

$X$ : the independent variable.

$\beta$ : coefficient of independent variables.

After projecting the data to the specified direction, the two projected coordinates are same and that happens along the decision boundary. Therefore, the linear combination can be formulated as:

$$W^T X_1 + W_0 = W^T X_2 + W_0 \Rightarrow W^T (X_1 - X_2) = 0. \quad (3.3)$$

Figure 3.2 shows the decision boundary between the new features.

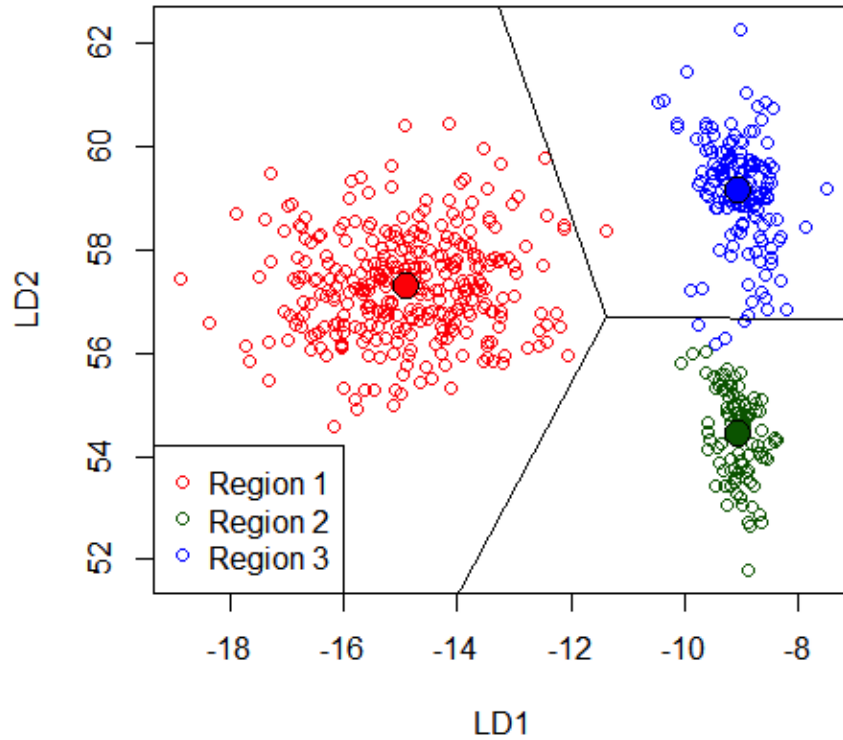


Figure 3.2: Linear Discriminant Analysis between Three Classes (Li et al., 2006)

The optimal classification function  $\hat{G}(X)$  between classes is:

$$\hat{G}(X) = \arg \max \delta_k(X) \quad (3.4)$$

$\hat{G}(X)$ : the optimal classification function in Linear Discriminant Analysis.

$\delta_k(X)$ : the decision boundary between classes in Linear Discriminant Analysis.

and the decision boundary between two classes  $k$  and  $l$  is (these classes are the same as the regions in the previous figure):

$$[x : \delta_k(x) = \delta_l(x)]. \quad (3.5)$$

To validate the LDA model and preserve the integrity of the statistical inference, cross-validation is used for estimating generalization error of a given model among various models and choosing an optimal model with the smallest generalization error. In order to fit the data through LDA, the data should be split into approximate equal size  $k$  subsets: train and test subsets. For each  $k$  subset, one fits the model with a tuning parameter (such as subset size) and computes the test error in predicting the  $k^{th}$  part as follows:

$$E_k(\lambda) = \sum_{i \in k_{th} part} (y_i - x_i \hat{\beta}^{-k}(\lambda))^2. \quad (3.6)$$

$k^{th}$ : subset of data.

$E_k(\lambda)$ : test error in predicting the  $k^{th}$  part.

The LDA validation could be adopted in different way. Leave-one-out cross-validation is used to estimate generalization error through leaving one observation out of dataset and fit the model. The predicted statistics with one out is then compared with the original entire dataset in order to estimate the variance given that observation (Ripley, 1996).



### 3.2.2 Kernel Support Vector Machines

The Support Vector Machines (SVM) is a supervised statistical learning algorithm that recognizes the discrete classes for the given data based on maximizing the margin around the separating hyperplane that is defined by the support vectors, and the decision function is fully specified by a subset of training samples, the support vectors. Therefore, the full classification requires solving the SVMs as a quadratic programming problem (Burges,, 1998). The quadratic optimization algorithms can identify which training points  $x_i$  are support vectors with non-zero Lagrangian multipliers  $\alpha_i$ . Both in the dual formulation of the problem and in the solution, training points appear only inside inner products. The linear classifier relies on an inner product between vectors (Ben-Hur and Weston, 2010):

$$K(X_i, X_j) = X_i^T X_j \quad (3.7)$$

$K$ : Kernel classifier between vectors. If every data point is mapped into high-dimensional space via some transformation  $\phi : x \rightarrow \phi(x)$ , the inner product can be corresponded to the kernel function in some expanded feature space:

$$K(X_i, X_j) = \phi X_i^T \phi X_j. \quad (3.8)$$

$\phi$ : vector transformation in SVM. In a two-class problem, the training set  $X = \{x_i, x_j\}_{i=1}^n$  is associated with:

$x_i \in R^d$  Points

$y_i \in \{-1, 1\}$  Labels.

SVM seeks an optimal separating hyperplane to separate two classes maximizes the distance to the closest point from either class (Vapnik, 1996). This is achieved by maximizing the margin between the two classes on the training data in order to obtain better classification performance on test data (Hastie et al., 2001). Since SVM is based on the notion that the

best canonical separating hyperplane needs be the farthest away from the training points, the larger margin the better hyperplane as illustrated in Figure 3.3.

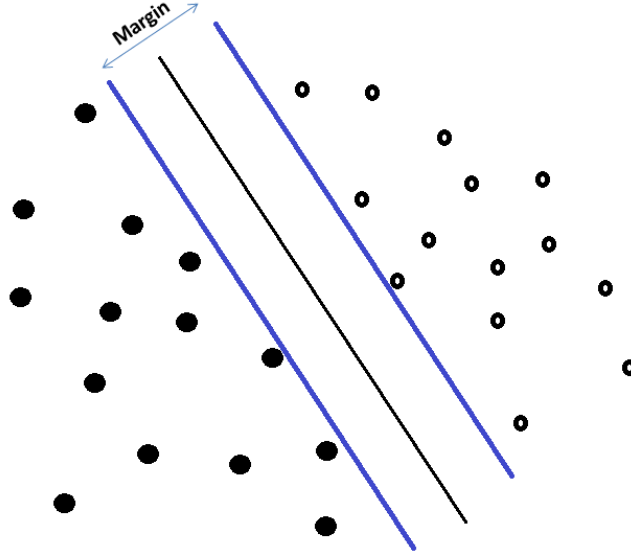


Figure 3.3: Separating Hyperplanes between Two Classes

A hyperplane in  $R^d$  consists of all  $x_i \in R^d$  and has a linear equation:

$$f(x) = \beta_0 + \beta^T X = 0. \quad (3.9)$$

By giving  $x_i$  and  $y_i$ , a separating hyperplane is:

$$y_i(\beta_0 + \beta^T X) \geq c. \quad (3.10)$$

The margin of a hyperplane is:

$$margin = 2 \times \min\{y_i d_i\} \quad (3.11)$$

$d_i$ : the signed distance between observation  $x_i$  and the hyperplane in SVM. *margin*: margin of a hyperplane in SVM. where  $d_i$  is the signed distance between observation  $x_i$  and the

hyperplane. The  $\beta$  can be computed from the following formula:

$$\beta = \sum_{[i=1]}^n \alpha_i y_i x_i \quad (3.12)$$

where  $\sum_{[i=1]}^n \alpha_i y_i = 0$  and  $\alpha_i \geq 0$ .

So the dual maximization objective function is:

$$\max \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j y_i y_j X_i^T X_j. \quad (3.13)$$

To adopt a kernel function into the SVM, the inner-product  $X_i^T X_j$  becomes  $K_h(x_i, x_j)$ , where  $h$  is a hyper parameter that is usually determined by cross-validation, the previous function can be written as:

$$\max \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j y_i y_j K_h(x_i^T, x_j). \quad (3.14)$$

Then, the decision function can be written as:

$$f(x) = \sum_{i=1}^n \alpha_i y_i K_h(x_i^T, x_j) + \beta_0 = 0. \quad (3.15)$$

$h$ : the hyper parameters determined by cross-validation in SVM.  $K_h$ : the hyper kernelization function in SVM. The kernelization of inner-product is achieved by assuming a data matrix with  $n \times d$  dimensions, where  $n$  is the number of observations  $x_1, x_2, \dots, x_n \in R^d$  and  $d$  is the row vector. The linear algorithm can be written as:

$$K = XX^T. \quad (3.16)$$

Then, it is simply a replacement of each of the inner-product entry of  $K$  with  $K_{ij} = K_h(x_i, x_j)$ .

Finally, the Kernel SVM maximizes the minimum distance from data to separator boundary (maximum margin with simple classifier) through representing nonlinear boundaries with feature expansion. Therefore, KSVMs usually achieve good error rates and can handle unusual types of data.

### 3.2.3 Probabilistic Neural Networks

Probabilistic Neural Networks (PNN) was adopted as a non-linear classification approach to model the Lithofacies given well log data.

A probabilistic neural network is principally a classifier to map any input pattern to a number of classifications. PNN is a multi-layer feedforward network that was derived from Bayesian network and Kernel Fisher discriminant analysis (Specht, 1988). PNN solves a common problem in automatic learning by classifying the observations in a given number of groups. After training the datasets, the group of any new observation can be guessed. PNN advantages come from its ability to train from a small set of input data. Convergence on an optimal classifier is ensured with increasing the size of the representative training set at no local minima constraints (Cheung and Cannons, 2003).

PNN operations are organized into a multi-layered feed forward network with four layers: input, pattern, summation and output (Cheung and Cannons, 2003). The architecture of PNN with the related mathematical formula of its procedure is depicted in Figure 3.4.

LDA, KSVM and PNN were applied on the available well data set to predict the discrete and posterior probability Lithofacies distributions for other wells. The same data set was considered for modeling by other methods such as Naive Bayes Classifier, Boosting Logistic Regression, Multinomial Logistic Regression, and Tree-based Classification Algorithms (Al-Mudhafar, 2016). It was investigated that PNN and KSVM are the most accurate classifications over the first aforementioned algorithm. To evaluate the most accurate Lithofacies classification among these algorithms, the total percent correct is computed based on the transition probability matrix that is derived from the predicted discrete Lithofacies. The transition probability matrix is constructed by dividing the number of facies sequences over

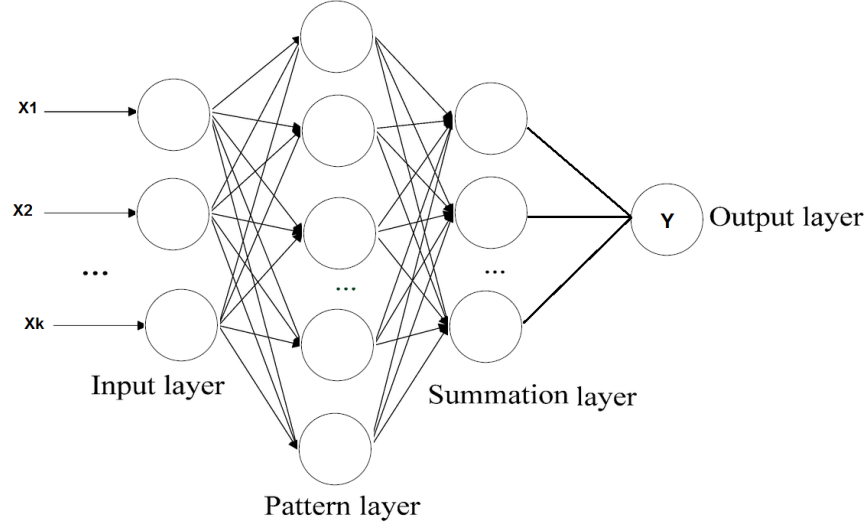


Figure 3.4: Structure of Probabilistic Neural Networks Algorithm

the total number of all predicted facies given each row. The PNN model was finally adopted to predict the discrete and continuous distribution of Lithofacies (Sand, Shaly Sand, and Shale) for the other 10 wells in the reservoir under study (Chasset, 2015). The discrete Lithofacies distribution was included in the multivariate petrophysical property modeling to obtain different prediction approaches for each lithofacies type.

### 3.3 Multivariate Petrophysical Modeling

The applied multivariate geostatistics encompasses the petrophysical properties of permeability, porosity, as well as rock facies modeling. The permeability should be modeled given well log and core measurements for wells in order to predict the petrophysical parameters at non-measured intervals and other wells. Modeling and prediction of formation permeability is a decisive step in reservoir characterization as it involves the sparseness of row data that has different scales and inferred from different sources (Al-Mudhafar, 2014). Efficiently combining the different sources of permeability especially from core and well log data leads to accurate estimation for other wells that have no core analysis and thus improve overall reservoir characterization and simulation.

The common method to model and predict the core permeability as a function of other core and well log data is multiple linear regression that considers the backward/forward or

stepwise methods for subset selection (Al-Mudhafar, 2014). Stepwise elimination depends on the probability of the null hypothesis (p-value) to eliminate predictors that have no effect on the response. The probability of the null hypothesis is calculated for each predictor through partial t-test. Null hypothesis claims that the factor has no effect on the response. When the hypothesis is rejected (P-value < 0.05), that factor should be kept at the model and should not be eliminated (Efron, 2004).

The core permeability modeling was comparatively carried out in this study through Generalized Linear Modeling (GLM) and Smooth Generalized Additive Model (sGAM). sGAM has more efficient subset selection technique through smooth terms than stepwise elimination. Both of the two algorithms were implemented in R statistical language (Hothorn and Everitt, 2009).

### 3.3.1 Generalized Linear Modeling

Generalized Linear Modeling, GLM, is similar to linear models as there is a continuous response variable given continuous and/or categorical predictors:

$$y_i \sim N(X_i^T \beta, \sigma^2) \quad (3.17)$$

where  $X_i$  represents the independent variables,  $\beta$ : the coefficients, and  $\sigma^2$ : the variance of error term  $E_i \sim N(0, \sigma^2)$ .

GLM includes multiple linear regression, Analysis of Variance (ANOVA), and Analysis of Covariance (ANCOVA). Parameter estimation does not depend on minimizing the sum of squared errors. However, maximum likelihood estimation is iteratively fitting the model in order to ensure most likely estimation of the observed data (Dobson, 1990).

The form of the linear predictor is:

$$\eta_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i}, \dots, \beta_p X_{pi}. \quad (3.18)$$

The link function relates the mean value of response with the linear predictors:

$$\eta_i = g(\mu_i) \quad (3.19)$$

GLM is generally solved through iterative weighted Fisher's Method of Scoring:

$$\hat{\beta}^{(t+1)} = \hat{\beta}^{(t)} + (X^T W^{(t)} X)^{-1} X^T W^{(t)} Z^{(t)}. \quad (3.20)$$

Model sub-selection in Generalized Linear Modeling also adopts the probability of null hypothesis to eliminate non-influential predictors that have no effect on the model.

### 3.3.2 Smoothed Generalized Additive Modeling

Smoothed Generalized Additive Regression (sGAM) considers a sum of nonparametric smoothing functions to identify nonlinear relationships depending on the degree of smoothing. The smooth function minimizes a penalized negative log-likelihood function, which represents the smooth terms, by minimizing of an internal generalized cross-validation function through iteratively re-weighted least squares (Hastie and Tibshirani, 1990).

The Generalized Additive Model is defined to be:

$$y = \alpha + \sum_{j=1}^n f_j(X_j) + \epsilon \quad (3.21)$$

where  $f_j$  is a non-parametric or nonlinear function in GAM (Brockett et al., 2014).

The function  $f_j(X_j)$  is estimated by summarizing the trend of Y, usually from the plot, as a function of the independent variables with less variation than the Y itself which reflects the property of smoothing (Friedman et al., 2000). Attaining less variation requires minimizing the cubic spline smoother by choosing the smoothing parameter  $\lambda$  by measuring the global

error which is the Average Mean Square Error (AMSE):

$$AMSE(\lambda) = \frac{1}{n} \sum_{i=1}^n E\{f_{\lambda}(X_i) - f(X_i)\}^2 \quad (3.22)$$

where  $AMSE(\lambda)$  is the global error which is the Average Mean Square Error (AMSE) and  $f_{\lambda}(X_i)$  indicates the fit at  $X_i$  is computed by leaving out the  $i^{th}$  data point in GAM. The smoother parameter  $\lambda$  can be also estimated by cross validation sum of square as follows:

$$CV(\lambda) = \frac{1}{n} \sum_{i=1}^n E\{y_i - f_{\lambda}(X_i)\}^2 \quad (3.23)$$

$\lambda$ : the smoother parameter.  $CV(\lambda)$ : cross-validation that estimates  $\lambda$ . The smoother parameter  $\lambda$  can be likewise estimated by generalized cross validation sum of square and the fit is computed by leave-one-out data (Guisan et al., 2002):

$$CV(\lambda) = \frac{1}{n} \sum_{i=1}^n E\left\{\frac{y_i - f_{\lambda}(X_i)}{1 - S_{ii}}\right\}^2 \quad (3.24)$$

Elimination of non-influential predictor(s) in sGAM also depends on the null hypothesis, but smoothing terms help to identify and deactivate these predictors through the modeling process.

### 3.4 Results and Discussion

The well log interpretations that were considered for modeling and estimation of the discrete and continuous Lithofacies sequences are neutron porosity, shale volume (from Gamma Ray Log), and water saturation, all given well depth. There are only 20 wells that have well log interpretations in the field sector. The vertical Facies sequences, Core porosity, and horizontal and vertical core permeability have been obtained for one well only with limited intervals. All data has been provided by the Iraqi Ministry of Oil in LAS file format.



There is only one well that has core measurement including the vertical lithofacies sequences in addition to the well log interpretations of neutron porosity, shale volume, and water saturation. The well log interpretations of the main well that is considered for Lithofacies and petrophysical modeling, were illustrated in Figure 3.5. The mean distri-

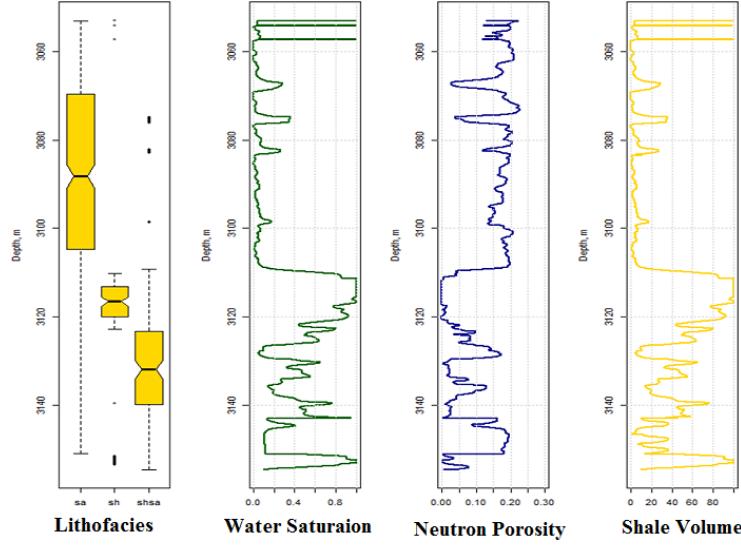


Figure 3.5: Vertical Lithofacies and Well Logs Distribution

bution and histograms of core permeability and porosity given the Lithofacies are depicted in Figure 3.6 and Figure 3.7, respectively. Unlike the porosity distribution, the left-hand-side bar for permeability histogram in Figure 3.7 represents both shale and shaly sand as both of them have values less than 200 md, as shown in the right-hand-side figure. The relationship between each Lithofacies type and the well log variables has been depicted in Figure 3.8. In this figure, the lines represent the density of property measurements given each lithotype. It also indicates the effect of the well log on each of the facies distribution. For instance, the log porosity has higher values in sand than water saturation and shale volume values. However, shale volume and water saturation have higher values in shale than log porosity.

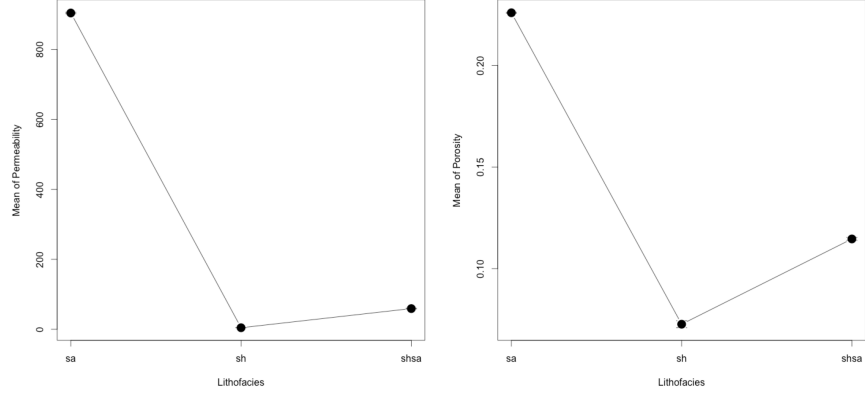


Figure 3.6: Mean Distribution of Permeability and Porosity given Sand, Shaly Sand, and Shale Facies

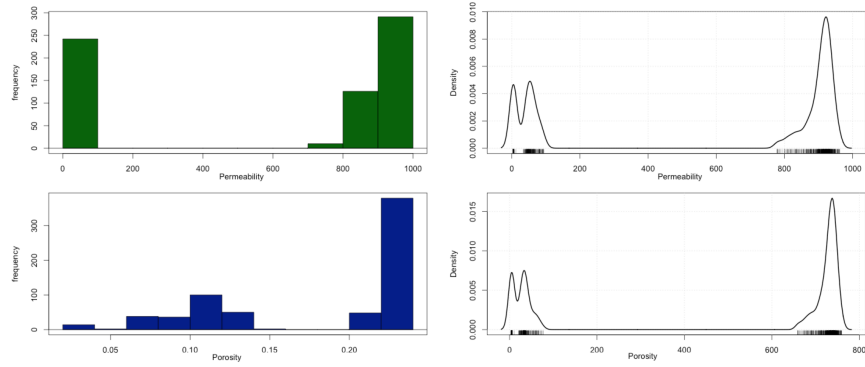


Figure 3.7: Histogram of Permeability and Porosity given Sand, Shaly Sand, and Shale Facies

### 3.4.1 Results of Multivariate Lithofacies Modeling

For full intervals Lithofacies distribution, a comparative workflow of Linear Discriminant Analysis (LDA), Kernel Support Vector Machines (KSVM), and Probabilistic Neural Networks (PNN) was conducted to model and estimate the discrete and continuous Lithofacies distribution prior to fit the petrophysical modeling. The well log interpretations were considered for Lithofacies modeling as independent variables in order to predict the Lithofacies for other wells that have same well log. In order to use LDA, it was needed firstly to split the data first into two parts: one for training and another for testing the classifier. For this problem, approximately 80% was considered for training and 20% for testing. Nevertheless, the LDA classification can be performed in different way without splitting the data into a training and testing, but with cross validation procedure. The

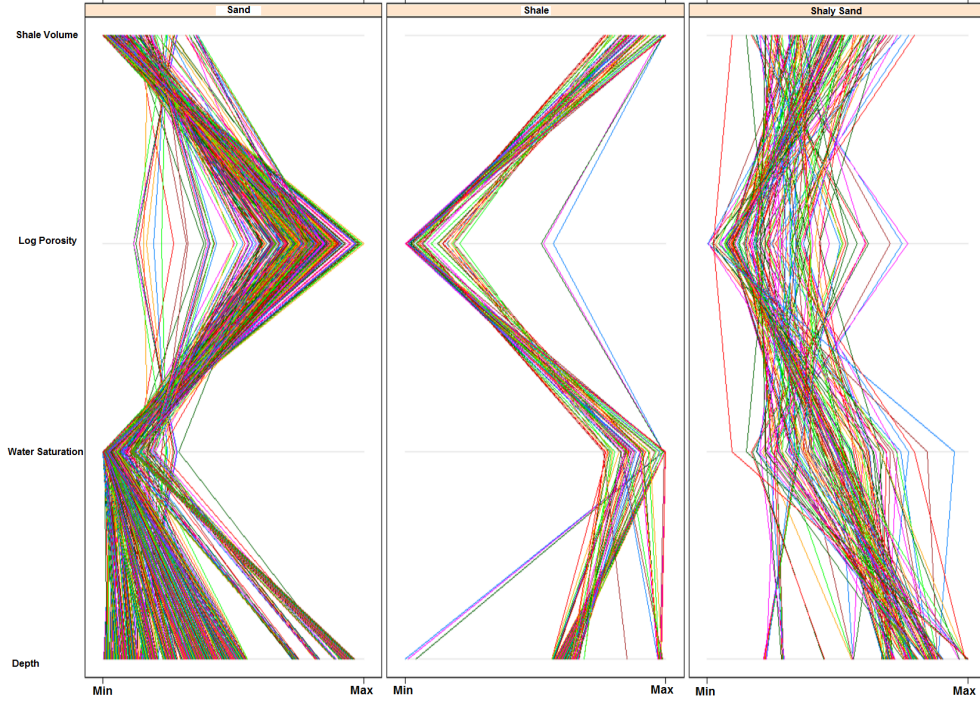


Figure 3.8: Relationship between each Lithotype and the Well Log Variables

beta distribution of original Lithofacies sequences was incorporated as prior distribution into LDA computations. The resulting discrimination of Lithofacies through Linear Discriminant Analysis was shown in the scatter matrix plot in Figure 3.9.

Figure 3.10 represents the discriminated Lithofacies given the created LDA components. It is noticeable that LDA is not efficient to discriminate between the three Lithofacies because the LDA considers a linear separation between the discriminated classes given the two components (LDA1 and LDA2).

To assess the accuracy of linear discriminant analysis modeling, the total correct percent for Facies classification was computed from the summation of the transition probability matrix diagonal that represents the correct percent for the three facies classification. The correct percent through LDA modeling is 94.76831% and the error in the facies classification is more than 5%. The error indicates that the Linear Discriminant Analysis for Lithofacies classification for other wells is not accurate. Therefore, it was necessary to look for another method that provide high accuracy value of classification. The resulting predicted discrete

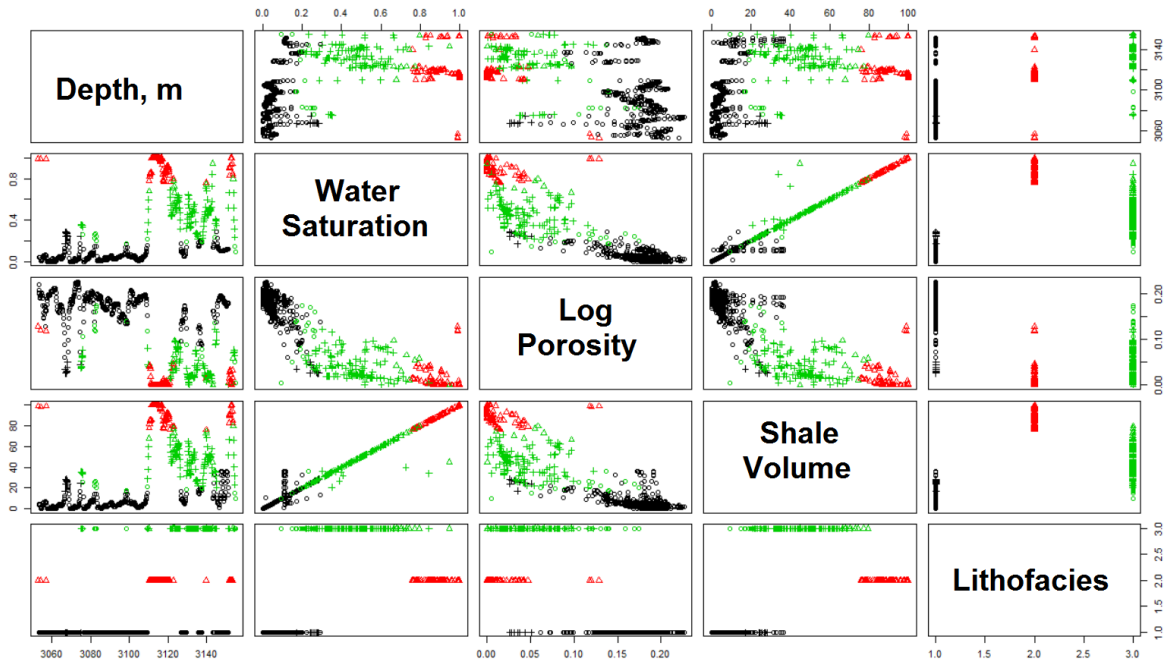


Figure 3.9: LDA Discrimination of Lithofacies Given the Well Logs Data

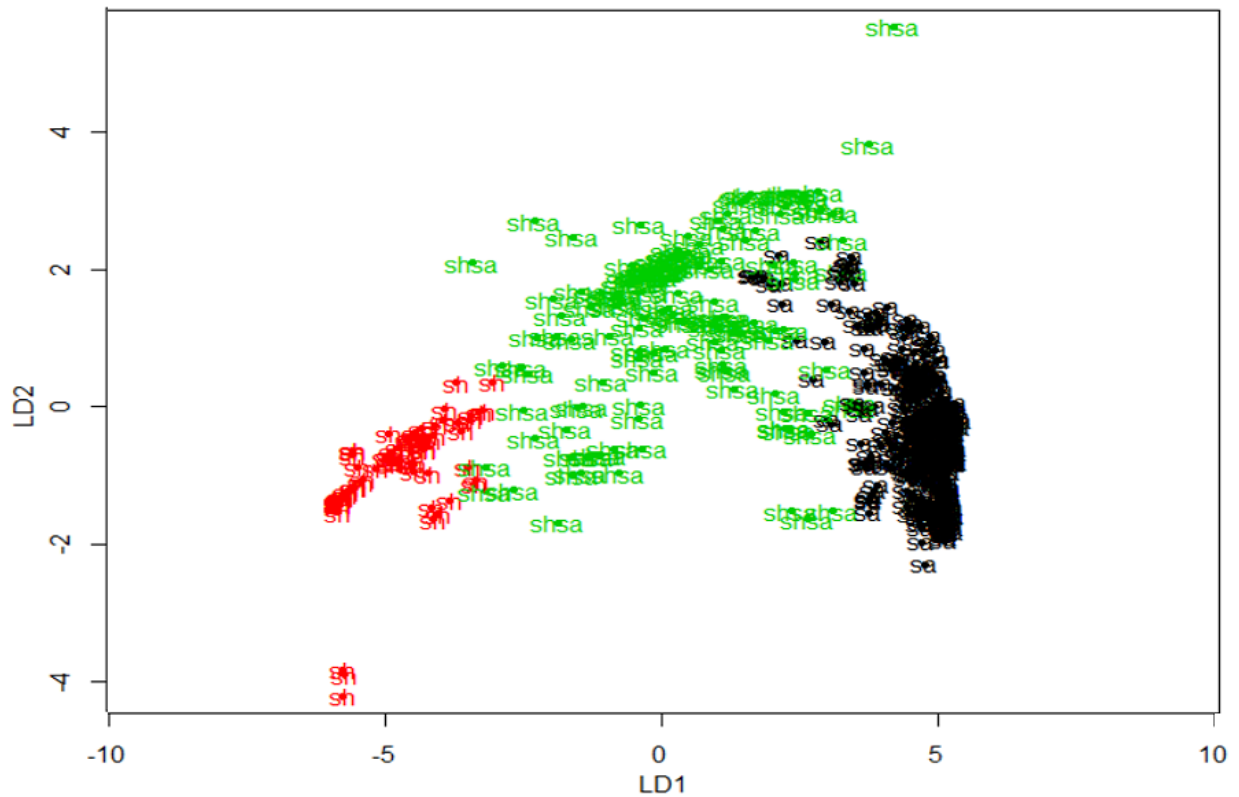


Figure 3.10: Scatter Plot of the Discriminated Lithofacies Given the LDA Components

and continuous posterior distribution for the three Lithofacies from LDA were illustrated in Figure 3.11. In this figure, the discrete lithofacies distributions were represented by boxplots. Each boxplot represents the location of a lithotype with respect to the reservoir depth. The same dataset for that well was fitted through Kernel Support Vector Machines

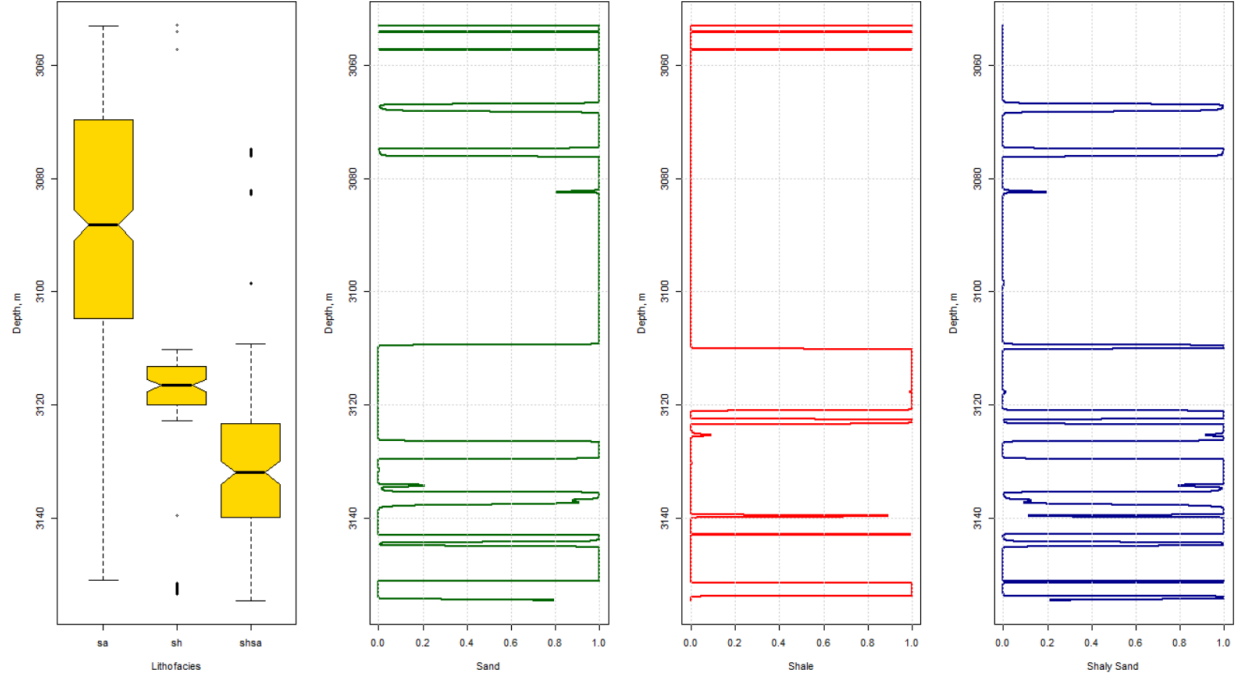


Figure 3.11: The Predicted Discrete and Probability Distribution of the LDA Classified Lithofacies

(KSVM). The accuracy of classification in KSVM is much better than the LDA because the total correct percent 99.55157% and there is very subtle and non-considerable error in Lithofacies modeling. Figure 3.12 shows the observed and predicted discrete with posterior probability distribution of the Lithofacies given the entire depth through KSVM. Finally, the PNN algorithm was adopted to repeat the same procedure of Lithofacies modeling, validation, and prediction. Figure 3.13 shows excellent matching between the observed and predicted discrete distribution of Lithofacies by PNN with a total percent correct was 99.81%. It should be noted that both Kernel Support Vector Machine and Probabilistic Neural Networks have the same accuracy and exceeds Linear Discriminant Analysis in Facies classification. Consequently, the PNN model was considered for predicting of the

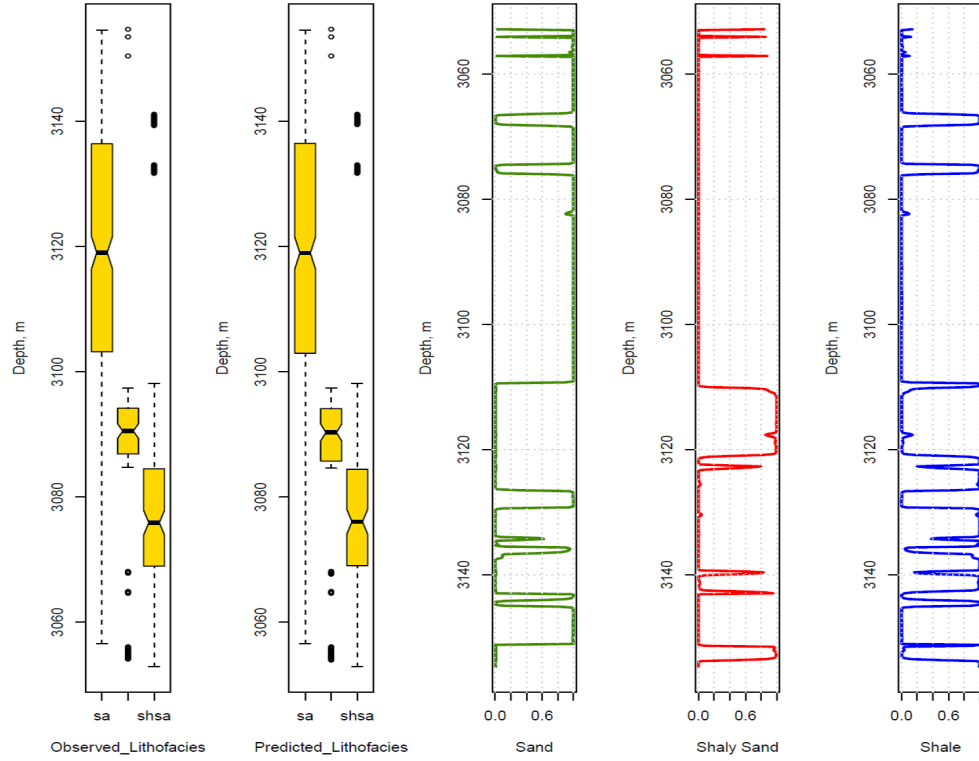


Figure 3.12: The Observed and Predicted Discrete and Posterior Distribution of KSVM Classified Lithofacies for the Entire Depth Intervals

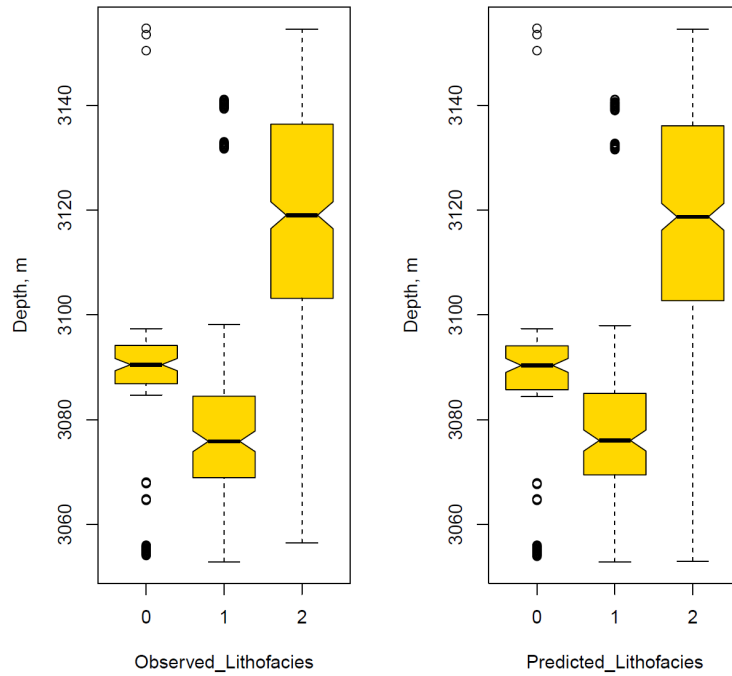


Figure 3.13: Box-Plot Comparison between Observed and Predicted Discrete Lithofacies Distribution by PNN

facies sequence in other wells to be incorporated in multivariate petrophysical modeling and Geostatistical simulation.

### 3.4.2 Results of Multivariate Petrophysical Modeling

The PNN predicted discrete lithofacies distribution of Sand, Shaly Sand, and Shale of the 20 wells was considered as an additional independent variable in the upcoming permeability correction procedure, core permeability, and porosity modeling.

In order to obtain the corrected core permeability, it was necessary to model core porosity given log porosity and Lithofacies which was conducted through linear regression modeling. After predicting corrected core porosity for the entire depth intervals, corrected core permeability was then obtained given the corrected core porosity since both share the same scale. Fitting core permeability as function of core porosity and Lithofacies was also implemented through linear regression. The regression models were statistically tested for their validity via partial t-test and coefficient of determination ( $R^2$ ) to ensure that p-values are less than 0.05 reflecting the rejection of null hypothesis for non-influential variables on the response and confidence intervals of more than 95%.

The scatter matrix plot of the predicted core porosity, corrected horizontal core permeability, the classified Lithofacies, in addition to other well log interpretations is depicted in Figure 3.14. The main diagonal of matrix represents the histogram distribution of the various parameters. The lithofacies have discrete distribution. However, other parameters have continuous distributions. The non-diagonal plots refer to the relationship between every two parameters. The off-diagonal plots represent the scatter plots between the parameters located at the diagonal plots. The red lines refer to the data trend of the relationship between every two parameters. There is high correlation between some parameters, such as water saturation and shale volume because both reflect the same rock property. In order to predict core permeability and core porosity for other wells, Generalized Linear Models and Smooth Generalized Additive Models were comparatively applied for modeling given well log data and discrete Lithofacies distribution, which was obtained by PNN. First,

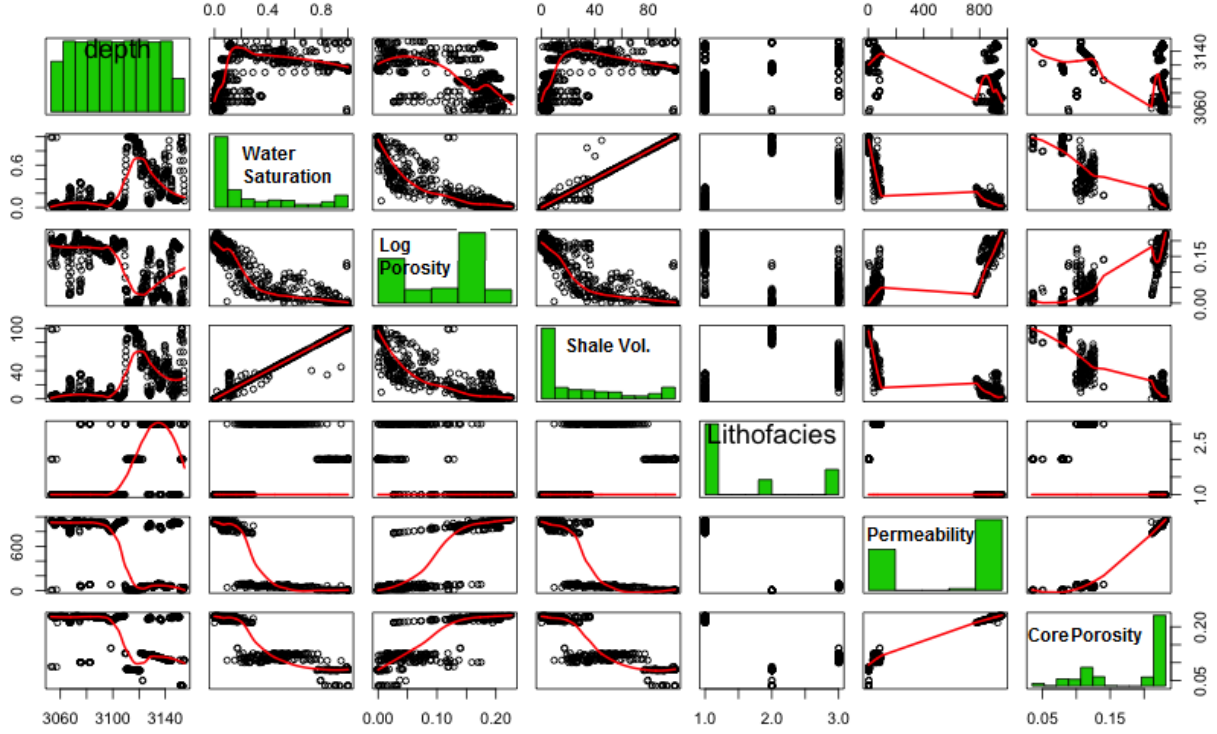


Figure 3.14: ScatterMatrix Plot for Predicted Lithofacies and Petrophysical Properties

the observations that we have for the main well were sampled and split into two groups: Train (80 %) and Test (20%). The data was first fit on training data in order to perform a prediction given test data to check the accuracy of internal prediction prior to predicting for the same well data or other wells. The procedure was compared between GLM and sGAM methods.

In the GLM algorithm, stepwise elimination was adopted to eliminate the non-influential factors on the response. Table 3.1 and Table 3.2 represent the partial t-test analyses for full and reduced GLM models, respectively. In these tables, the 2nd column represent the estimated coefficient values of the independent variables, which were shown in the 1st column. The 3rd and 4th columns refer to the standard error and t-values. The  $\Pr(>|t|)$  represents the probability of null hypothesis given each independent variable. The stars in last column refer to the strength of null hypothesis rejection. For instance, three stars mean stronger rejection of null hypothesis ( $0.001 > p\text{-value} \approx 0.00$ ) than one star ( $0.05 > p\text{-value} > 0.01$ ). Figure 3.15 shows the scatter-plot between observed and predicted permeability for the



reduced GLM model of train data that has ( $AdjustedR^2 = 0.9969$ ) and low Root Mean Square Prediction Error ( $RMSPE = 21.2184$ ).

Table 3.1: Partial t-test for the Full Generalized Linear Model

Variables	Estimate	Std. Error	t value	Pr(> t )	
Depth	-0.0134	0.0569	-0.24	0.8135	
Log Porosity	635.0998	41.4557	15.32	0.0000	***
Shale Volume	-0.9535	0.3413	-2.79	0.0054	**
Water Saturation	62.8455	31.3839	2.00	0.0458	*
Sand	836.7464	176.5308	4.74	0.0000	***
Shale	64.1099	175.0159	0.37	0.7143	
Shaly Sand	90.4722	177.1178	0.51	0.6097	
<hr/>					
Residual standard error:	23.52 on 493 DF				
Multiple R-squared:	0.9969	Adj.R-squared:	0.9968		
F-statistic:	2.632e+04	p-value:	< 2.2e-16		

Table 3.2: Partial t-test for the Reduced Generalized Linear Model

Variable	Estimate	Std. Error	t value	Pr(> t )	
Log Porosity	636.3787	41.0609	15.50	0.0000	***
Shale Volume	-0.9591	0.3401	-2.82	0.0050	**
Water Saturation	62.6843	31.3465	2.00	0.0461	*
Sand	795.1208	7.9043	100.59	0.0000	***
Shale	23.0420	18.8999	1.22	0.2234	
Shaly Sand	48.7241	9.3198	5.23	0.0000	***
<hr/>					
Residual standard error:	23.49 on 495 DF				
Multiple R-squared:	0.9969	Adj.R-squared:	0.9969		
F-statistic:	3.957e+04	p-value:	< 2.2e-16		

However, GLM is non-accurate modeling because the predicted values are not fully fit to the theoretical straight line. Therefore, the same sampled and split data was re-fitted through the smooth Generalized Additive Models (sGAM). The smoothed terms were adopted for all variables prior to re-fit the model. If the smoothed term is applied on a variable and still has no rejection for the null hypothesis ( $p-value > 0.05$ ), that variable has no effect on the response and should be deleted from the model. That procedure has led to accurate modeling and prediction for porosity and permeability. The smooth terms

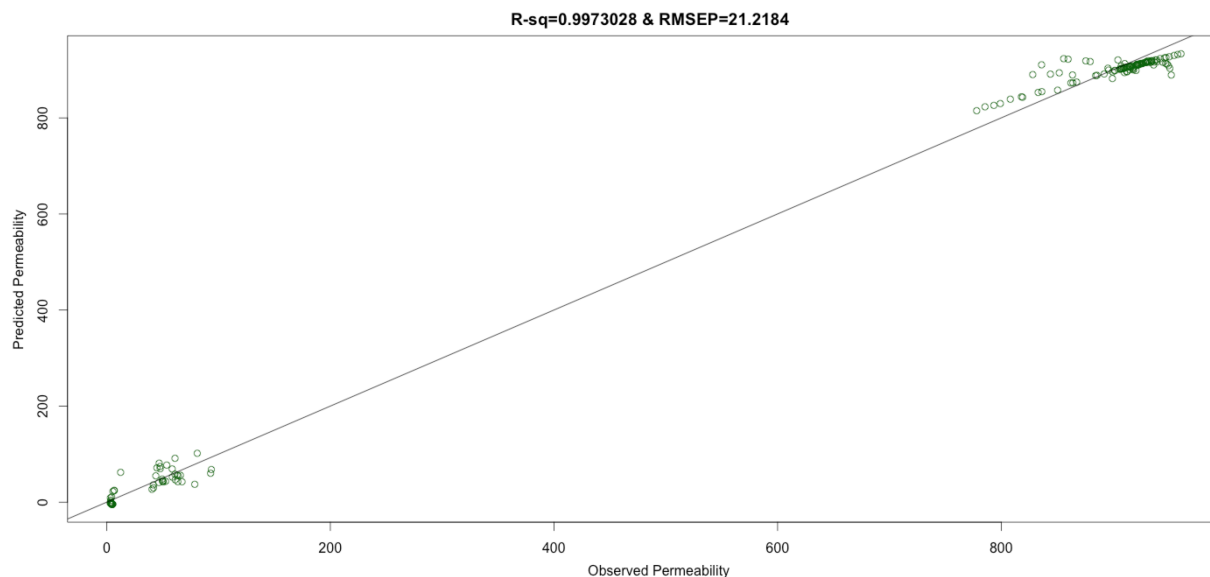


Figure 3.15: Scatterplot of Observed and Predicted Permeability-Reduced GLM Model

have resulted in more accurate modeling once the null hypotheses were fully rejected for all independent variables leading to high confidence interval ( $> 95\%$ ). The sGAM model better models than the GLM with ( $AdjustedR^2 = 0.999$ ) and ( $RMSPE = 15.75$ ). Table 3.3 and Table 3.4 represent the partial t-test analyses for full and reduced sGAM modeling, respectively. The detaining of these tables were provided in the previous GLM modeling procedure.

Perfect modeling and prediction was obtained from sGAM as the smooth terms have shown accurate effects of the continuous predictors on the response in full and reduced sGAM models as shown in Figure 3.16. This figure shows the relationship between the parameter before (x-axes) and after applying the smooth term (y-axes). Shale volume parameter has no effect on the response since it failed to reject null hypothesis and it has horizontal trend lines with the x-axes. Therefore, it was eliminated from the model as shown in Figure 3.17. The shale volume is eliminated because the collinearity between it with water saturation as they have similar data behavior. Therefore, one of them should be removed from the modeling. Figure 3.18 shows the scatter-plot between observed and predicted permeability for the reduced sGAM model based on train data.

Table 3.3: Partial t-test for Full Permeability sGAM Model

Family: Gaussian		Link function: Identity				
Factor	Estimate	Std. Error	t value	Pr(> t )		
Sand	870.506	2.081	418.31	<2e-16	***	
Shale	-796.946	9.194	-86.68	<2e-16	***	
Shaly Sand	-760.793	4.551	-167.19	<2e-16	***	
Smoothed Term	edf	Ref.df	F	p-value		
s(Depth)	9.000	9	43.317	< 2e-16	***	
s(Log Porosity)	8.232	9	28.801	< 2e-16	***	
s(Water Saturation)	4.707	9	4.189	6.61e-10	***	
s(Shale Volume)	2.164	9	0.264	0.259		
R-sq.(adj)	= 0.999	Deviance explained =	99.9%			
GCV score	= 230.17	Scale est. =	217.69	n = 500		

Figure 3.19 shows comparative matching between the observed and predicted per-

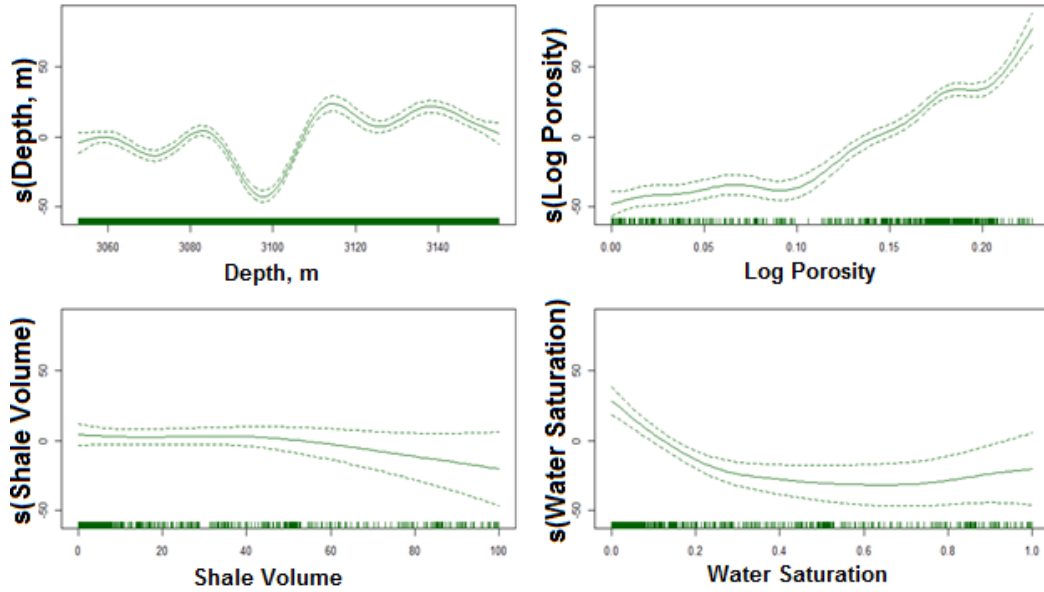


Figure 3.16: Smooth Term Effect of the Continuous Predictors in Full sGAM model

meability distribution through reduced GLM and sGAM models on the train/test data. It can be shown that the sGAM model is more accurate than the GLM because it captured the subtle changes in the observed curves readings. Over the formation depth, there are many intervals have exact matching between the measured and predicted permeability.

Table 3.4: Partial t-test for the Reduced Permeability sGAM Model

Family: Gaussian		Link function: Identity				
Factor	Estimate	Std. Error	t value	Pr(> t )		
Sand	870.839	2.036	427.62	<2e-16	***	
Shale	-800.259	8.660	-92.41	<2e-16	***	
Shaly Sand	-760.178	4.520	-168.19	<2e-16	***	
Smoothed Term	edf	Ref.df	F	p-value		
s(Depth)	9.000	9	43.446	< 2e-16	***	
s(Log Porosity)	8.217	9	29.085	< 2e-16	***	
s(Water Saturation)	4.978	9	9.349	6.61e-10	***	
R-sq.(adj)	= 0.999	Deviance explained =	99.9%			
GCV score	= 229.94	Scale est. =	218.35	n = 500		

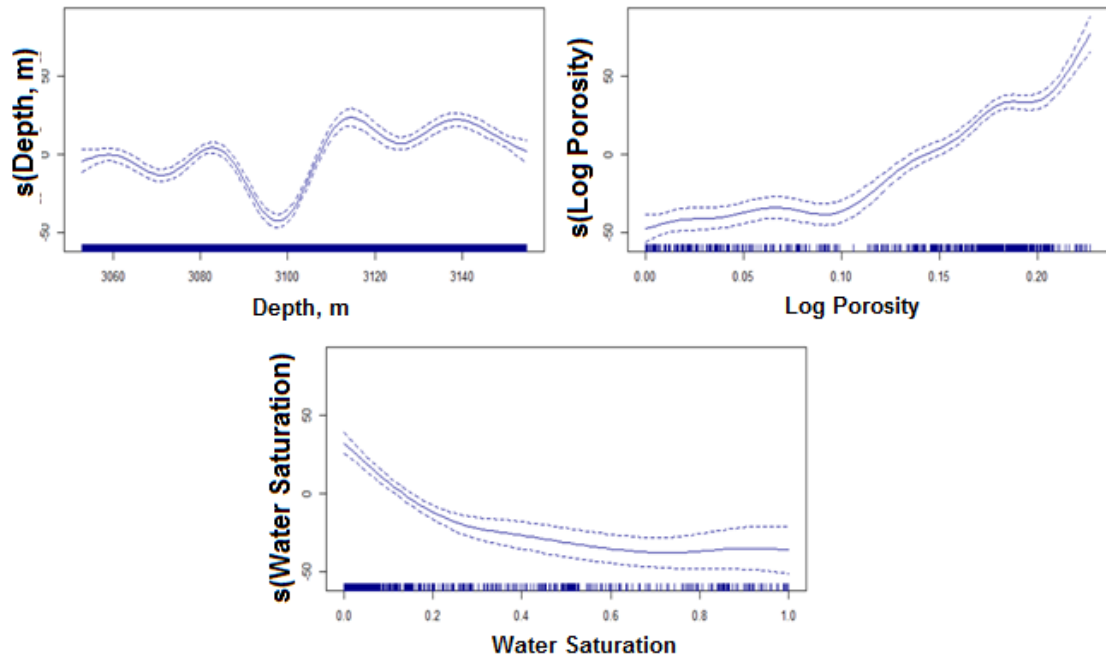


Figure 3.17: Smooth Term Effect of the Continuous Predictors in Reduced sGAM model

Additionally, There are mismatch intervals in the reduced Generalized Linear Model larger than the reduced Genralized Additive Model.

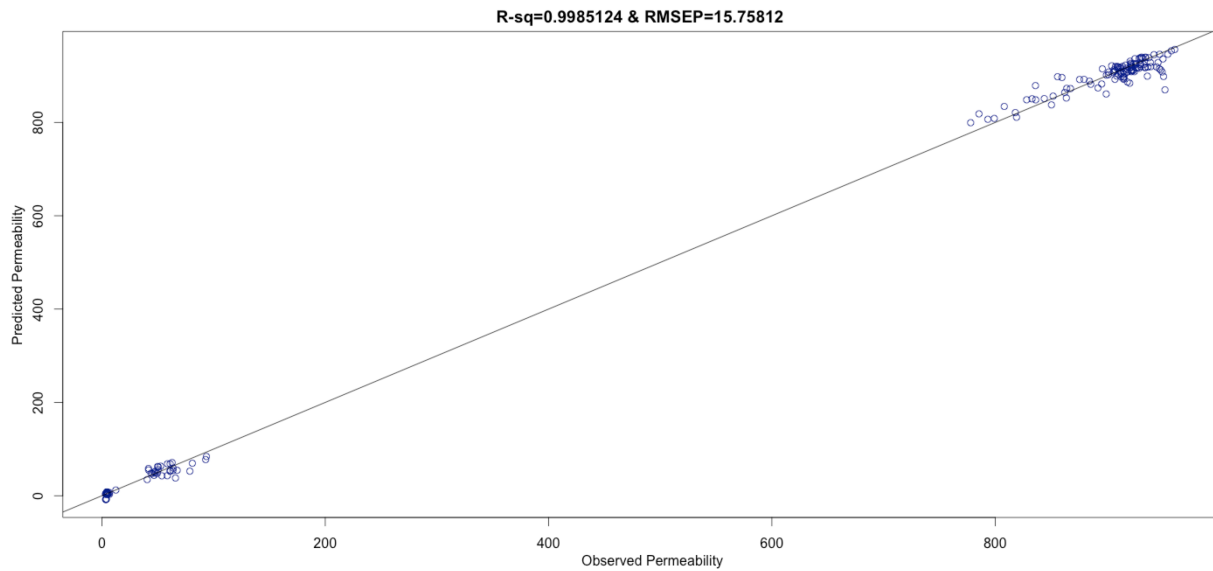


Figure 3.18: Scatterplot of Observed and Predicted Permeability-Reduced sGAM Model

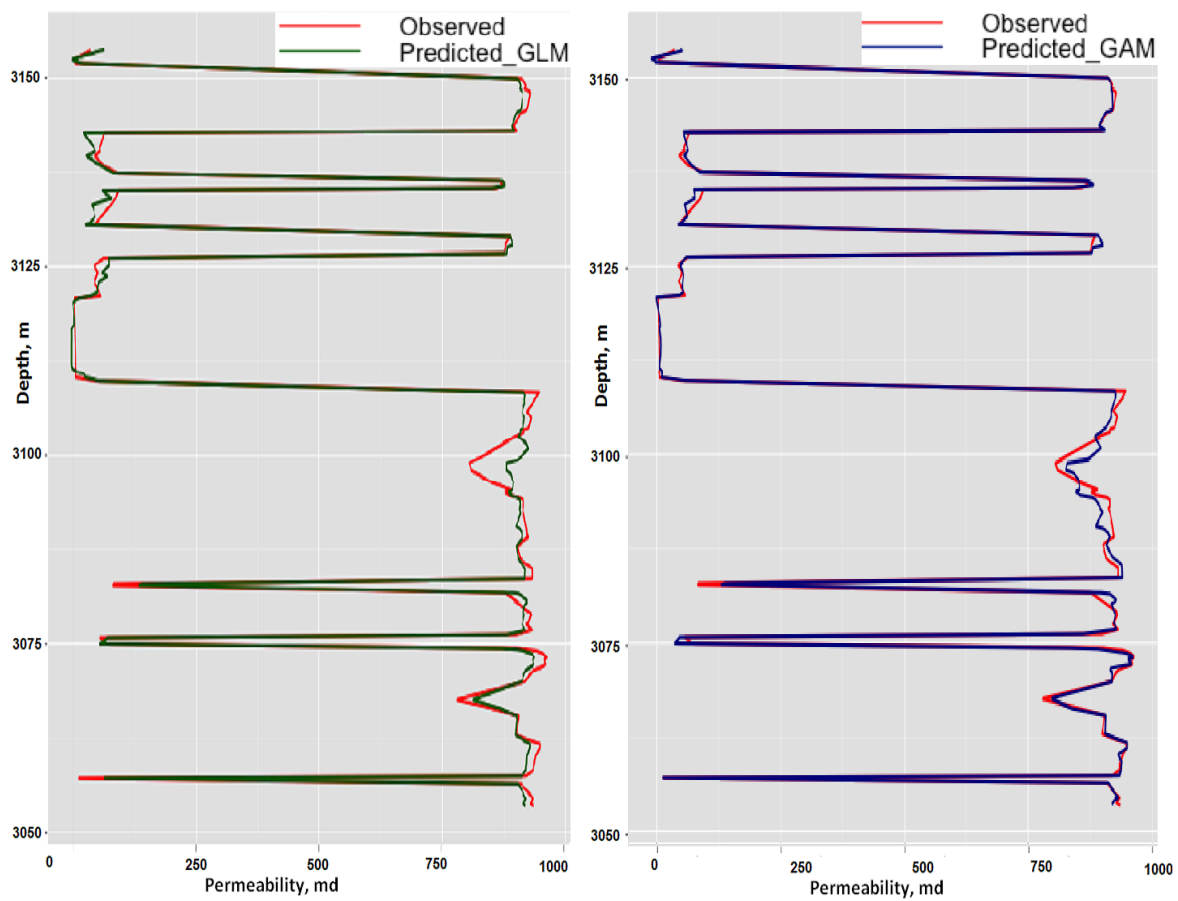


Figure 3.19: Observed and Predicted Permeability-Reduced GLM and sGAM Models

The aforementioned figures and tables were also created for Train/Test data for the main well. Figure 3.20 shows the matching between Observed and Predicted Permeability through Reduced GLM and sGAM Models for the entire-depth dataset. It is noticeable that the reduced sGAM matching is much better than the reduced GLM for the entire depth intervals of the formation.

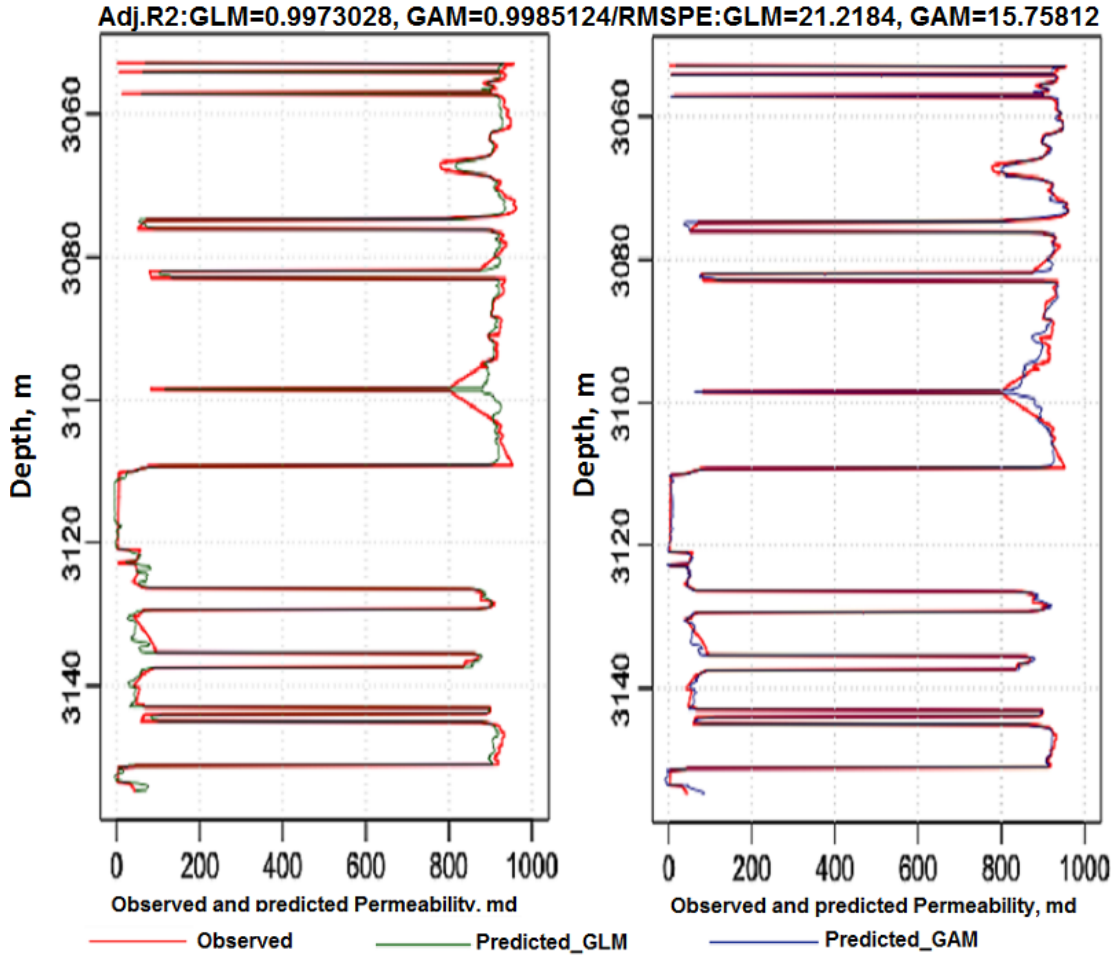


Figure 3.20: Full Depth Permeability Matching-Reduced GLM and sGAM Models

The full procedure of the comparison between Generalized Linear Modeling and Smooth Generalized Additive Modeling was repeated for core porosity modeling given the same train/test dataset. Table 3.5 and Table 3.6 represent the partial t-test analyses for full and reduced GLM of porosity modeling, respectively. Figure 3.21 shows the scatter-plot

between observed and predicted porosity for the reduced GLM model of train data with ( $AdjustedR^2 = 0.9884265$ ) and ( $RMSPE = 0.006493749$ ). The modeling of core porosity

Table 3.5: Partial t-Test for Full Porosity GLM Model

Variable	Estimate	Std. Error	t value	Pr(> t )	
Depth	-0.0001	0.0000	-4.84	1.71e-06	***
Water Saturation	-0.0021	0.0089	-0.23	0.8170	
Log Porosity	0.0685	0.0123	5.58	3.96e-08	***
Shale Volume	0.0002	0.0001	2.32	0.0209	*
Sand	0.4393	0.0468	9.38	< 2e-16	***
Shaly Sand	-0.1558	0.0036	-43.13	< 2e-16	***
Shaly Sand	-0.1066	0.0019	-57.11	< 2e-16	***
Residual standard error: 0.007999 on 493 DF					
Multiple R-squared:	0.9843	Adj.R-squared:	0.9842		
F-statistic:	3.957e+04	p-value:	< 2.2e-16		

Table 3.6: Partial t-Test for reduced Porosity GLM Model

Variable	Estimate	Std. Error	t value	Pr(> t )	
Depth	-0.0001	0.0000	-4.84	1.70e-06	***
Log Porosity	0.0693	0.0117	5.90	6.71e-09	***
Shale Volume	0.0002	0.0000	4.11	4.69e-05	***
Sand	0.4383	0.0466	9.40	< 2e-16	***
Shale	-0.1560	0.0034	-45.52	< 2e-16	***
Shaly Sand	-0.1067	0.0018	-58.86	< 2e-16	***
Residual standard error: 0.007991 on 494 DF					
Multiple R-squared:	0.9843	Adj.R-squared:	0.9842		
F-statistic:	5166	p-value:	< 2.2e-16		

resulted from smooth Generalized Additive Models since ( $AdjustedR^2 = 0.9923537$ ) and ( $RMSPE = 0.005278232$ ) were both better than GLM as shown in Figure 3.23. All factors in sGAM porosity modeling are influential and reject null hypothesis. Therefore, all should be kept in the model as shown in Figure 3.22. Table 3.7 illustrates partial t-test analyses for full sGAM on porosity modeling. Figure 3.24 shows a comparative matching between observed and predicted porosity distribution between reduced GLM and sGAM models on the train/test data. Figure 3.25 shows the matching between observed and

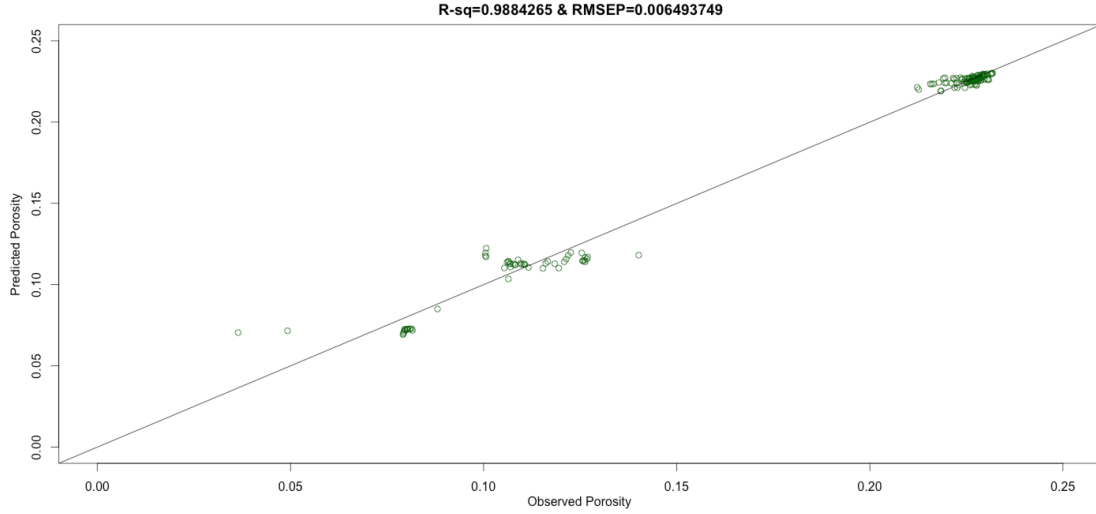


Figure 3.21: Scatterplot of Observed and Predicted Core Porosity-Reduced GLM Model

predicted porosity through Reduced GLM and sGAM Models for the entire-depth dataset.

Based on the results of smooth Generalized Additive Models in multivariate petrophysical

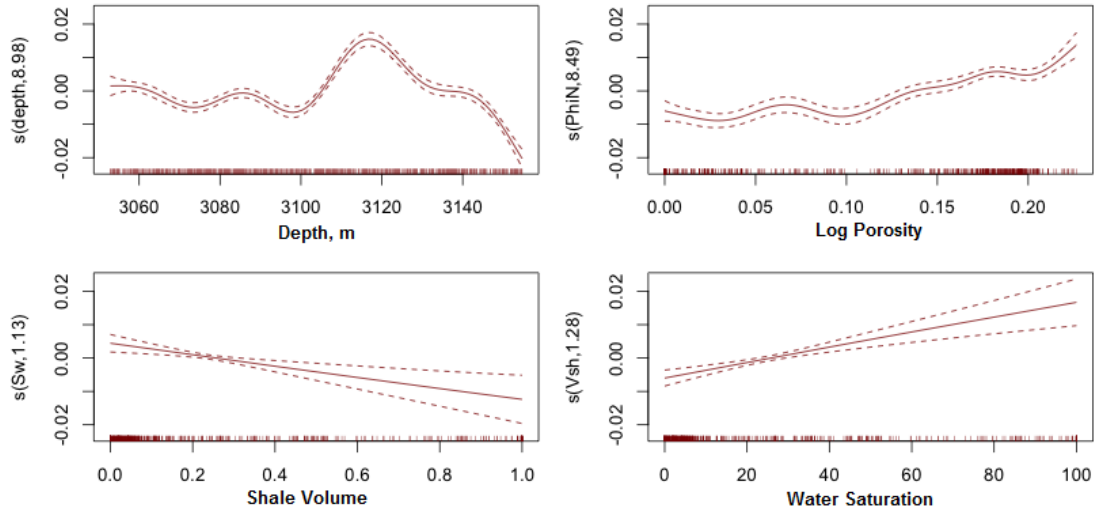


Figure 3.22: Smooth Term Effect of the Continuous Predictors in Full sGAM model

modeling, the scatter plot between log and core porosity given lithofacies for the 20 wells were included in Appendix A. The final predicted Lithofacies through Probabilistic Neural Networks in addition to predicted porosity and permeability for all other wells were depicted in Appendix B.



Table 3.7: Partial t-Test for Full Porosity sGAM Model

Family:	Gaussian	Link function:	Identity			
Variable	Estimate	Std. Error	t value	Pr(> t )		
Sand	0.2260011	0.0006154	367.23	<2e-16	***	
Shale	-0.1613318	0.0024290	-66.42	<2e-16	***	
Shaly Sand	-0.1073110	0.0013007	-82.50	<2e-16	***	
Smoothed Term	edf	Ref.df	F	p-value		
s(Depth)	8.980	9	82.523	< 2e-16	***	
s(Log Porosity)	8.556	9	13.526	< 2e-16	***	
s(Water Saturation)	1.143	9	0.697	0.00744	**	
s(Shale Volume)	1.301	9	2.036	7.62e-06	***	
R-sq.(adj)	= 0.994	Deviance explained =	99.4%			
GCV score	= 2.5106e-05	Scale est.	= 2.3952e-05	n = 500		

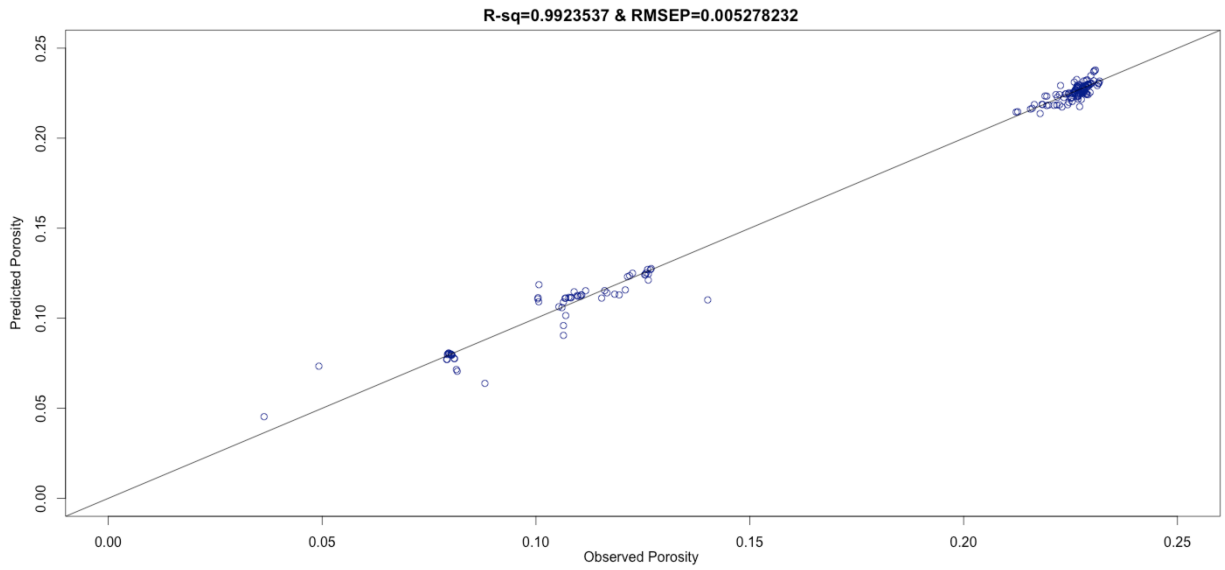


Figure 3.23: Scatterplot of Observed and Predicted Core Porosity-Reduced sGAM Model

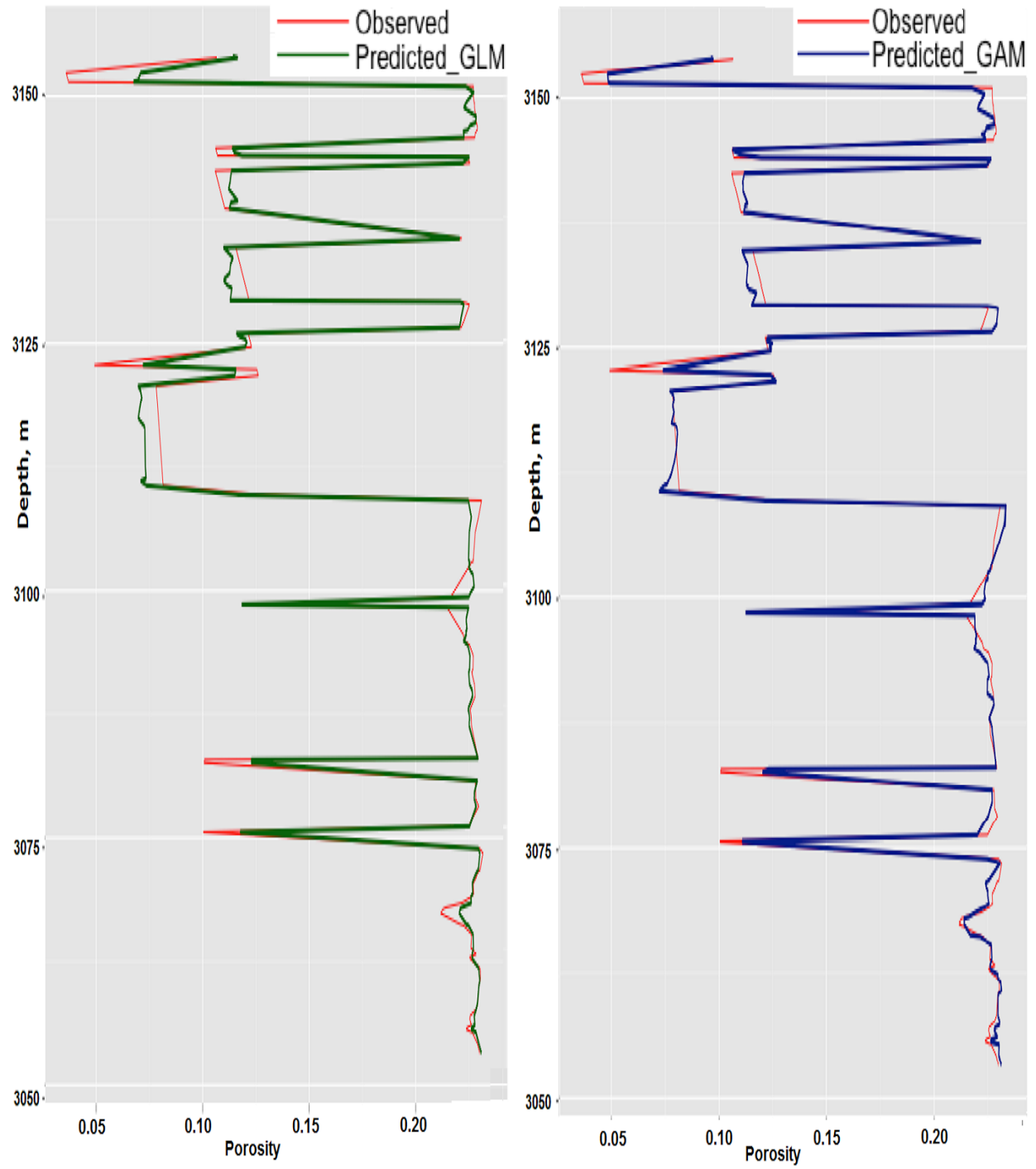


Figure 3.24: Matching of Core Porosity-Reduced GLM and sGAM Models

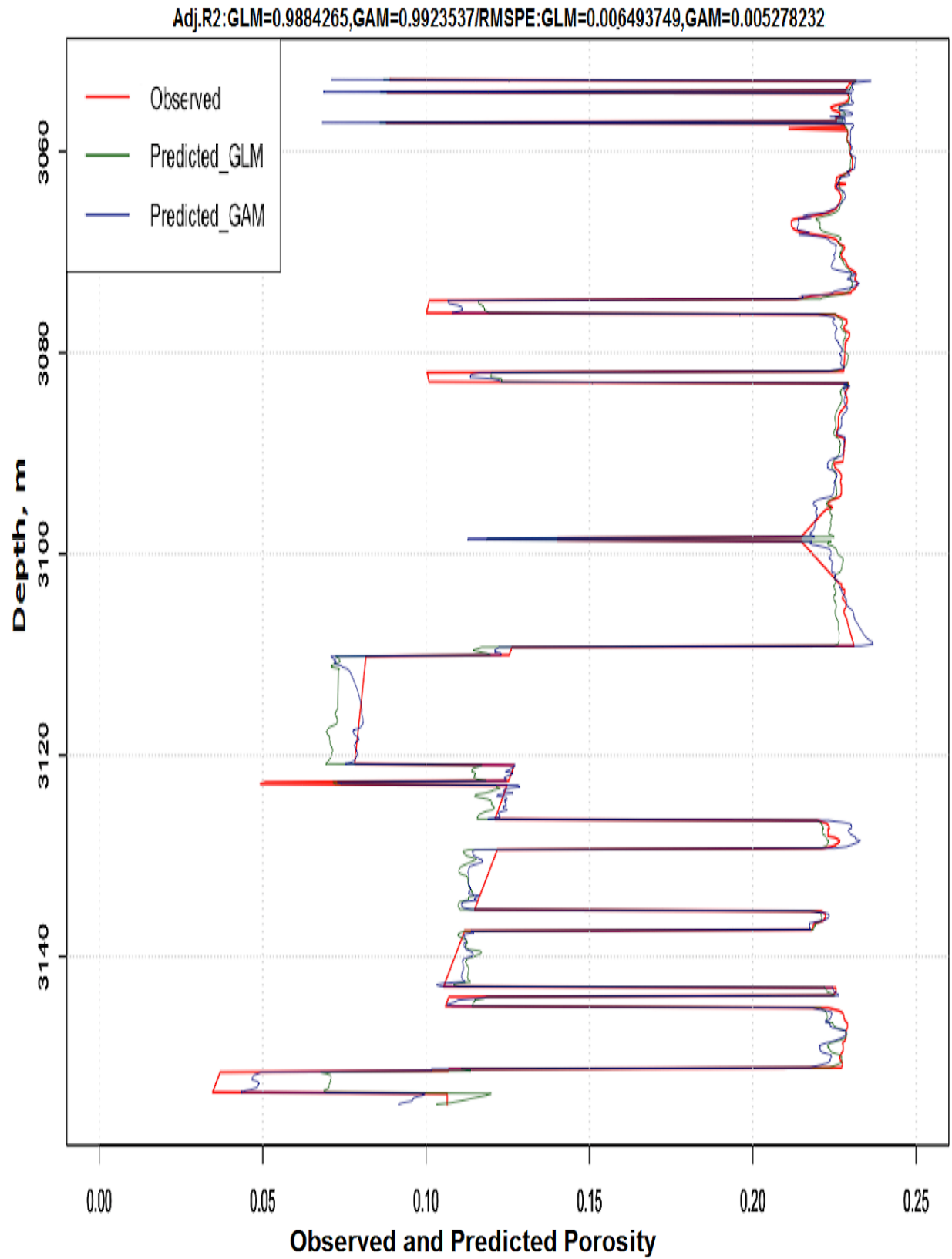


Figure 3.25: Full Depth Matching of Core Porosity-Reduced GLM and sGAM Models

### 3.5 Summary and Significance

In this chapter, an efficient workflow of multivariate geostatistical analysis was implemented for petrophysical and lithofacies modeling and estimation. The workflow includes adopting different statistical data mining and machine learning tools for modeling and prediction the lithofacies classifications and petrophysical properties in a well in the main pay/South Rumaila oil field.

In lithofacies modeling, Linear Discriminant Analysis, Kernel Support Vector Machine, and Probabilistic Neural Networks were comparatively adopted in to model the lithofacies, sand, shaly sand, and shale, given the well log interpretations: neutron porosity, shale volume, and water saturations. Based on the total correct percent of the predicted lithofacies classification, it was noted that Probabilistic Neural Networks and kernel Support Vector Machine are the most accurate algorithms as the total correct percent exceeds 99%. However, the Linear Discriminant Analysis led to not more than 95%. Therefore, the Probabilistic Neural Networks was adopted for prediction of the discrete lithofacies distributions at other wells.

Next, the lithofacies sequence was included in the petrophysical modeling as an independent discrete variable to predict the permeability and porosity given each lithotype. Prior to model the petrophysical properties, a correction procedure was conducted to correct the core porosity given log porosity. Then, the corrected core porosity was used to obtain the corrected core permeability in order to obtain a calibration between the core and log scales. Then, the permeability and porosity was modeled given the well log interpretations and discrete lithofacies distribution in order to predict the core permeability and porosity at 11 wells in the reservoir. The modeling was carried out through Generalized Linear Modeling (GLM) and smooth Generalized Adaptive Models (sGAM). It was concluded that the sGAM is more accurate than GLM because it led to more accurate matching between the observed and predicted petrophysical property distributions than GLM. That matching was also identified by the Root Mean Square Prediction Error (RMSPE).

## 4. Geostatistical Reservoir Modeling

Isaaks and Srivastava (1989) stated that "Geostatistics offers a way of describing the spatial continuity of natural phenomena and provides adaptations of classical regression techniques to take advantage of this continuity". Geostatistics integrates mathematical concepts, computer technology, and stochastic modeling to generate multiple equiprobable realizations that keep the reservoir heterogeneity through honoring all the available data (Journel, 1990; Deutsch and Journel, 1998; Liu et al., 2004; Caers and Zhang, 2004). These alternative stochastic images (realizations) are created through conditional stochastic simulation that has the ability of reproduction of extreme values, both high permeability channels and low permeability barriers, in order to quantify the geological spatial uncertainty (Caers, 2005). The geological uncertainties come because of incomplete information regarding the modeled phenomenon (Goovaerts, 1997; Pyrcz and Deutsch, 2014). Most Geostatistical reservoir characterization models are variogram-based algorithms such as: sequential indicator simulation for facies modeling and sequential Gaussian simulation for continuous petrophysical parameters.

### 4.1 Variogram-based Algorithms

Variogram-based geostatistics adopt variogram models to describe the variation between any two spatial data locations (Gringarten and Deutsch, 1999). That's why it is also called two-point statistics (Zhang, 2008, Liu et al., 2004) or traditional two-point geostatistics (Deutsch and Journel, 1998). The variogram describes the geometry and continuity of reservoir properties and has direct impact on the flow behavior. The variogram can be

defined as the measure of dissimilarity between known and unknown data as the distance increases and mathematically is the expected squared difference between two data values separated by a distance vector  $h$  (Gringarten and Deutsch, 1999).

$$2\gamma(h) = E[Y(u) - Y(u + h)]^2 \quad (4.1)$$

The variogram terminology includes sill (the plateau that variogram reaches the range); range (The distance at which the variogram is no longer increases with distance increases); and nugget (measurement error when distance is zero).

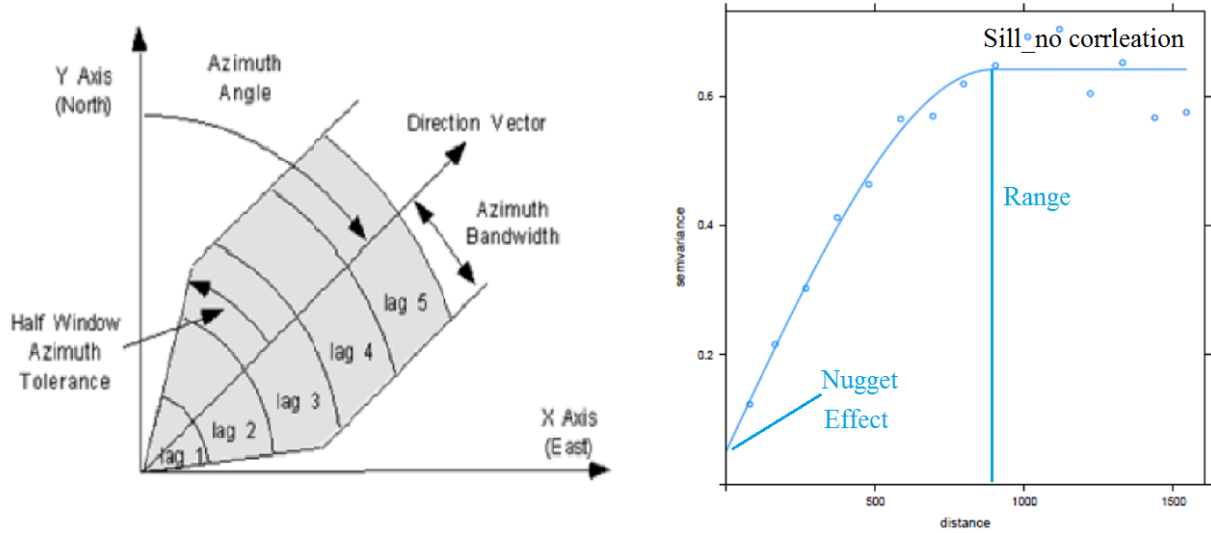


Figure 4.1: Variogram Structure

$$\gamma(h) = var(z) - cov(h) \quad (4.2)$$

$$Dissimilarity(h) = Sill - Similarity(h)$$

Simple kriging is then used for spatial data distribution based on best variogram fit.

1.  $K$  is the covariance matrix between the known data.

$$K = \begin{pmatrix} cov(Z_1, Z_1) & cov(Z_1, Z_2) & \cdots & cov(Z_1, Z_n) \\ cov(Z_2, Z_1) & cov(Z_2, Z_2) & \cdots & cov(Z_2, Z_n) \\ \vdots & \vdots & \ddots & \vdots \\ cov(Z_n, Z_1) & cov(Z_n, Z_2) & \cdots & cov(Z_n, Z_n) \end{pmatrix} \quad (4.3)$$

2.  $k$  is the covariance vector between the known and unknown data.
3.  $\lambda$  is the kriging weights.

$$k = \begin{pmatrix} cov(Z_1, Z_0) \\ cov(Z_2, Z_0) \\ \vdots \\ cov(Z_n, Z_0) \end{pmatrix} \sim \lambda = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{pmatrix} \quad (4.4)$$

4. The final formula of simple kriging is:

$$z_0 = \sum_{i=1}^n \lambda_i z_i \quad (4.5)$$

Where  $\sum_{i=1}^n \lambda_i = 1$ .

The variogram-based conditional simulation algorithms include Sequential Gaussian Simulation for continuous variables, and Sequential Indicator Simulation for categorical variables.

## 4.2 Geostatistical Lithofacies Modeling

The spatial Lithofacies model is very important because it accurately reflects the stratigraphic reservoir structure as it includes all the hydraulic elements. The Lithofacies model

also helps an individual to understand the description of different depositional environments and capture all the heterogeneity levels and scales to be integrated in the reservoir flow model (Mike and Geel, 2006; Walker, 1992). In addition, the facies model can serve as a guide to predict the reservoir properties in different locations with the same depositional environment, by controlling the reservoir heterogeneities and fluid flow characteristics (Walker, 1992). Since the sGAM model for petrophysics utilizes lithofacies indicator, these indicators are needed to predict petrophysical properties away from wells.

The most common Geostatistical facies simulations are sequential indicator simulation, object-based model, and multiple-point geostatistics.

#### 4.2.1 Sequential Indicator Simulation

The most common Geostatistical modeling for reservoir simulation is Sequential Indicator Simulation (SISIM) that is designed for modeling the spatial distribution of facies based on the indicator variogram (Massonnat et. al., 1992; Journel and Alabert, 1990). The indicator variogram is used to build up a discrete cumulative density function (CDF) for the individual facies types and the node is assigned a lithotype,  $k$ , selected at random from this discrete CDF (Journel and Gomez-Hernandez, 1993). After encoding Facies into elementary samples 0, 1 given threshold values, the indicator variogram then can be formulated as:

$$I(Z_k; x) = \begin{cases} 0 & \text{if } Z(x) > Z_k \\ 1 & \text{if } Z(x) \leq Z_k \end{cases} \quad (4.6)$$

Where  $I(Z_k; x)$  is the indicator random variable that is associated with random function  $Z(x)$  for a threshold value  $Z_k$ .

The expected value of the indicator random variable  $I(Z_k; x)$  is equal to the cumulative probability  $Pr \{ Z(x) < Z_k \}$  as shown below:

$$E(I(Z_k; x)) = 0 \times Pr \{ Z(x) > Z_k \} + 1 \times Pr \{ Z(x) \leq Z_k \} \quad (4.7)$$



$$E(I(Z_k; x)) = Pr \{ Z(x) < Z_k \} \quad (4.8)$$

where  $Pr \{ Z(x) < Z_k \}$  represents the cumulative probability in SISIM. According to the logic described about SISIM, the SISIM is considered as a non-parametric sequential simulation (Caers, 2000). The main SISIM procedural steps for the Facies spatial modeling after establishing the grid network and coordinate system are (Goovaerts, 1997; Journel and Gomez-Hernandez, 1993; Pyrcz and Deutsch, 2014):

1. Create the indicator variogram for each lag distance based on the indicator lithofacies.

$$\gamma(h) = \frac{1}{2N_h} \sum_{i=1}^{N_h} (facies_{(h+i)} - facies_{(h)})^2 \quad (4.9)$$

where  $N_h$  is number of points included in indicator variogram.

The prior distribution function represents the density distribution of the facies and it is calculated:

$$F(z_i) = \sum_{j=1}^{i-1} P(z_j) \quad (4.10)$$

where  $F(z_i)$  represents the prior distribution function of facies and  $P(z_j)$  refers to the density distribution of the facies.

2. Select randomly all the un-sampled locations to be simulated.
3. Consider the indicator kriging to estimate the probability that un-sampled location prevails given the indicators values at the surrounding locations.
4. Specify the indicator values to be either 0 or 1 randomly.
5. After adding the simulated indicator value to the sampled data group, repeat the procedure for the remaining un-sampled locations.
6. Repeat the procedure for all the realizations.

### 4.2.2 Multiple-Point Geostatistics

To overcome the limitations of variogram-based stochastic modeling, multiple-point geostatistics (MPS) was introduced to incorporate geological analogues, outcrops, seismic, production data, and well petrophysical data to reconstruct the depositional facies environments and model the most realistic reservoir heterogeneity (Deutsch and Journel, 1998; Strebelle, 2003; Caers and Zhang, 2004). The variogram stores the geological structures in a mathematical formula based on two-point statistics. However, it cannot capture the complex formation patterns. Additionally, three different heterogeneities from pixel-based and object-based methods (Deutsch and Journel, 1998) can have approximately the same variogram (Caers and Zhang, 2004) as shown in Figure 4.2. MPS combines the strength

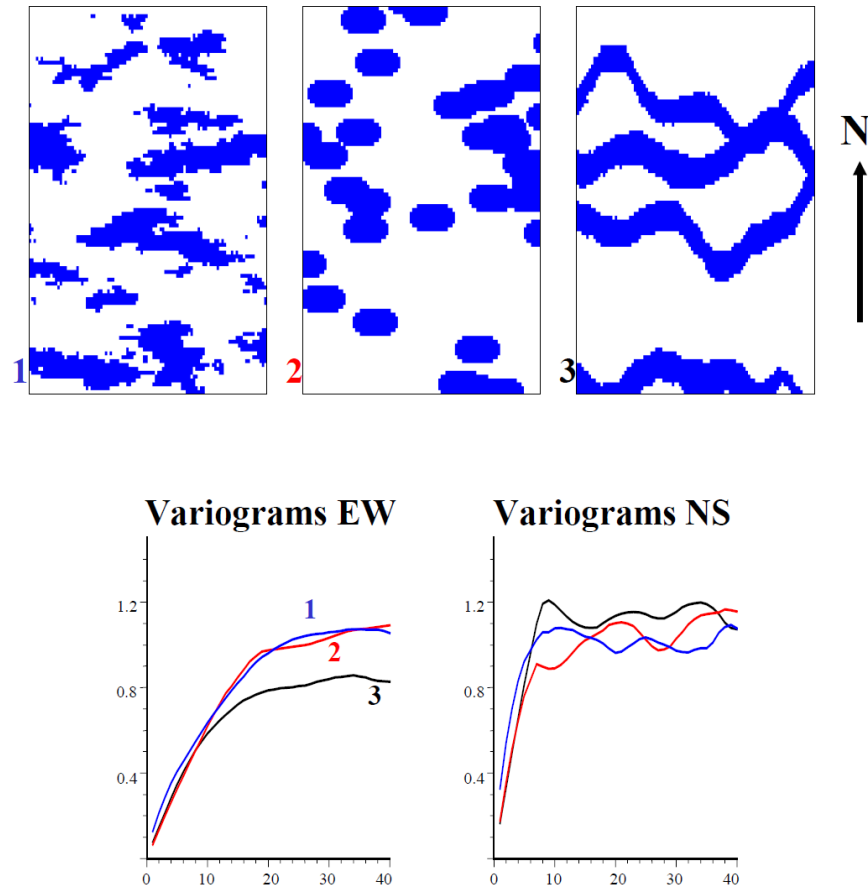


Figure 4.2: Three Different Geological Models with Approximately Same variograms (Caers and Zhang, 2004)

of pixel-based and object-based methods as it reproduces the curvilinear facies geometry and distribution which is depicted through multiple-point training images, conditioning to hard data (well data) that includes discrete facies distribution. As a substitution for the variogram, training images work as a synthetic framework of geologic body that is integrated to capture the depositional facies patterns through obtaining the local conditional probability function given in each grid cell (Caers and Zhang, 2004; Wu et al., 2007). The training images can be created in several ways such as user-defined images based on geological analogues, object-based model (Deutsch and Wang, 1996), sequence stratigraphy, satellite surface images, and process-based models (Zhang, 2008; Petrel, 2014). Meanwhile, deposits photographs or facies structure drawings cannot be considered as training images unless they meet the principles of stationarity and ergodicity (Caers and Zhang, 2004), as it will be explained in the following description.

Similar to variograms, the training images are restricted to the assumption of stationarity and ergodicity to be able to make estimations for the uncertain/random variables given in all grid blocks in a finite domain formation. Figure 4.3 shows three different types of training images: elliptical shapes, a fluvial channel, and a deltaic deposit (Caers and Zhang, 2004). It is notable that the assumption of stationarity is available only in image 2 as there is similar repetition of channel patterns everywhere in the reservoir that facilitate extraction of multiple-point statistics. However, it is clear that the 1st image does not meet the stationarity assumption where there is no consistency of pattern repetition over the image. The 3rd image is better than the 1st one to be used as training image, but the thickness of channels are not the same across the entire image. Making a statistical simulation on a finite domain leads to large fluctuations (ergodicity), and patterns will lose continuity as distance increases. Therefore, a large size training image, at least double the size of the reservoir being simulated, should be adopted for multiple-point Geostatistical modeling (Caers and Zhang, 2004).

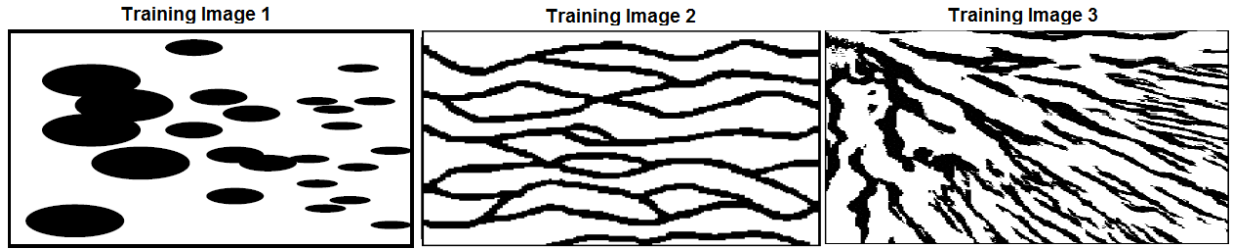


Figure 4.3: Three Possible Candidate Training Images (Caers and Zhang, 2004)

The geostatistical algorithm that can reproduce the patterns of training images conditioning to well discrete facies distribution is Single Normal Equation Simulation, SNESIM (Strebelle, 2003). SNESIM is a sequential pixel-based algorithm that recaptures the spatial facies through scanning training images over each grid block node sequentially, along a random path with well and/or seismic data to derive the facies conditional probability function (CDF). It is then stored in a template (Arpat and Caers, 2005), and all unsampled node is revisited sequentially using a random path, to assign node value (0...1) from the CDF, and simulated nodes (Caers and Zhang, 2004; Liu et al., 2004; De Vries et al. 2009). The SNESIM algorithm workflow can be described into successive main points (Arpat and Caers, 2005, Caers and Zhang, 2004):

1. Retrieving all the spatial patterns from training image through scanning it to be stored in a discrete template.
2. Smoothing or trimming of the pattern statistics should be applied on the resulting template in order to and capture important patterns and to get rid of undesirable characteristics such as infrequent fine patterns of the training image.
3. The sequential simulation consists of the following steps:
  - (a) Construct a fine 3D grid with well-data assigned to closest grid cells.
  - (b) Define a random path on the simulation grid to visit each node one at a time until each node on the random path is visited.
  - (c) Search for closest nearby well data and previously simulated cells.

- (d) Construct a probability model for the property to be simulated based on the data found in the previous step.
- (e) Draw an outcome from the probability model in the previous step and assign that value to the current grid cell.
- (f) The steps from 3-5 should be repeated considering another node of the random path until all grid nodes along the random path are visited.

Capturing all the facies structure from the training images can be performed through sampling the image using the neural networks tool after creation of a surface from the well data. Then, the surface is re-sampled into a 3D grid and the discrete template is activated through geometrical modeling . Next, the 3D pattern is created honoring the 3D grid and discrete template to be included in the final step of multiple-point geostatistics of facies simulation. In addition to the hard (well) data, the MPS can also be conditioned to the geometrical trend model, based on layer thickness or depth, or the spatial distribution of posterior facies simulation as was accomplished in this study and resulting in providing a facies channels continuity (Petrel, 2014).

### **4.3 Geostatistical Petrophysical Modeling**

Spatial petrophysical property modeling is a crucial step in reservoir characterization as it directly affects heterogeneity and flow modeling. The conditional stochastic simulation, such as sequential Gaussian simulation (SGSIM), is the most common Geostatistical algorithm for generating stochastic random fields of continuous variables such as permeability and porosity. It has been successfully adopted to create multiple equiprobable realizations of petrophysical properties distribution in order to capture the uncertainty (Deutsch, 2014). The SGSIM is more efficient than the Truncated Gaussian Simulation because it honors only values within a search neighborhood through the cumulative probability function computation. SGSIM works better especially if it is not necessary to model the orderly transition between rock types.

### 4.3.1 Sequential Gaussian Simulation

The process of conditional simulation includes transforming data into a normal distribution, computing and modeling variograms, and generating properties by sequentially sampling from conditional distributions by representing the variable at each grid as a random variable following Gaussian distribution (Deutsch and Journel, 1998; Goovaerts, 1997; Deutsch, 2014). For the first data points, the SGSIM adopts the simple kriging algorithm procedure after construction of the variogram.

However, the Sequential Gaussian Simulation then randomly selects the location of estimation based on the seed number to generate multiple equiprobable stochastic images. Also, the estimated values are used for estimations of the next missing values. As a result, the main procedure steps for the Conditional SGSIM can be summarized by the following steps (Pyrcz and Deutsch, 2014):

1. Transform all the original sampled data into Gaussian distribution with mean  $\mu = 0$  and variance  $\sigma^2 = 1$  via the normal score transformation (Z-distribution).
2. Generate the variogram for the sampled data.
3. Select randomly the location of estimation based on the seed number.
4. Estimate the variable value and the related error variance at that location by the simple kriging.
5. Create the local conditional cumulative distribution function for the variable at that location for randomly selection the values based on the that function.
6. Repeat the process for all other locations of estimation.
7. Repeat the procedure for other realizations

### 4.3.2 Model-Based Geostatistical Simulation (Bayesian Kriging)

Instead of estimating the parameters in simple kriging or universal kriging, a prior knowledge about the data distribution can be integrated into a bayesian framework in order to update the estimated data distribution which leads to efficient spatial prediction (Diggle and Lophaven, 2004). Bayesian kriging can be considered as a new form of the universal kriging, which considers a data trend in x and y directions in the calculation, but the trend in Bayesian kriging is the prior distribution of the variogram parameters such as coefficients, data variance, range, nugget, or any other geological source (Cui et al., 1995).

It is important to note that prior distribution with the estimated covariance structure results in taking into account uncertainty, is what differentiates simple kriging from bayesian kriging (Diggle and Lophaven, 2006; Diggle et al., 2007). The model's parameters in bayesian kriging is estimated from the posterior distribution based on Bayes' theory:

$$p(\theta|Z) \propto f(Z|\theta)p(\theta) \quad (4.11)$$

where  $p(Z|\theta)$  is the likelihood function and  $p(\theta)$  is the prior distribution.

For the parameter Z, the bayesian kriging model is given by:

$$(Z | \theta) \sim N(\beta, \sigma^2 C(\phi) + \tau^2 I) \quad (4.12)$$

Since the prior distribution could be based on coefficients  $\beta$ , sill(variance)  $\sigma^2$ , range  $\phi$ , or nugget  $\tau^2$ , it is represented by the following choices:

$$f(\theta) = f(\beta)f(\sigma^2)f(\phi)f(\tau^2) \quad (4.13)$$

Consequently, the posterior distribution is:

$$f(\beta|Z = z) = \int \int \int f(\beta, \theta, \sigma^2, \phi, \tau^2|z) d\sigma^2 d\phi d\tau^2 \quad (4.14)$$

$$f(\beta|Z = z) \propto f(\beta) \int \int \int f(z|\theta) f(\sigma^2) f(\phi) f(\tau^2) d\sigma^2 d\phi d\tau^2 \quad (4.15)$$

The predictive distribution  $p(x; \theta)$  of variable X at location d is:

$$f(Z(d) | X = x) = \int f(p(x; \theta)) f(\theta|x) d(\theta) \quad (4.16)$$

The classical empirical variogram, which uses a random variable  $Z(h)$  at location  $h$  that is constructed prior to fitting the bayesian kriging, is:

$$\hat{\gamma}(t) = \frac{1}{2N(t)} \sum_{h_i, h_j \notin N(t)} [Z(h_i) - Z(h_j)]^2 \quad (4.17)$$

However, the modulus variogram that was suggested by Hawkins and Cressie is (Cressie, 1993):

$$\hat{\gamma}(t) = ([\frac{1}{N(t)} \sum_{i=1}^N |Z(h_i) - Z(h_j)|^{1/2}]^4) / (0.914 + (0.988/N)) \quad (4.18)$$

And the final Model basic model that incorporates all the possible spatial (sill and range) and non-spatial effects (nugget):

$$Z(h) = \mu(h) + \omega(h) + \epsilon(h) \quad (4.19)$$

where  $\omega(h)$  refers to the effect of sill and range (spatial variation) and  $\epsilon(h)$  represents the effect of nugget (non-spatial variation).

Based on the above bayesian procedure, the semivariogram model is estimated many times after transforming the data to Gaussian distribution rather than one time like simple



kriging. The first estimated semivariogram is used for unconditional simulation of new values at each of the input data locations with back transformation. Then, a new semivariogram model is estimated from the new Gaussian simulated data in order to calculate a weight for a new semivariogram based on Bayes' rule to show how likely the observed data can be generated from the semivariogram. The procedure is repeated many times and each repetition produces a new transformation and semivariogram. The calculated weight is then used to produce predictions and prediction standard errors at the unsampled locations (Krivoruchko, 2012).

In a special case study, Bayesian Kriging was used for 2D petrophysical data distribution through R, a powerful open-source statistical language. The R implementation of Bayesian Kriging has advantages of generating multiple stochastic reservoir images with the ability to efficiently rank them for geological uncertainty assessments. In addition, Bayesian statistics of spatial property modeling can be beneficial to study the influence of heterogeneities on flow through porous media. However, Bayesian Kriging has two limitations regarding the inability to distribute petrophysical data given the lithofacies distribution and in only 2D distribution.

#### **4.4 Results and Discussion**

Among 60 wells in the main pay of the South Rumaila Oil Field, only 19 wells have the well log data that were incorporated in the Probabilistic Neural Networks and Smooth Generalized Additive Modeling to predict the multivariate Lithofacies and Petrophysical properties: permeability and porosity. Figure 4.4 shows the well locations that have well log data, permeability, porosity, and Lithofacies (Sand, Shaly Sand, and Shale).

The first step in building the 3D property model is to construct the structural model that includes grid structure and horizon modeling. The grid structure involves setting the grid system for the reservoir to be considered for all the upcoming geological and reservoir modeling. The chosen dimensions of each grid in the orthogonal corner gridding was 50 m  $\times$  50 m and the number of grids (cells) in I direction was 210, and in J direction

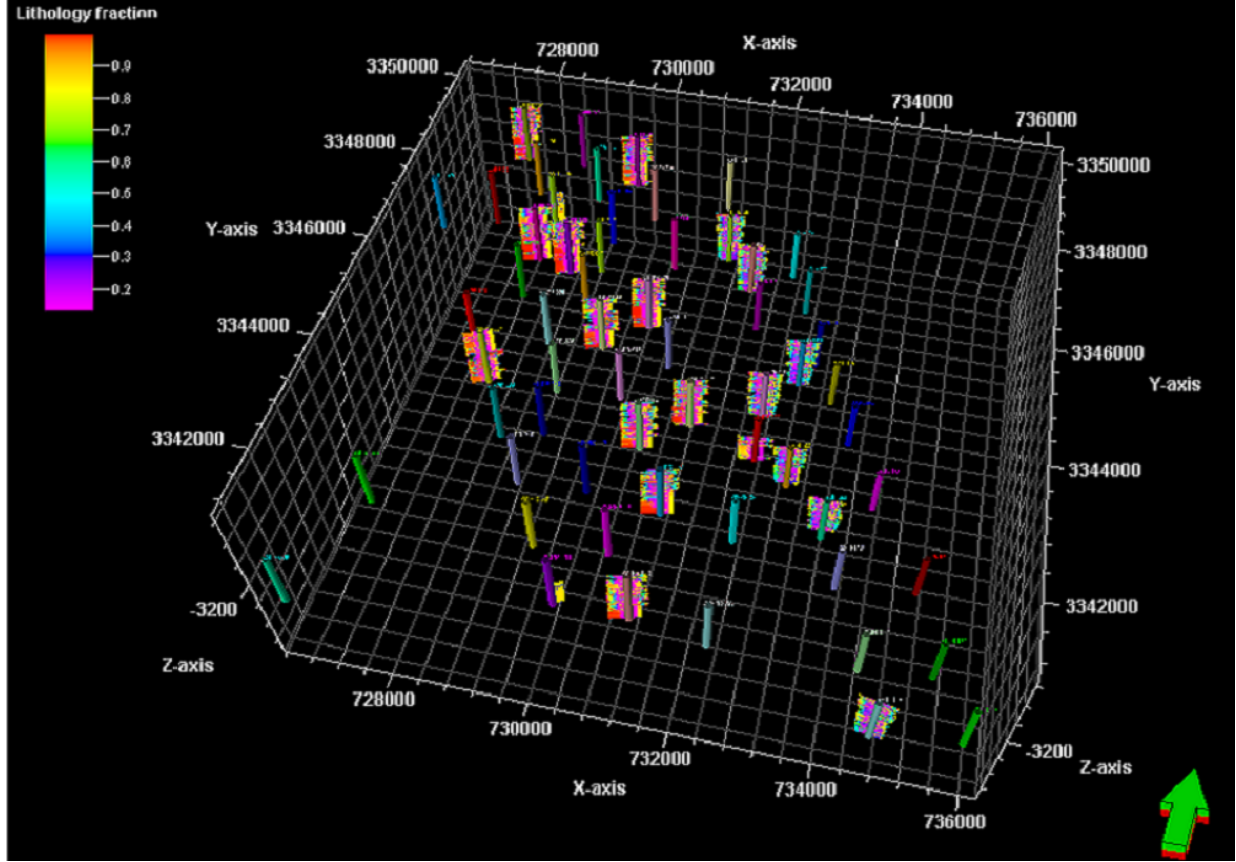


Figure 4.4: Well Locations of Available Petrophysical Data

was 202 grids. In the horizon modeling, the main reservoir has five zones. In order to capture a more realistic geological structure, the zones were subdivided into 45 layers to have approximately 2 m depth for each layer. Thus, the number of grids in K direction was 45 grids. The final Geostatistical model has 1,908,900 total number of grids for all the layers. Figure 4.5 shows the structure gridded surface for the reservoir including the layering structure.

#### 4.4.1 Geostatistical Lithofacies Modeling Results

For the Geostatistical Lithofacies modeling, the Sequential Indicator Simulation and Multiple-point geostatistics were adopted to reconstruct a 3D Lithofacies distribution for the sector of South Rumaila Oil Field/Main Pay. The well log data was upscaled, given each layer prior to starting the Geostatistical modeling for Lithofacies and Petrophysical properties.

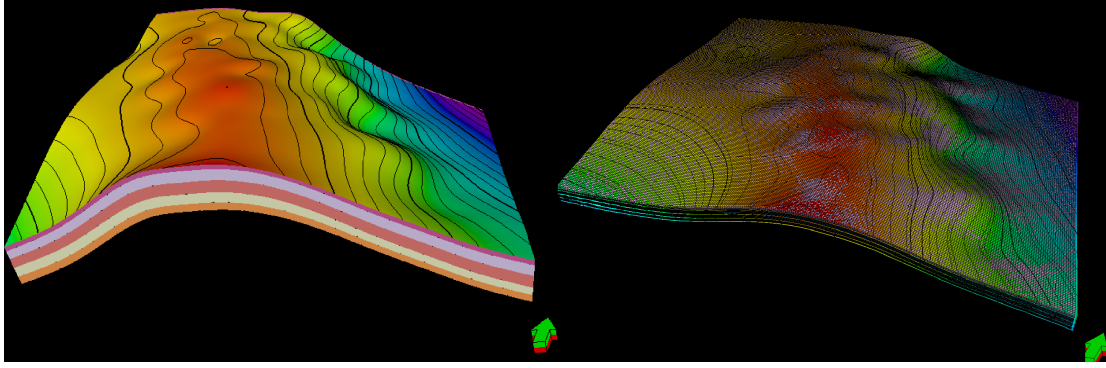


Figure 4.5: Zones Structure Gridded Surface for the Main Pay

The main steps for implementing procedure of Sequential Indicator Simulation are outlined below:

1. Upscale well log data.
2. Construct and fit indicator variogram.
3. Random seed number.
4. Frequency distribution of upscaled data points.

The Sequential Indicator Simulation considers the indicator variogram to obtain the spatial correlation of each Lithotype in four different angles:  $0^\circ$  (horizontal),  $45^\circ$ ,  $90^\circ$  (vertical), and  $135^\circ$ ). The spatial correlation of each lithotype was accomplished versus lag distance in those directions. Choosing the four main directions is to test and capture the different spatial correlations in different directions (Anisotropy). The number of lags in all the cases was 20, the approximate average range was 4000 m, and the bandwidth of the search cone was around 2000 m. These parameters were chosen to capture all possible wells with sufficient data.

After construction and fitting the empirical indicator variograms, the standard covariance parameters: sill, nugget, and range were collected to be used in 3D spatial distribution algorithm (solving the kriging equation). The spherical modeling is the best fit for all the indicator variograms for all the lithotypes in the four directions. Table 4.1 illustrates the

indicator variogram parameters for the Lithofacies Modeling in the four directions.

The constructed and modeled indicator variograms for all the facies types given each direction were depicted in Figure C1 in Appendix C.

The outputs of indicator variogram fit such as sill, nugget, and range are necessary

Table 4.1: Indicator Variogram Parameters for All the Lithofacies in the Four Direction

	Range					
Facies	Major	Minor	Range	Azimuth	Nugget	Variogram Type
Sand	1800	500	100	0	0.197	Spherical
Shaly Sand	1795	500	100	0	0.201	Spherical
Shale	1340	500	100	0	0.00	Spherical
Sand	1658	500	100	45	0.067	Spherical
Shaly Sand	1326	500	100	45	0.063	Spherical
Shale	1320	500	100	45	0.0	Spherical
Sand	1550	500	100	90	0.1	Spherical
Shaly Sand	1215	500	100	90	0.1	Spherical
Shale	2376	500	100	90	0.1	Spherical
Sand	1872	500	100	135	0.1	Spherical
Shaly Sand	1753	500	100	135	0.1	Spherical
Shale	1997	500	100	135	0.1	Spherical

for the 3D Facies modeling through the Sequential Indicator Simulation. In order to honor all the data, histogram, variogram, and correct smoothing, the Sequential Indicator Simulation was adopted for 3D Facies modeling rather than the indicator kriging that does not represent more than a facies plotting (deterministic). Figure 4.6 shows the 3D Lithofacies modeling in four different directions for the search cone which was set for the indicator variogram.

Prior to judging how accurate is the resulting 3D Facies modeling for the main pay/South Rumaila field, one should consider the depositional environment for that reservoir. The depositional environmental studies have shown that there are coarsening upward black claystones at the upper sandstone member. That is clear in the resulting Facies modeling that has shale distribution mainly in the top layers of the reservoir, and has been clearly

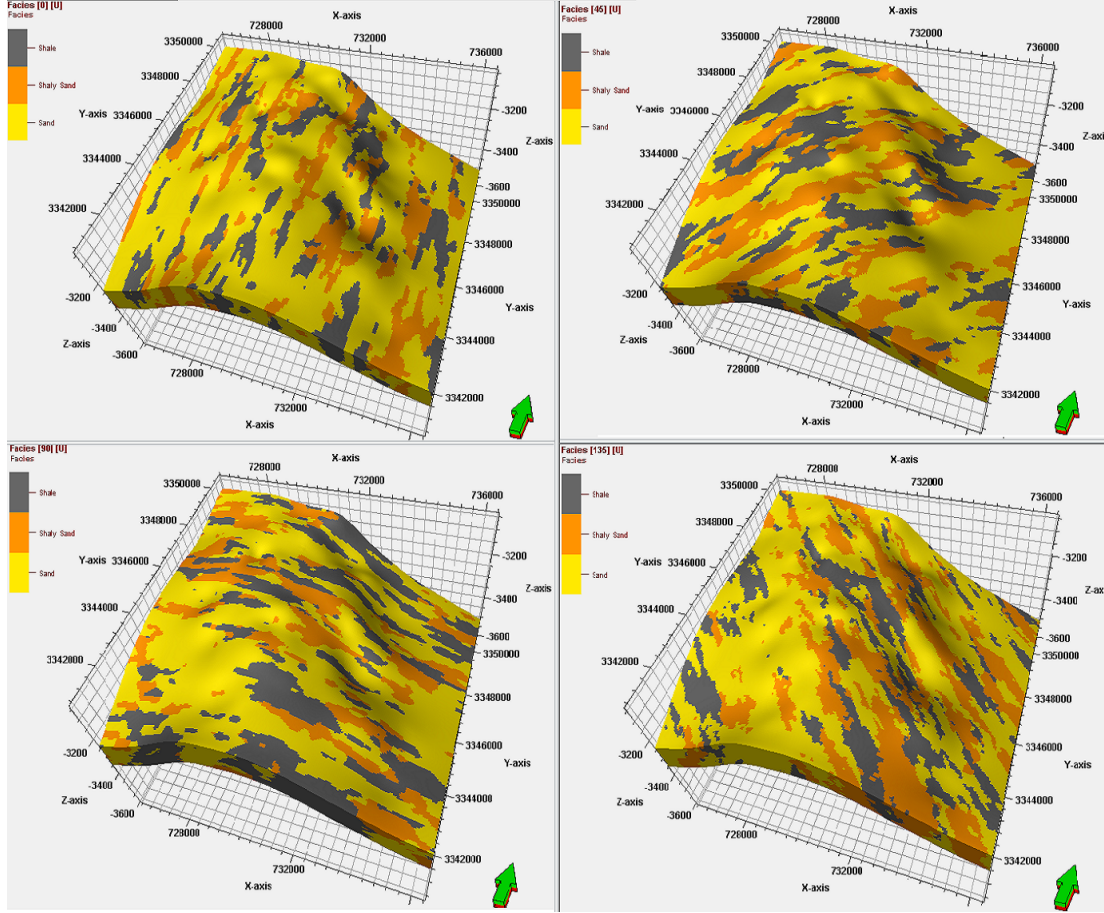


Figure 4.6: 3D Geostatistical Lithofacies Simulation in Four Different Directions

concluded from the well log and Lithofacies characterization in Appendix B.

However, the depositional environment studies have also indicated that the Main Pay was a fluvially dominated from sand-rich tidal/deltaic environment; the Sequential Indicator Simulation has not resulted in appropriate modeling for the real fluvial tidal/deltaic environment for the reservoir. As a result, multiple-point geostatistics was considered for 3D Lithofacies simulation in order to capture the most realistic depositional environment and reservoir heterogeneity. To perform multiple-point facies simulation, a geological training image should be created either by object-based model or a user-defined one that has the same characteristics of the depositional environment for the formation being studies. Since the main pay of the South Rumaila oil field exhibits a fluvial/deltaic deposition, a user-defined training image was adopted for MPS modeling. The training image was given

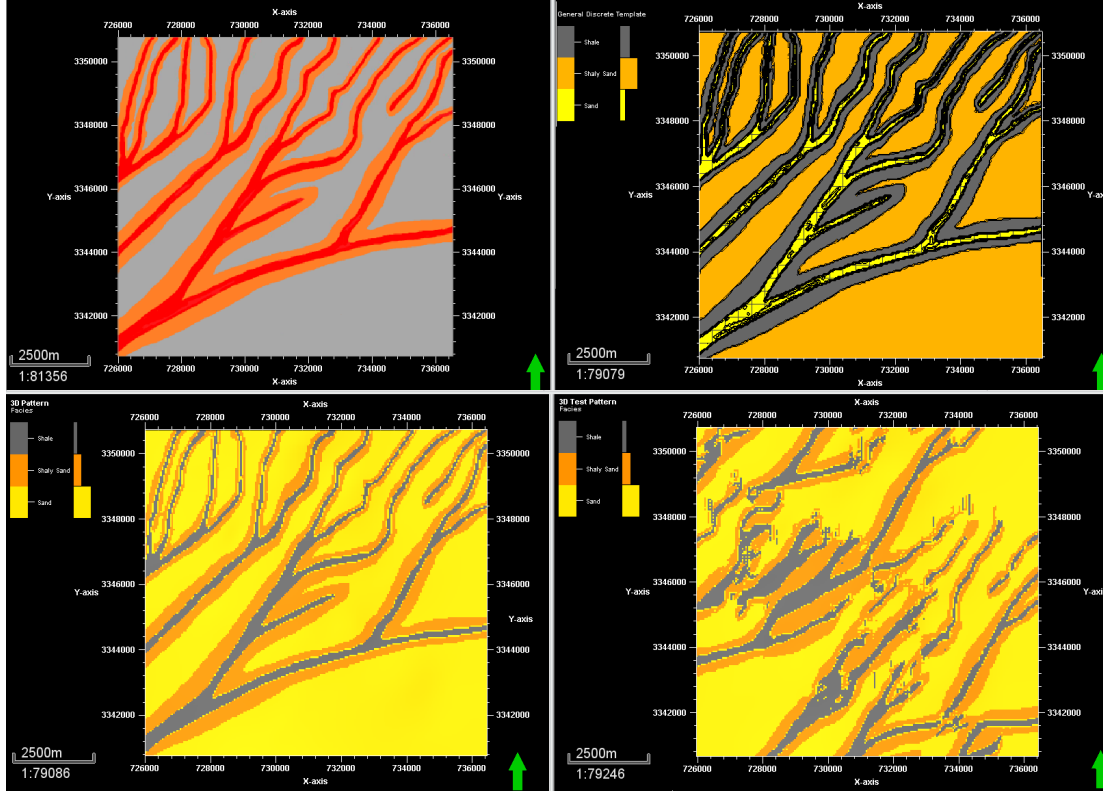


Figure 4.7: Four Main Prior Steps of Multiple-Point Facies Simulation

the coordinates of the formation in order to simulate the global reservoir locations. Next, a 2D surface was created from the training image itself for further processing. A Neural networks algorithm was used to sample the training image and generate a discrete facies template that has the original Lithofacies in the reservoir. A 3D pattern was then created, honoring the 3D grid set, and generated a discrete template. Prior to starting the multiple-point model, the 3D pattern should be tested to match the resulting pattern with the conceptual training image. Figure 4.7 shows the four main steps of maps: training image, sampling, pattern, and pattern test.

After testing the 3D pattern, a multiple-point Lithofacies simulation was adopted to reconstruct the measured spatial Lithofacies data into 3D modeling based on the conceptual geological training image. Based on the first MPS modeling trial, the original training image does not fully meet the stationarity. Therefore, 3D trend models of the three Lithotypes, Sand, Shaly Sand, and Shale, were created and incorporated in the MPS modeling.



The trend model supports the channels continuity and reduces the possible breaks. The conventional way to create lithotypes trend models is sampling using neural networks; this approach has never supported the current MPS modeling regarding channels continuity. Another way was adopted considering a sequential Gaussian simulation for the well posterior Lithofacies probability distributions, which were created from the Probabilistic Neural Networks in multivariate geostatistics. The sampled and spatial posterior Lithofacies trend were illustrated in Figure 4.8. These models support the channels continuity as shown in Figure 4.9. The model was rotated to have the model direction towards the sea shore-line at the Southern-east of the reservoir. To test whether the generated nominal MPS

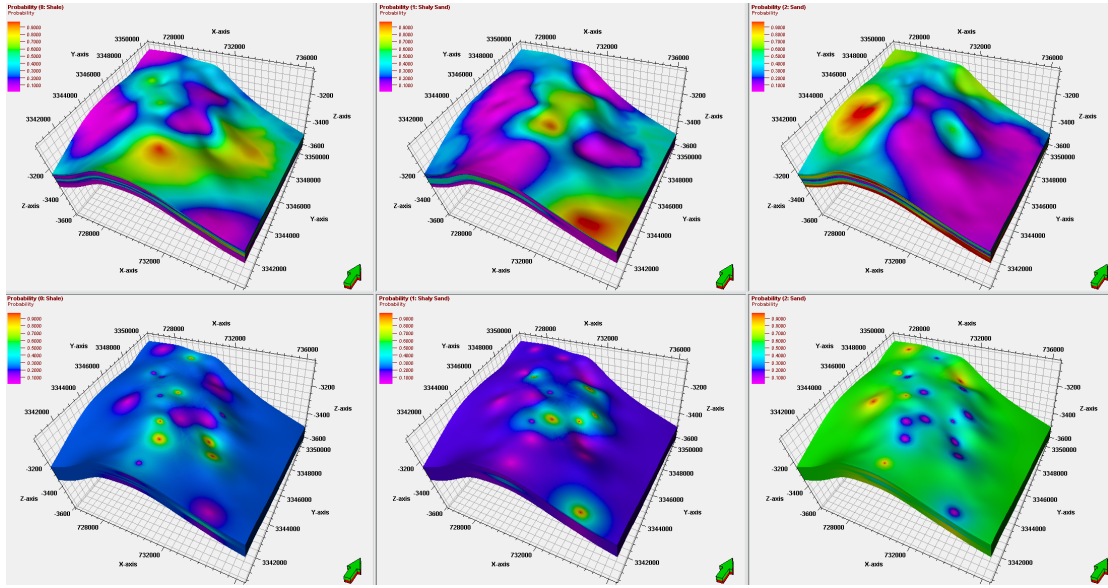


Figure 4.8: Top: Conventional Sampled Trend Models. Bottom: Spatial Posterior Lithotypes Trend Models

Lithofacies model has captured the accurate description of fluvial/deltaic structures, multiple MPS simulation runs were created considering data subdivision into two groups. The Lithofacies data are available for only 19 wells in the reservoir. There are groups of two-well pairs, located everywhere in the reservoir, mainly on the crest. The pairs were split to have two subset groups of 9 and 10 wells. Multiple realizations were generated given each group to check how the MPS model differs when the data is split and if the models still capture the depositional description. Figure 4.10 shows the two subset well data locations in the

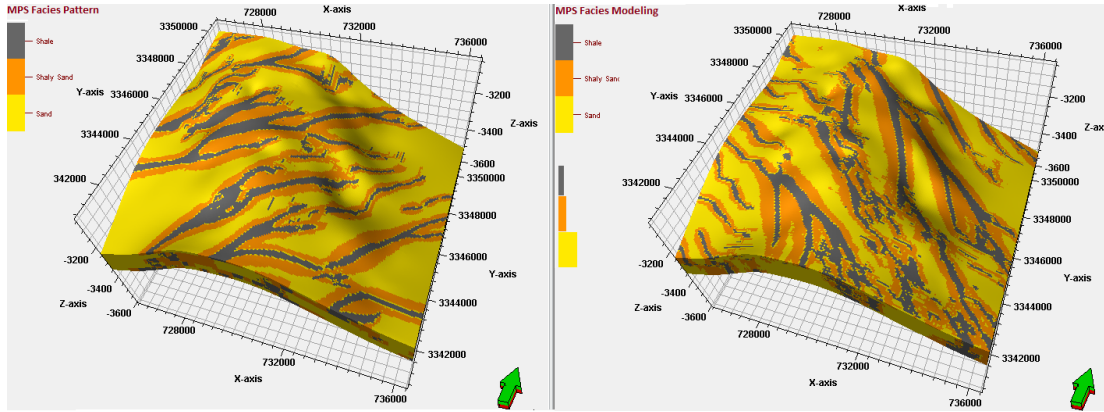


Figure 4.9: Nominal Multiple-Point Geostatistical Simulation of South Rumaila Oil Field/Main Pay

reservoir. The figures that show the multiple realizations of MPS Lithofacies models given each data group are depicted in Appendix C. There is approximate matching between the two subset models.

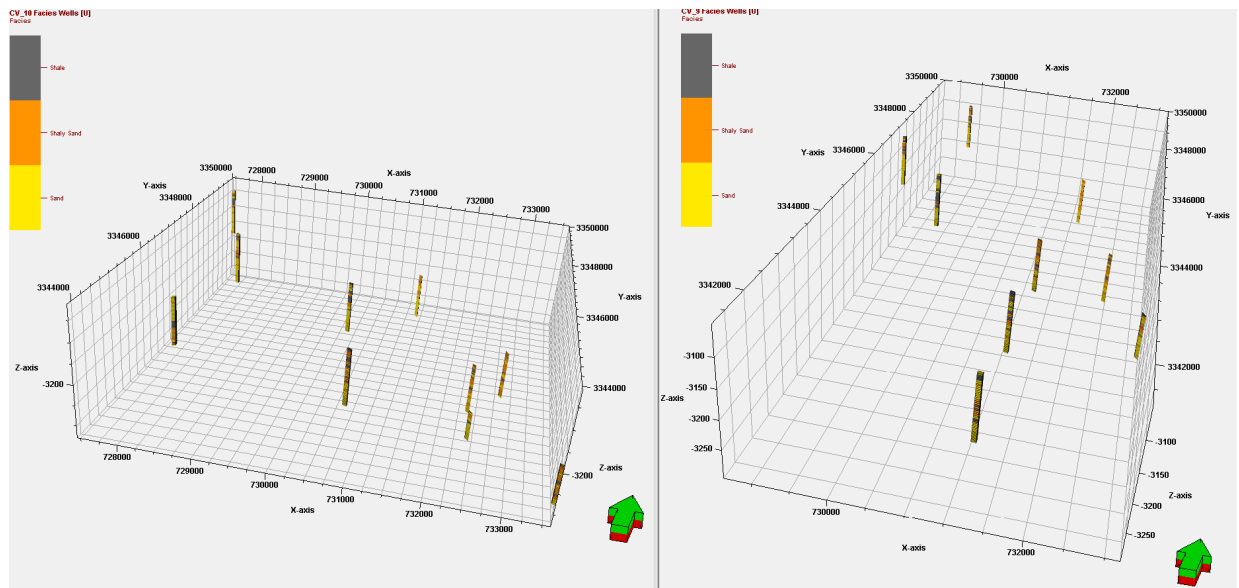


Figure 4.10: Distribution of Wells in the Two Subset Data Groups

#### 4.4.2 Geostatistical Petrophysical Modeling Results

The stochastic simulation algorithm known as Sequential Gaussian Simulation (SGSIM) is the most used algorithm for petrophysical modeling to capture the reservoir heterogeneity. It honors all the data, parameter distribution, variogram, and trends. The data should



first be upscaled in order to be used for spatial modeling. Based on the property-related histogram, the arithmetic mean was considered for well porosity upscaling, and the harmonic mean was adopted for well permeability upscaling. The harmonic mean was considered for permeability upscaling based on its distribution that seems to be highly heterogeneous according to the Heterogeneity Index or Dykstra-Parsons coefficient (Tiab and Donaldson, 2004).

Prior to using the SGSIM, the variogram for the petrophysical properties, given the three facies, was constructed and modeled. After fitting the variogram, the standard values of sill, range, and nugget, in addition to the seed number, are then honored in the SGSIM for Spatial modeling. The variograms were constructed conditioning to the three Facies (Sand, Shaly Sand, and Shale) that have been created by the sequential indicator simulation and multiple-point geostatistics. It was necessary to construct distinct variogram for each facies because the spatial distribution of the parameters is different from one lithotype to another. Therefore, 12 different variograms were constructed for each parameter to capture the variance in four different directions ( $0^\circ, 45^\circ, 90^\circ$ , and  $135^\circ$ ) given the Sand, Shaly Sand, and Shale. That procedure ensures that the property is distributed at a lithotype in a way that is different than its distribution at other facies based on the distinct variogram modeling. Tables 4.2, 4.3, and 4.4 have the variogram parameters for the log porosity, core porosity, and horizontal permeability given the three Lithofacies types; respectively. Figures 4.11, 4.12, and 4.13 represent the 3D Geostatistical reservoir characterization models for log porosity, core porosity, and horizontal permeability; respectively. In appendix C, Figures C2, C3, and C4 depict the constructed Variograms for log porosity, core porosity, and horizontal permeability. The core porosity was also modeled given given each lithotype just to provide a comparison with log porosity modeling.

The Sequential Gaussian Simulation was also adopted for 3D permeability and log porosity distribution given the multiple-point Facies simulation model. The MPS model is the model adopted into the reservoir simulation and later into the GAGD process applica-

Table 4.2: Log Porosity Variogram Parameters Given All Lithofacies in the Four Directions

	Range					
Facies	Major	Minor	Range	Azimuth	Nugget	Variogram Type
Sand	2333	500	100	0	0.1	Spherical
Shaly Sand	2014	500	100	0	0.166	Spherical
Shale	2705	500	100	0	0.115	Spherical
Sand	2162	500	100	45	0.0	Spherical
Shaly Sand	2125	500	100	45	0.0	Spherical
Shale	1876	500	100	45	0.0	Spherical
Sand	2153	500	100	90	0.0	Spherical
Shaly Sand	1631	500	100	90	0.0	Spherical
Shale	1874	500	100	90	0.178	Spherical
Sand	2756	500	100	135	0.0	Spherical
Shaly Sand	1942	500	100	135	0.0	Spherical
Shale	950	500	100	135	0.0	Spherical

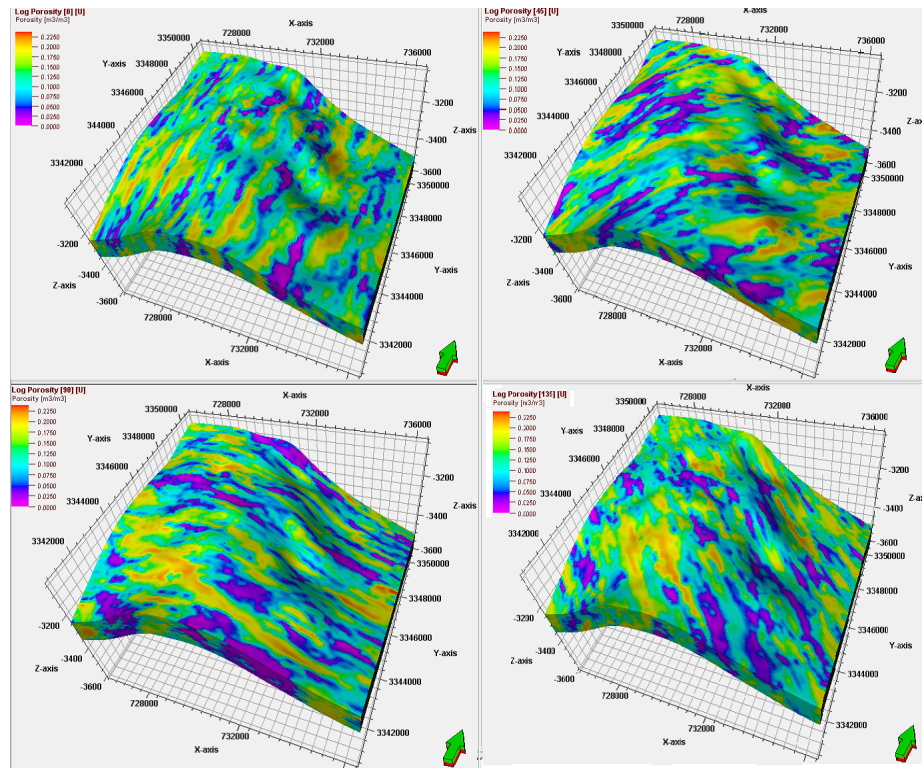


Figure 4.11: Geostatistical Log Porosity Simulation Given the Facies in Four Directions

tion. Table 4.5 illustrates the variogram fitting values for porosity, horizontal and vertical permeability in direction  $135^\circ$  only (the direction of sea shoreline). Figure 4.14 shows

Table 4.3: Core Porosity Variogram Parameters Given All Lithofacies in the Four Directions

	Range					
Facies	Major	Minor	Range	Azimuth	Nugget	Variogram Type
Sand	2619	500	100	0	0	Spherical
Shaly Sand	2624	500	100	0	0	Spherical
Shale	6511	500	100	0	0	Spherical
Sand	3144	500	100	45	0.014	Spherical
Shaly Sand	1912	500	100	45	0	Spherical
Shale	2636	500	100	45	0	Spherical
Sand	3723	500	100	90	0	Spherical
Shaly Sand	3011	500	100	90	0	Spherical
Shale	3285	500	100	90	0	Spherical
Sand	5836	500	100	135	0.019	Spherical
Shaly Sand	4735	500	100	135	0.028	Spherical
Shale	7861	500	100	135	0.0	Spherical

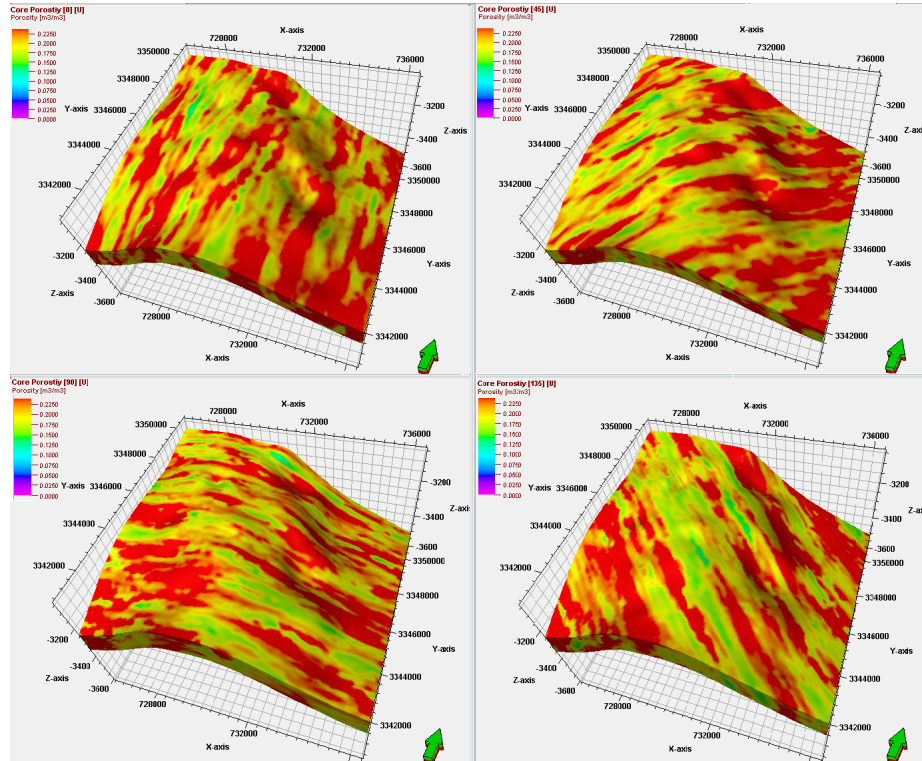


Figure 4.12: Geostatistical Core Porosity Simulation Given the Facies in Four Directions

Table 4.4: Horizontal Permeability Variogram Parameters Given All the Lithofacies in the Four Directions

	Range					
Facies	Major	Minor	Range	Azimuth	Nugget	Variogram Type
Sand	2243	500	100	0	0.001	Spherical
Shaly Sand	3318	500	100	0	0.0	Spherical
Shale	4158	500	100	0	0.0	Spherical
Sand	3389	500	100	45	0.038	Spherical
Shaly Sand	2630	500	100	45	0.0	Spherical
Shale	3811	500	100	45	0.0	Spherical
Sand	3923	500	100	90	0.007	Spherical
Shaly Sand	3182	500	100	90	0.0	Spherical
Shale	2857	500	100	90	0.0	Spherical
Sand	7222	500	100	135	0.0	Spherical
Shaly Sand	5289	500	100	135	0.01	Spherical
Shale	5119	500	100	135	0.0	Spherical

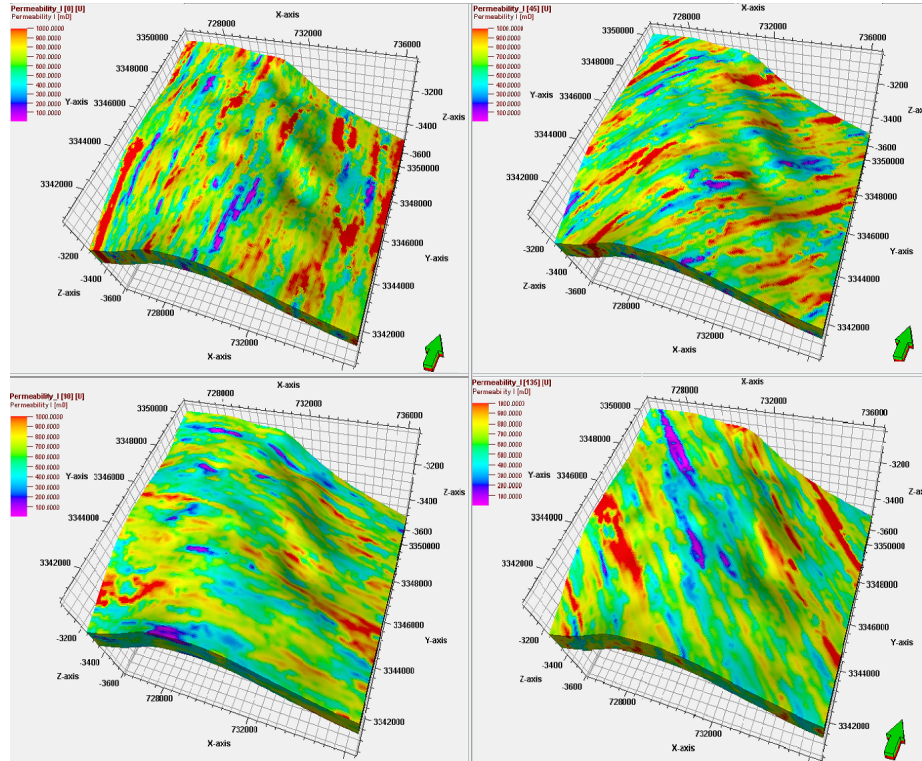


Figure 4.13: Geostatistical Permeability Simulation Given the Facies in Four Directions

the 3D modeling of log porosity, horizontal and vertical permeability given the MPS facies simulation. The variogram structures for log porosity, horizontal and vertical permeability given the MPS facies simulation are illustrated in Appendix C (Figures C5, C6, and C7).

In addition to Sequential Gaussian Simulation, bayesian kriging was also adopted for

Table 4.5: Porosity, Horizontal and Vertical Permeability Variogram Parameters Given the Lithotypes in MPS Facies Model

Property	Range					
	Facies	Major	Minor	Vertical	Nugget	Variogram Type
Porosity	Sand	2179	2392	100	0.1516	Spherical
	Shaly Sand	1999	2051	100	0.1556	Spherical
	Shale	4037	3845	100	0.1419	Spherical
Permeability-I	Sand	3817	4162	100	0.2442	Spherical
	Shaly Sand	4767	2276	100	0.7142	Spherical
	Shale	1957	8142	100	0.2965	Spherical
Permeability-K	Sand	4000	1885	100	0.2428	Spherical
	Shaly Sand	3968	24618	100	0.6331	Spherical
	Shale	1731	3530	100	0.3146	Spherical

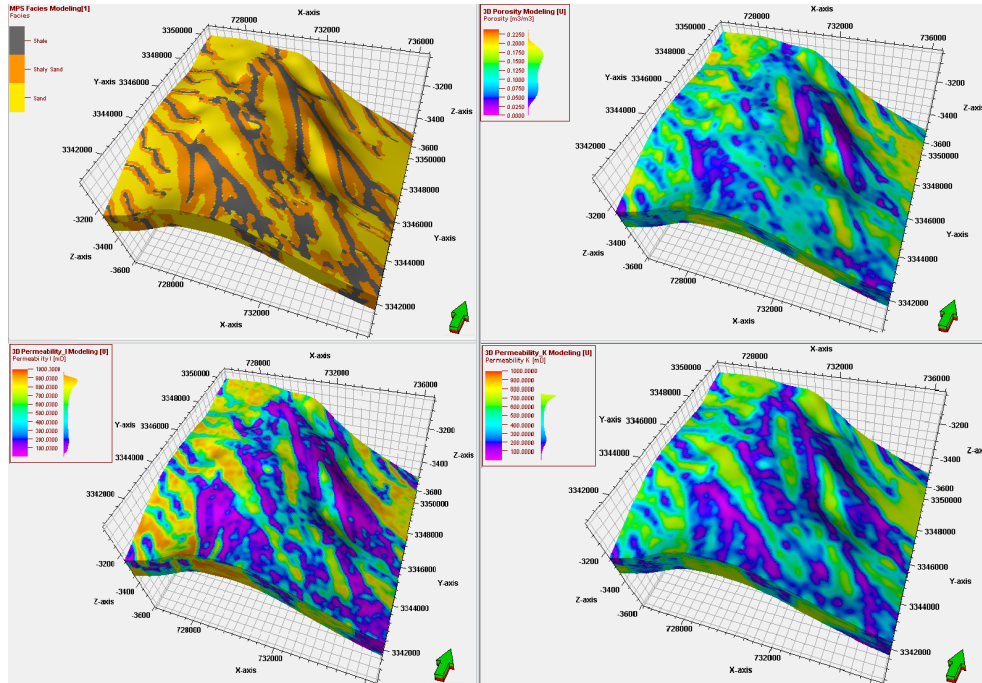


Figure 4.14: 3D Petrophysical and Lithofacies Modeling given MPS Simulation



2D permeability modeling only for one layer (1st layer). The same data that was used for Sequential Gaussian Simulation, was employed in bayesian kriging algorithm without conditioned to the different Lithofacies. That is another limitation of Bayesian kriging implementation in R in addition to the 2D modeling only.

The bayesian kriging procedure also includes creating many equiprobable reservoir stochastic realizations as well as ranking them based on the sampling techniques to select the three quartiles ( $P_{10}$ ,  $P_{50}$ , and  $P_{90}$ ). This example was performed through the parallel programming in R statistical language using *Rmpi* package. *Rmpi* is the Message Private Interface (MPI) for R parallel programming . In *Rmpi* package, the code distributes the simulation jobs (realizations) on multiple CPUs of one machine for more quick results.

Figure 4.15 decorates the different variogram models that were fitted for permeability data. The spherical model has the best fit with the permeability data and as a result it was adopted in spatial modeling. Figure 4.16 shows the Posterior Distributions of

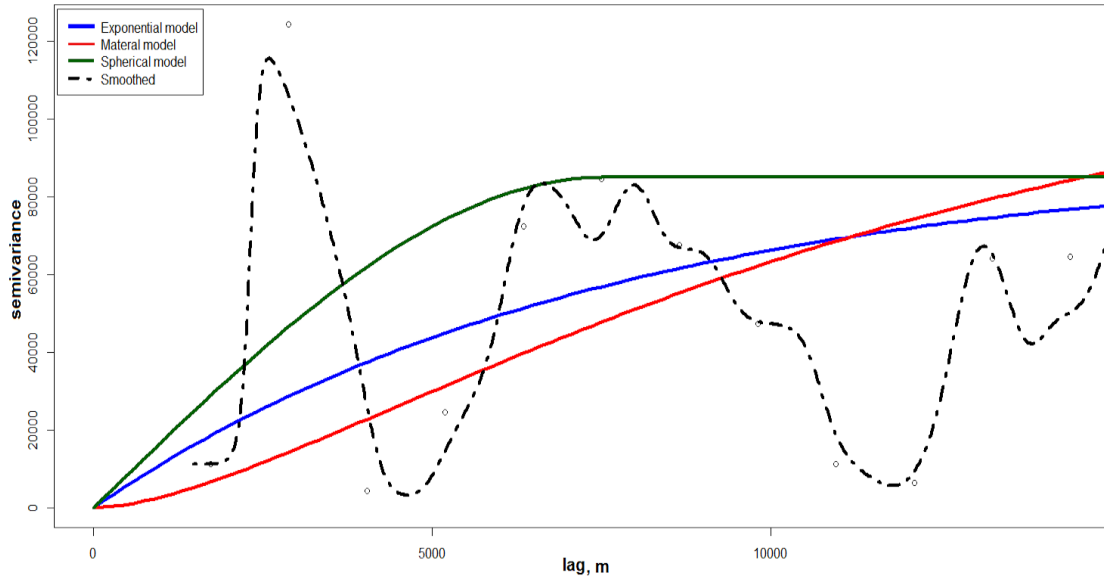


Figure 4.15: Various Variogram Fitting for the Permeability Data

Model's Parameters (coefficient ( $\beta$ ), sill ( $\sigma^2$ ), and range ( $\phi$ ) for the permeability modeling. Figure 4.17 depicts the sample of data prior and posterior distributions in bayesian kriging. Figure 4.18 illustrates the 2D Plots of the nine quantiles of permeability simulation

( $P_{10}, P_{20}, \dots, P_{90}$ ) from the predictive distribution. Figure 4.19 decorates the histograms of the nine quantiles of permeability simulation. Figure 4.20 illustrates 2D and histogram plots of the sampling-based ranked realizations  $P_{10}, P_{50}, \text{and } P_{90}$  of permeability simulation from the predictive distribution.

The Leave-One-Out cross validation of bayesian kriging procedure for the permeability distribution are illustrated in Figure 4.21 that shows the variance spatial map and histogram of standard error. There are two points have high variance values as they are located in shale regions. The bayesian kriging algorithm smoothly simulates data without conditioned to the different facies.

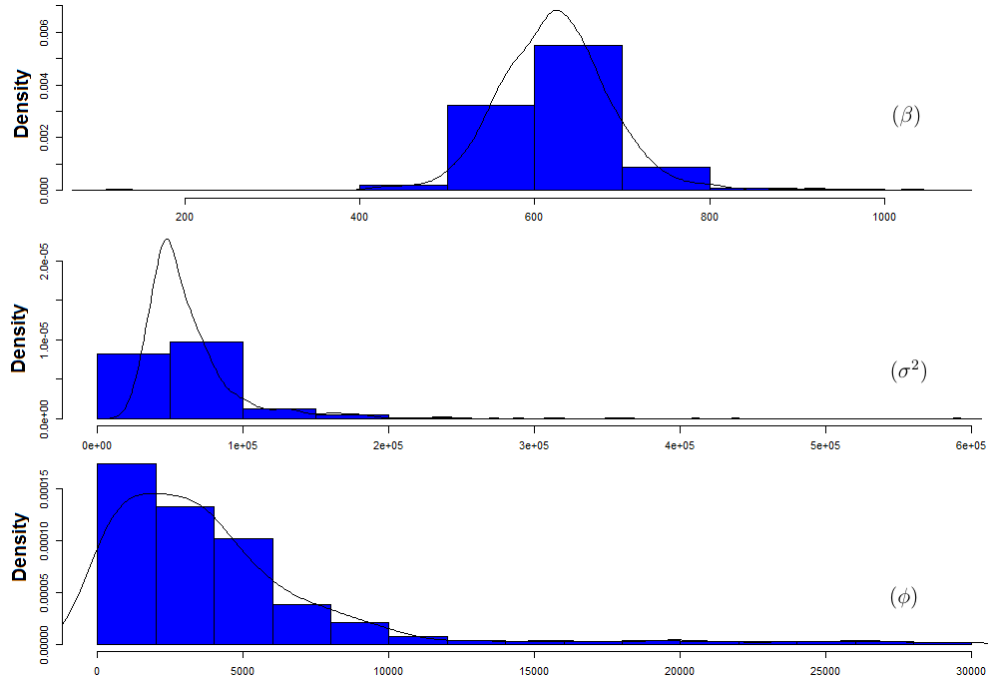


Figure 4.16: Posterior Distributions of Model's Parameters Coefficients

## 4.5 Cross-Validation Results

To check the validity of Sequential Gaussian Simulation for petrophysical property modeling, a Leave-One-Out (L-O-O) cross-validation should be carried out. Doing Geo-statistical modeling in Petrel leads to preserving of the actual data given their locations. So there is no mismatch between measured and spatial predicted values. Figure 4.22: a cross-plot showing the strong correlation between predicted and measured of permeability

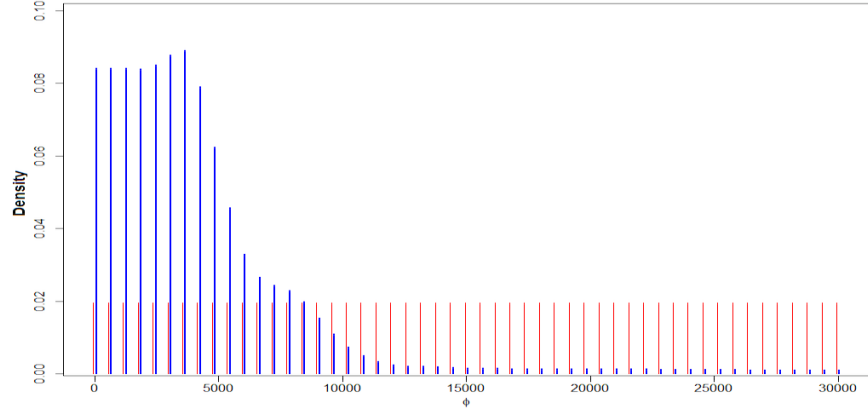


Figure 4.17: Prior and Posterior Distributions Sampling in Bayesian Kriging

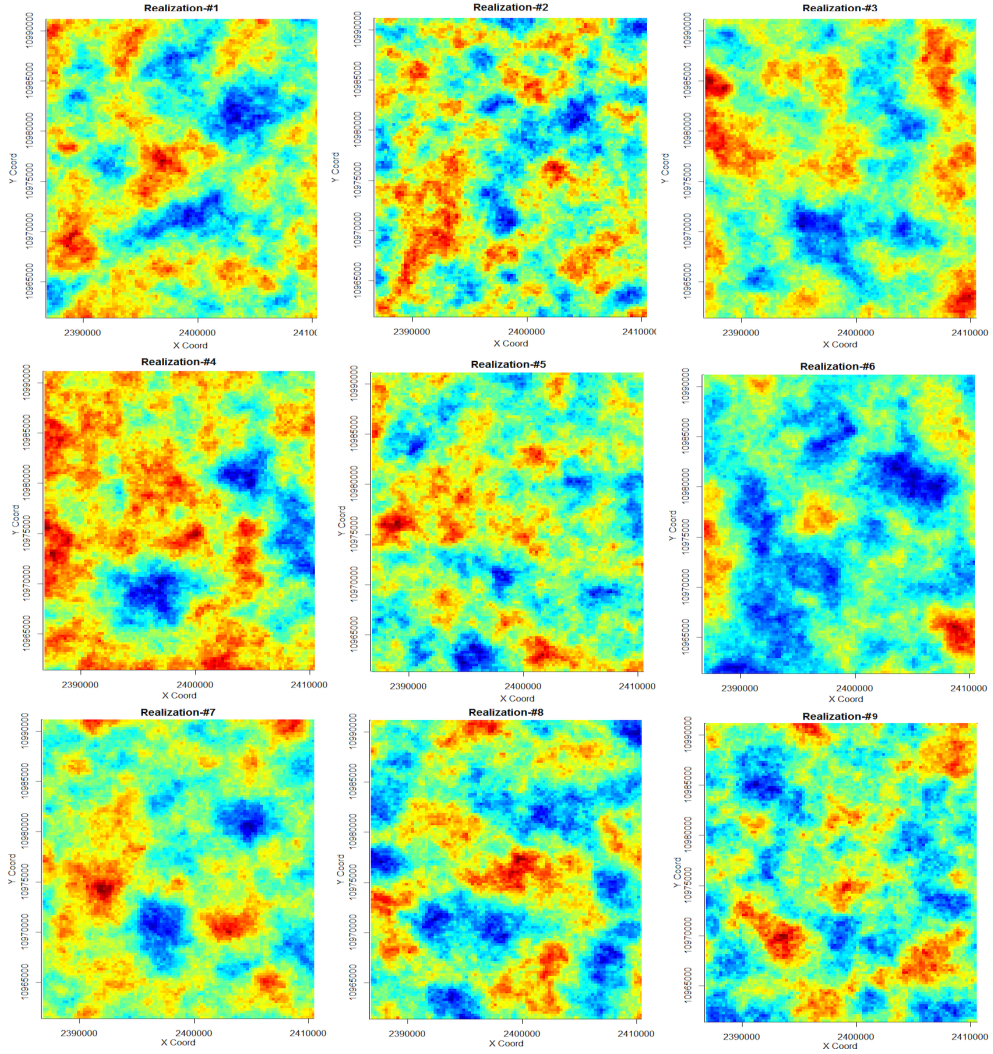


Figure 4.18: 2D Distribution of the Nine Quantiles of the Permeability Simulation



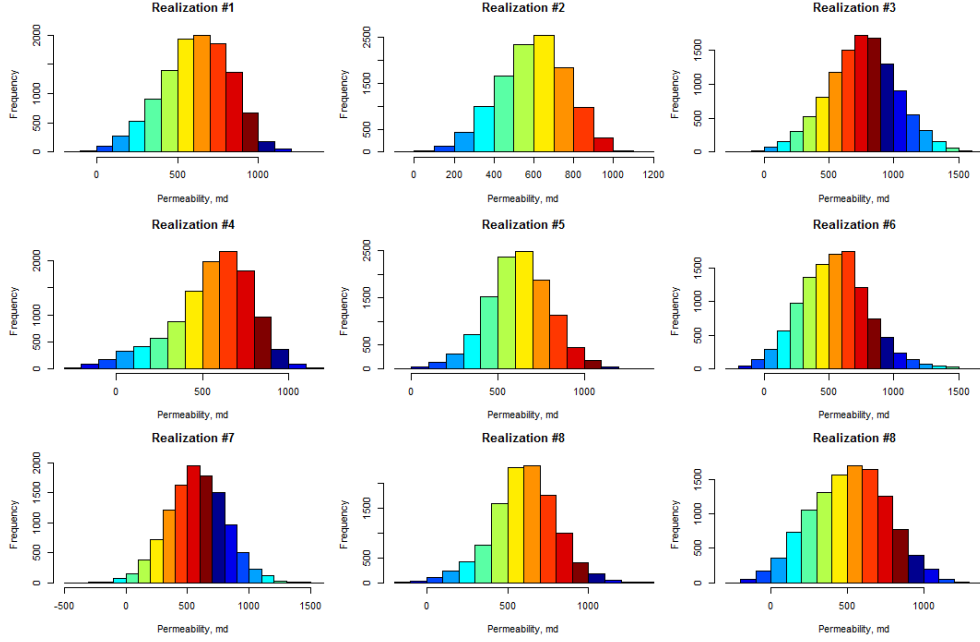


Figure 4.19: Histograms of the Nine Quantiles of the Permeability Simulation

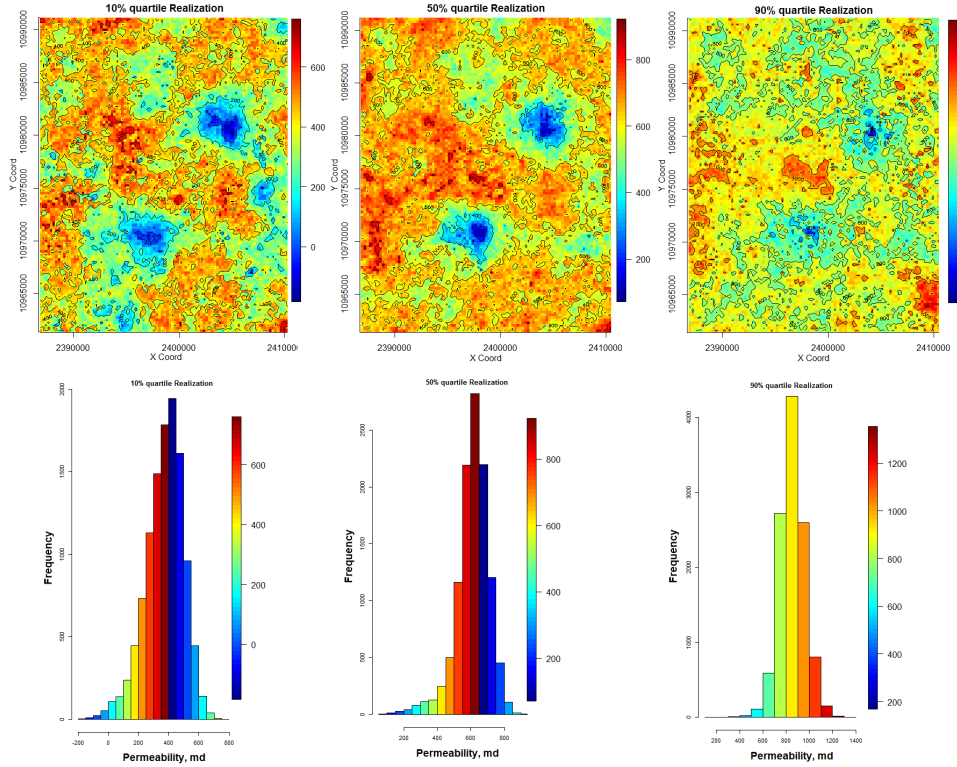


Figure 4.20: 2D and Histogram Plots of the Ranked Realizations  $P_{10}$ ,  $P_{50}$ , and  $P_{90}$  of Permeability Distribution.

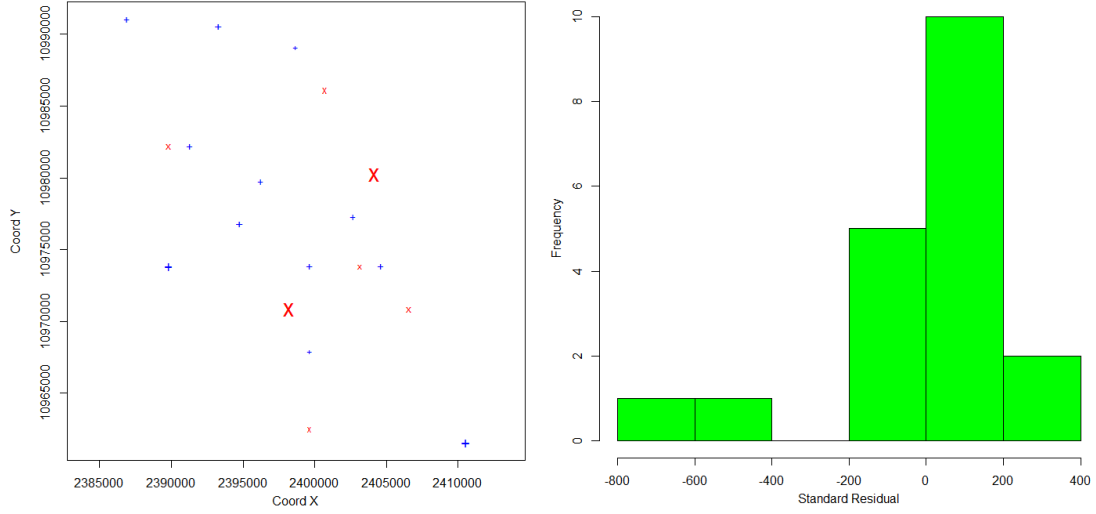


Figure 4.21: Leave-One-Out Cross Validation of Permeability Modeling-Bayesian Kriging and porosity by Sequential Gaussian Simulation given the MPS Facies model.

Figure 4.23 shows the cross-plot between porosity and permeability given the MPS Litho-

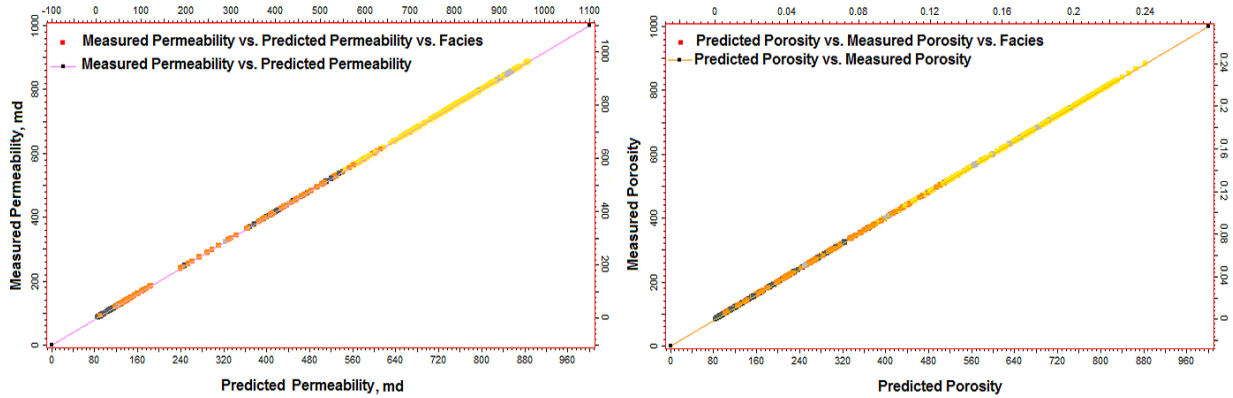


Figure 4.22: Predicted-Measured Permeability and Porosity Given Lithofacies Cross-Plots

facies model. Since we have data available only for 19 wells, so 19 subset data was upscaled for 3D geostatistical modeling for permeability and porosity. Then, a comparison was achieved between the measured property data including all wells and the modeled without one well. The Leave-One-Out procedure was also applied on the Lithofacies modeling. The figures of the entire procedure was illustrated in Appendix C from Figure C10 to Figure C19.

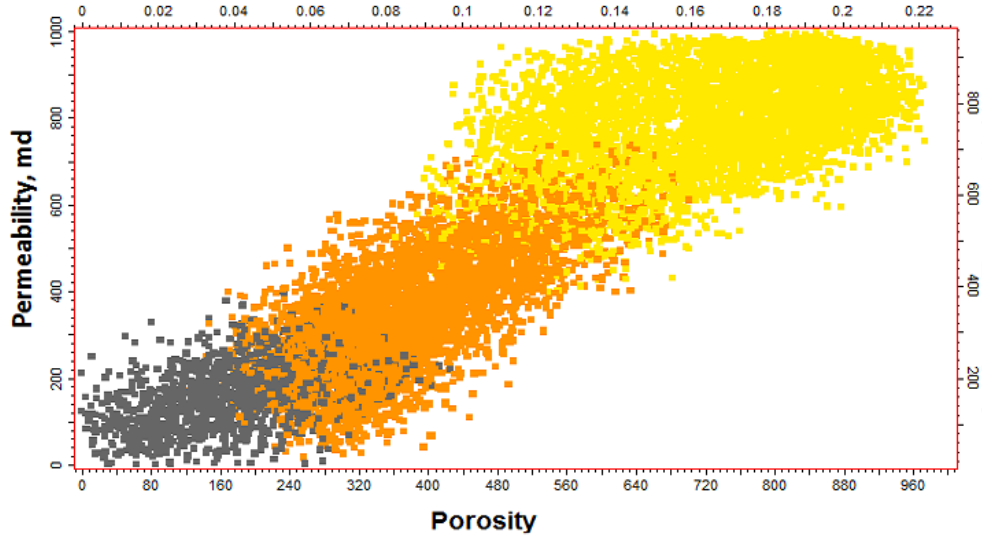


Figure 4.23: Cross-Plot between Porosity and Permeability Given Lithofacies

#### 4.6 Realizations Ranking

To capture geological uncertainties, 100 of stochastic equiprobable reservoir models(realizations) were created honoring geological constraints for permeability and porosity through sequential gaussian simulation. Since it is impractical to simulate all these realizations in the reservoir study, ranking was applied considering cumulative oil production for permeability and initial oil in place for porosity to determine the P10, P20,..., and P90 that represent the overall reservoir uncertainty. All these permeability property realizations were evaluated in the compositional reservoir model for the ranking process with short production period. Figure 4.24 illustrates the histogram of the created 100 stochastic porosity realizations based on the calculated initial oil in place (IOIP). Additionally, the figure shows the selected nine quantiles(P10, P20,..., and P90) that were sampled based on the cumulative probability function. The histogram of the ranked nine sampled of the reservoir porosity realizations was illuminated in Figure 4.25. Furthermore, the 3D distribution images of the nine reservoir porosity ranks were depicted in Figure 4.26. In this figure, there is significant difference between the nine realizations that reflects the overall uncertainty range. As previously stated, the same procedure was conducted to rank the

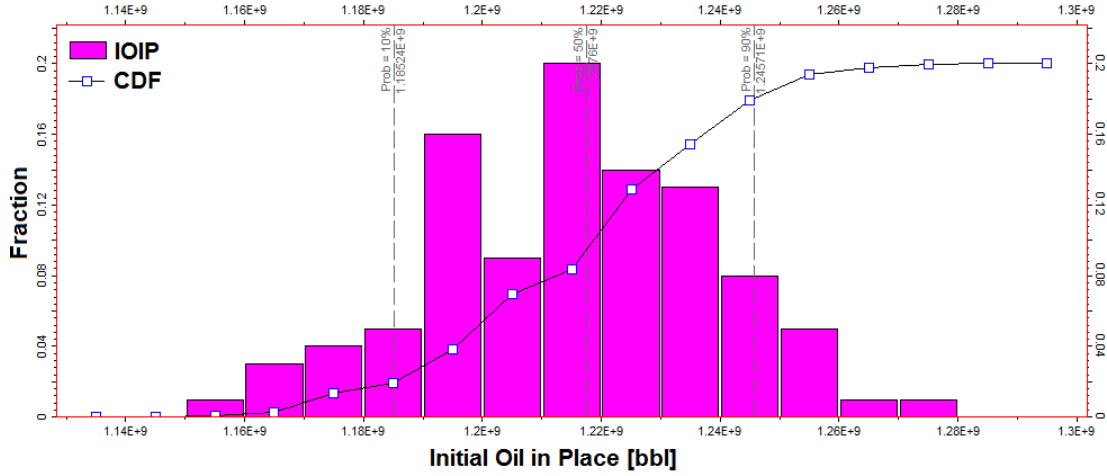


Figure 4.24: Histogram of the 100 Reservoir Realizations with P10, P20,..., and P90

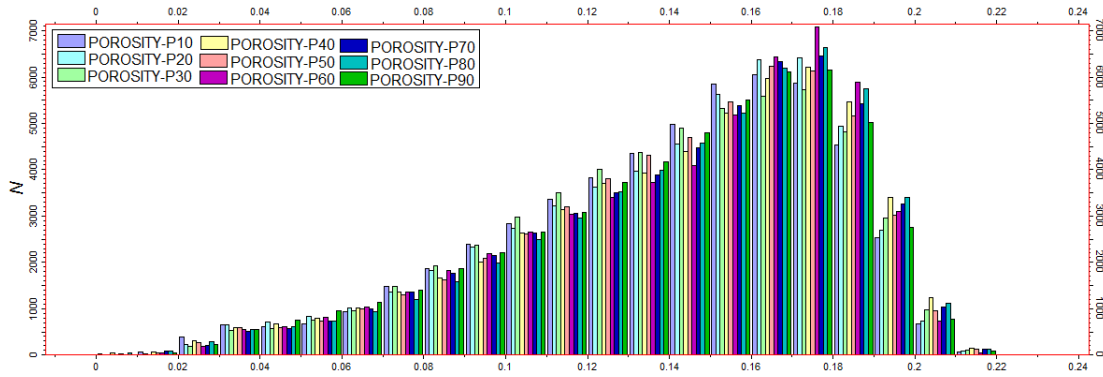


Figure 4.25: Histogram of the P10, P20,..., and P90 Reservoir Porosity Realizations

reservoir permeability stochastic models based on calculating cumulative oil production for a specific period. The histogram of the ranked nine sampled of the reservoir permeability realizations was depicted in Figure 4.27.

Moreover, the 3D distribution images of the nine reservoir permeability ranks were depicted in Figure 4.28, which shows the dissimilarity between these stochastic images with regard to the range of porosity values levels.

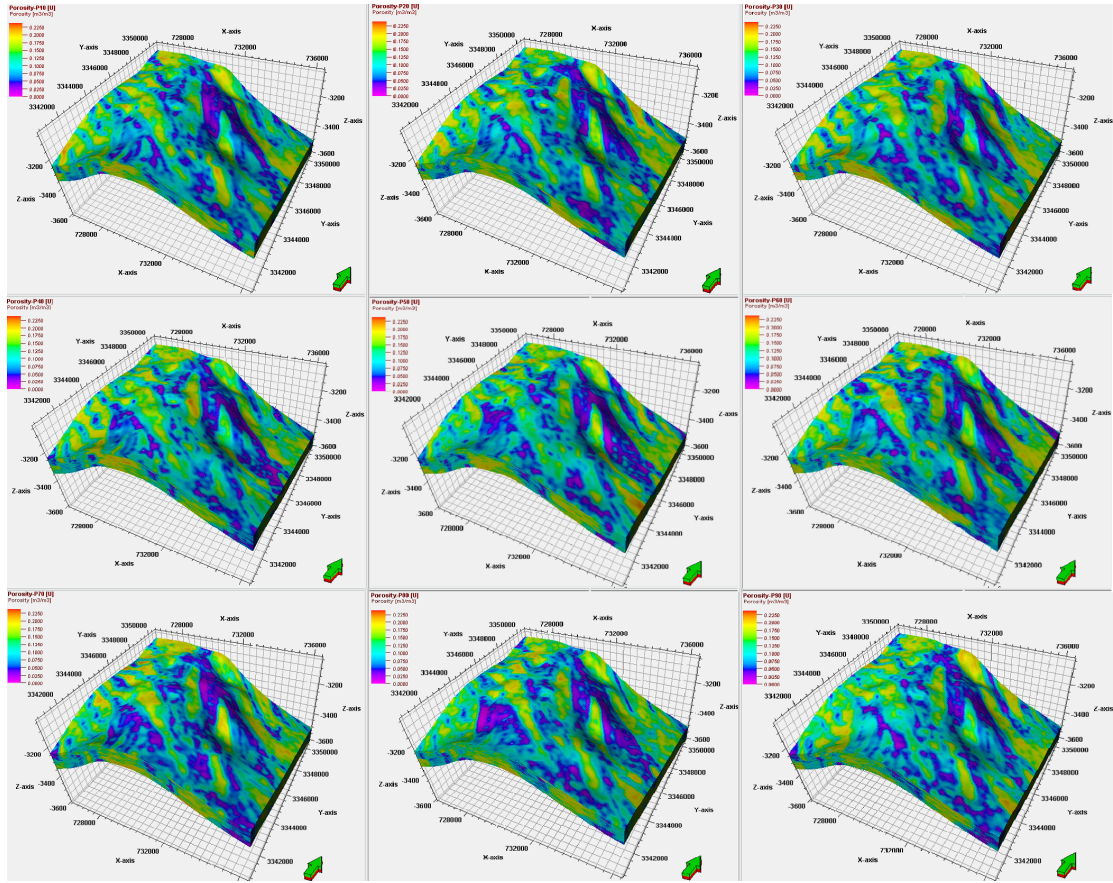


Figure 4.26: 3D Distribution Images of the Ranked Porosity Stochastic Models

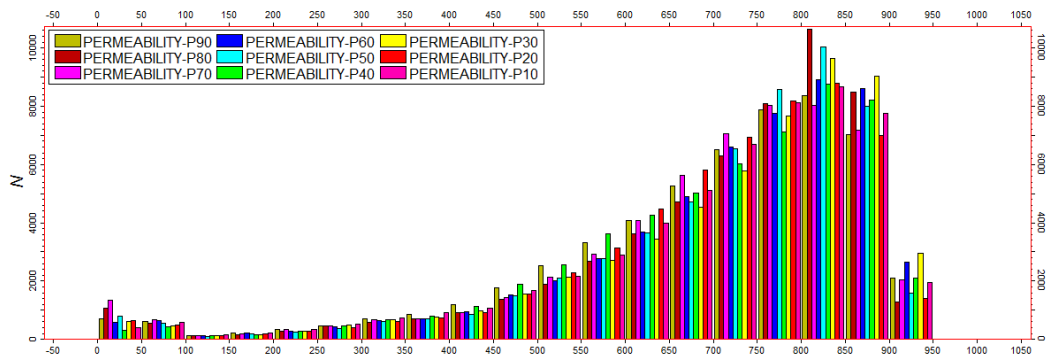


Figure 4.27: Histogram of the P10, P20,..., and P90 Reservoir Permeability Realizations

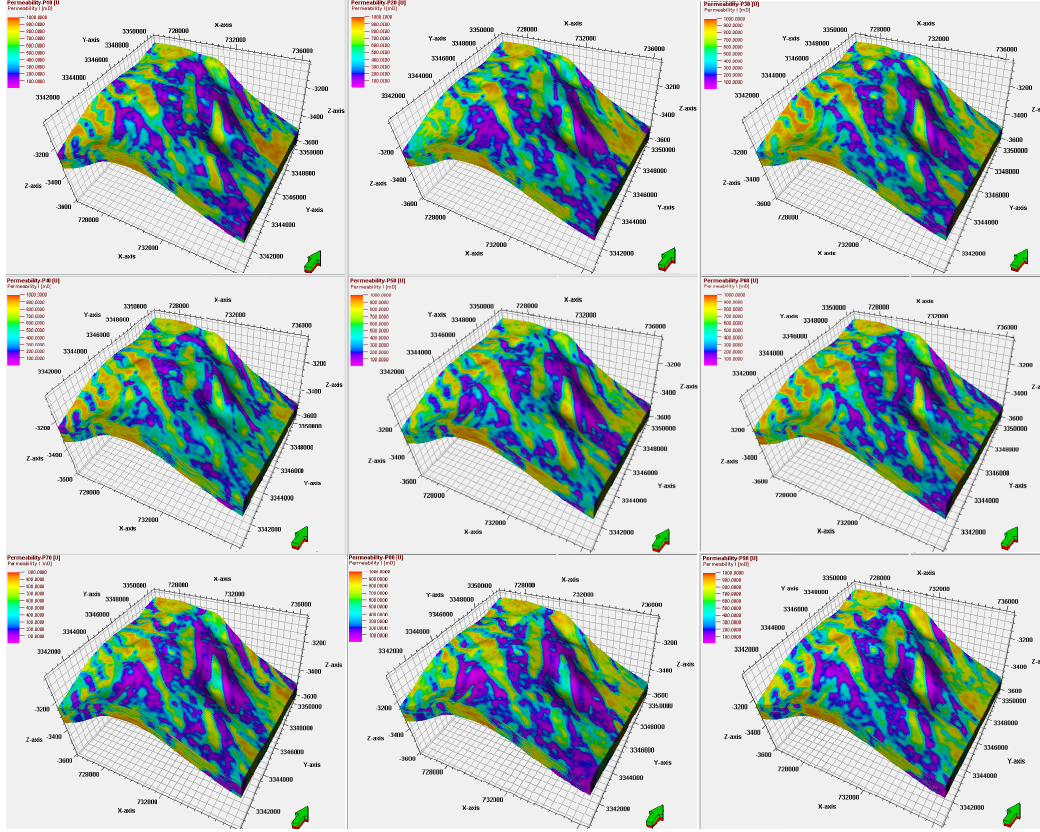


Figure 4.28: 3D Distribution Images of the Ranked Stochastic Permeability Models

Only the three quantiles (P10, P20, and P90) were employed in the robust design of Huff-and-Puff GAGD process optimization, as will be explained in Chapter 7. However, the nine quantiles (P10, P20,..., and P90) were all incorporated in the geological uncertainty assessment, as will be illustrated in Chapter 8.

#### 4.7 Summary and Significance

To capture the most realistic geological model, a comparative geostatistical lithofacies models were conducted through sequential indicator simulation (SISIM) and multiple-point geostatistics (MPS) for 3D lithofacies modeling. The two algorithms were implemented to capture the fluvial deposition environment of the main pay/South Rumaila oil field. The multiple-point geostatistics led to create more plausible fluvial lithofacies model than sequential indicator simulation. Since MPS considers training image as a substitution for the variogram, the resulting 3D lithofacies model corresponds to the deposition environment

that has been described in the literature.

The MPS-based lithofacies modeling was then considered for petrophysical property modeling. The sequential gaussian simulation (SGSIM) was adopted for permeability and porosity distribution given each lithofacies. To validate the resulting SGSIM models, the cross-plots and Leave-One-Out cross-validation showed excellent matchings between predicted and measured petrophysical properties.

The validated models were then upscaled for compositional reservoir simulation of the GAGD process. Multiple reservoir stochastic images (realizations) were created and ranked to capture the geological uncertainty space.



# 5. Integrated Reservoir Modeling

## 5.1 Introduction

In petroleum reservoir simulation, mathematical and numerical techniques are combined to solve governing differential equations in order to build a model that studies the behavior of fluid flow through porous media (Fanchi, 1997). The reservoir simulation helps to understand and validate the reservoir characteristics through integrating production data to improve reservoir characterization. Without a validated model, reservoir management objectives cannot be achieved to monitor different production activities during the field life. The main objective of reservoir simulation is to predict the future field performance for decision making. The future performance includes many optimization schemes such as infill drilling, waterflooding, and production control that are necessary to improve oil recovery (Aziz and Settari, 1979). In this study, the compositional reservoir simulation was adopted to evaluate the reservoir performance through the Gas Assisted Gravity Drainage (GAGD) process in South Rumaila oil field. The South Rumaila oil field started producing in 1954 and the water injection processes were initiated in 1980's to support the pressure from the east flank as the aquifer strength from the west flank is 20 times higher its strength from the east one. Both aquifer and water injection have a combined effect on reducing the production rates in the last decade with high ranges of water cut. In addition, the current recovery factor in the sector under study is 68% and there is more than 2 billions barrels of oil left in the reservoir. Consequently, the GAGD process was suggested as an efficient process to reduce water cut and enhance oil recovery.



## 5.2 Workflow of Reservoir Simulation Study

A typical workflow used to build a 3D general reservoir model encompasses the following steps (Aziz and Settari, 1979):

1. Define the reservoir engineering problem with possible economic benefits.
2. Gather all static and dynamic data required to conduct a reservoir simulation study.
3. Interpret and integrate the geological, geophysical, and petrophysical data into geological modeling.
4. Analyze the PVT behavior of fluids involved.
5. Select the type of simulator to be used to solve the defined problem.
6. Set the grid system and time-step sizes.
7. Adjust the reservoir model parameters for sensitivity analysis and history matching.
8. Conduct a future reservoir performance prediction process through different optimization techniques

The steps are followed in this study as presented below.

### 5.2.1 Defining the Simulation Problem

The reservoir simulation is an effective approach to evaluate the performance of carbon dioxide (CO<sub>2</sub>) injection through the Gas Assisted Gravity Drainage (GAGD) process. After obtaining an acceptable history matching, the GAGD process was simulated on a sector in the main pay of South Rumaila oil field, which was fully described in Chapter 2. In the GAGD process, a compositional reservoir simulator was constructed to model the CO<sub>2</sub> injection through 20 vertical injectors, which are located at the top two layers of the reservoir. Additionally, 11 horizontal producers were placed at the middle layers (5th, 6th, 7th, and 8th) for oil production. The effectiveness of CO<sub>2</sub> injection through the GAGD process was evaluated for 10 year of the prediction period (from 2016 to 2026).

### 5.2.2 Static and Dynamic Reservoir Data

In addition to geological and petrophysical data, there are important dynamic data that need to be incorporated to build a complete reservoir simulation model. The data includes well production/injection history. Also, wells perforation intervals and completion status history are required. PVT data such as solution gas oil ratio, formation volume factor, fluid viscosity in addition to gas and oil component and compositional analyses are also necessary for reservoir modeling of tertiary oil recovery operations. Capillary pressure and relative permeability curves are needed for all rock types in the reservoir.

For South Rumaila oil field, the physical and thermodynamic properties of the reservoir fluids have been obtained from the results of analyzing large numbers of oil samples taken from many wells for more than 50 years. The main pay in South Rumaila oil field is an undersaturated water-oil reservoir as the initial and current reservoir pressure values are greater than the unique bubble point pressure, which exists throughout the reservoir. More specifically, bubble point pressure is 2660 psi while the initial reservoir pressure of 5186 psi measured at 3154.7 m datum. In addition, original oil-water contact (OWC) is 3269 m. Table 5.1 shows the components and phase properties for South Rumaila oil field/main pay.

Table 5.1: Component and Fluid properties of main pay/South Rumaila Oil Field

Phase Property	Oil	Gas
Density, $lb/ft^3$	72.79	51.5
Compressibility, $psi^{-1}$	2.7E-6	9.1E-6
Formation Volume Factor, vol/st.vol	1.12	1.025
Viscosity, cp	0.65	0.52
Reservoir Temperature, F	210	
Solution Gas-Oil Ratio, SCF/STB	800	
API, degrees	34	

The capillary pressure and relative permeability curves are available for six rock types according to their absolute permeability ranges. However, only three curve types were

employed in the current model according to permeability ranges of sand, shaly-sand, and shale lithotypes, as included in Table 5.2.

Table 5.2: Capillary pressure and Relative Permeability for Reservoir Lithofacies

Sand: Swi=9.5, Sor=34				Shaly-Sand: Swi=12.5, Sor=40				Shale: Swi=17.5, Sor=44			
Sw	Krw	Kro	Pc(psi)	Sw	Krw	Kro	Pc(psi)	Sw	Krw	Kro	Pc(psi)
0.0950	0	0.790	29.862	0.125	0	0.736	29.88	0.175	0	0.6830	29.86
0.1515	4e-006	0.639	5.746	0.1725	0.000017	0.593	6.91	0.213	16E-05	0.549	10.20
0.208	5.2e-005	0.500	2.498	0.22	0.000325	0.458	6.0	0.252	0.0009	0.423	7.270
0.2645	0.00017	0.374	1.984	0.2675	0.001504	0.340	5.18	0.290	0.0015	0.310	6.250
0.321	0.00125	0.265	1.690	0.315	0.004529	0.239	4.45	0.329	0.0045	0.214	5.560
0.3775	0.0087	0.174	1.426	0.3625	0.01102	0.156	3.8	0.367	0.0108	0.137	4.930
0.434	0.0146	0.102	1.190	0.41	0.0237	0.0915	3.23	0.406	0.0226	0.078	4.350
0.4905	0.0292	0.050	0.985	0.4575	0.0472	0.044	2.72	0.444	0.0428	0.037	3.830
0.547	0.065	0.017	0.808	0.505	0.0895	0.0155	1.88	0.483	0.0766	0.012	3.260
0.638	0.1572	0.0037	0.8	0.5525	0.164	0.00231	1.55	0.521	0.1311	0.002	2.930
0.7	0.2322	0.0006	0.66	0.60	0.294	0	0.66	0.560	0.2177	0	2.540
0.8	0.4038	0	0.53	-	-	-	-	-	-	-	-
1	1	0	0	-	-	-	-	-	-	-	-

### 5.2.3 Geological Modeling

For geological modeling, a full geostatistical reservoir characterization model was built for multivariate statistical analysis and 3D spatial distributions of permeability, porosity, and lithofacies. The complete description of the geological model was described in Chapter 3 and Chapter 4.

### 5.2.4 PVT behavior of fluids

A full PVT model was constructed through the WinProp package within the CMG reservoir simulator. WinProp is a compositional equation-of-state based phase behavior and properties calculation package. In WinProp, the fluid are characterized in terms of their components in order to allow compositional fluids interaction through porous media. The calculation steps implemented to build the complete PVT model, for the main pay/South Rumaila oil field, include:

1. Fluid component properties.

2. Plus fraction characterization.
3. Multiple-contact miscibility calculations and ternary diagram generation.
4. Flash, saturation, and envelope calculations.
5. Recombination and constant composition expansion calculations.
6. Differential liberation and constant volume depletion.
7. Separator and swelling tests.

### 5.2.5 Property Upscaling and Grid System

For field-scale reservoir modeling, there are some factors that should be taken into consideration to decide the property upscaling and grid dimensions selection. The factors are cost, time available to conduct the study, as well as the processing speed of the CPU. Type of simulator is also a significant factor to be considered in upscaling because the compositional model (GEM), for instance, requires a much more running time than the black-oil models (CMG, 2011).

Since the field has a 60 year production history with 40 producers and 20 injectors, it requires considerably large time to run the model. The created fine-scale geostatistical model has 210, 202, and 45 grids in I, J, and K directions, respectively. The regular dimensions of grid blocks are constant: 50 m x 50 m and the total number of grids was approximately 2 million. Because of the high computation time for large models, the model was upscaled to obtain a coarse-scale grid system with a 150 m x 150 m regular grid dimension. The grid numbers in the upscaled model in I, J, and K directions are 69, 66, and 12 grids (total=54648 grids). Based on the cross-sectional well log, the 45 vertical layers were upscaled to only 12 layers. Arithmetic and harmonic means formula were considered for well porosity and permeability upscaling, respectively for their efficient preserving of data variations and for meeting the acceptance upscaling criteria, which are shown soon. For spatial property upscaling, the root mean square (RMS), also known as quadratic means,

was adopted to obtain spatial coarse-scale reservoir model for the same reason. Figure 5.1 illustrates the 3D coarse-scale reservoir models that include horizontal and vertical permeability, porosity, and lithofacies.

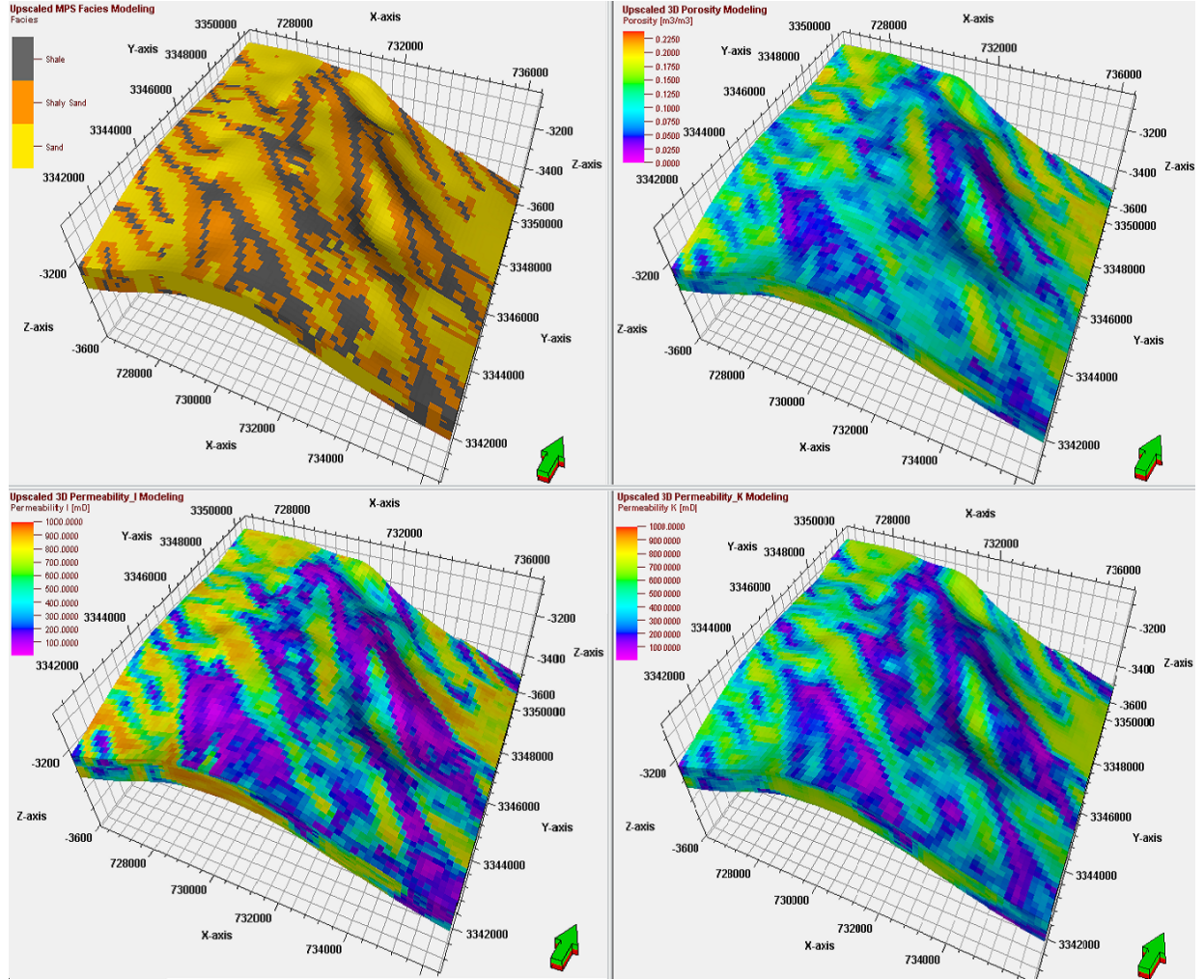


Figure 5.1: Coarse-Scale Geostatistical Models for Facies, Porosity, and Permeability

To check the quality of the upscaled model and ensure reduced simulation error, a matching between the fine and coarse-scale reservoir models was conducted in terms of grid bulk volume, cell angles, and cell inside out factors. Approximate similarity is required between pore volume and initial oil in place for the two models. The accepted percentage difference in volume needs be less than 7% (Petrel, 2014). For the current upscaled model, the percentage volume difference was 5.4%. A geometrical property cell angle

that represents the maximum deviation from 90 degrees needs to be visualized for upscaled model to check if some cells have angles far from 90 degrees. In Petrel, the deviation is acceptable if it does not exceed 15 degrees and the current upscaled model achieved this criterion (Petrel, 2014). The cell-inside-out was quantified to assure having zero values at all the reservoir areas. Matching histogram properties between fine and coarse scale models is likewise considered a good indicator for the quality of upscaling process, as illustrated in Figures 5.2, 5.3, 5.4, and 5.5. In these figures, there are non-significant difference in density distribution between the fine and coarse scale model with regards to petrophysical properties and lithofacies. Therefore, the coarse-scale models is an accurate enough to build the reservoir simulation model. Time step sizes, used in simulation, are normally

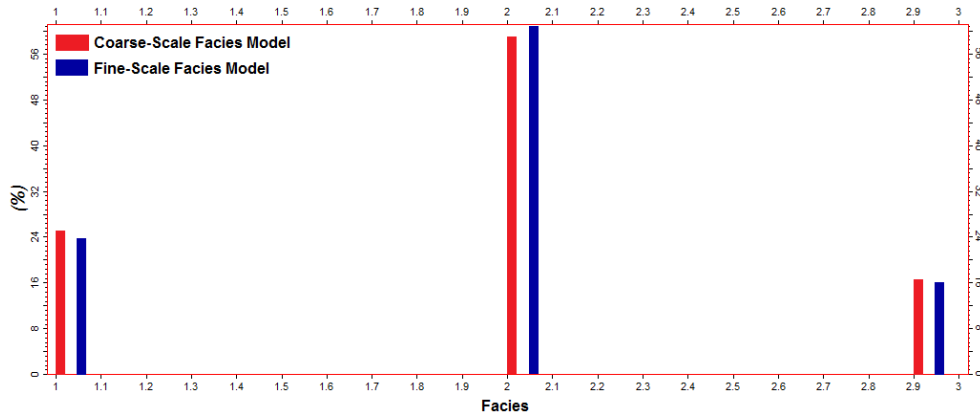


Figure 5.2: Histogram of Lithofacies Distribution of Coarse and Fine Scale Models

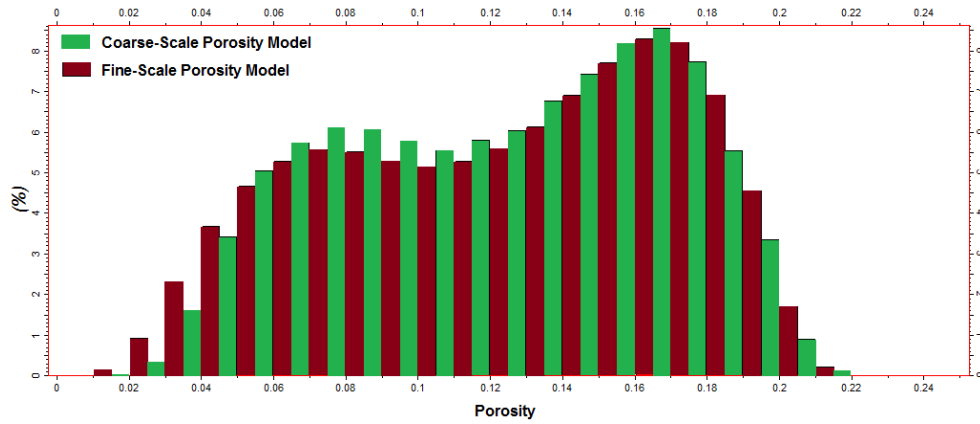


Figure 5.3: Histogram of Porosity Distribution of Coarse and Fine Scale Models

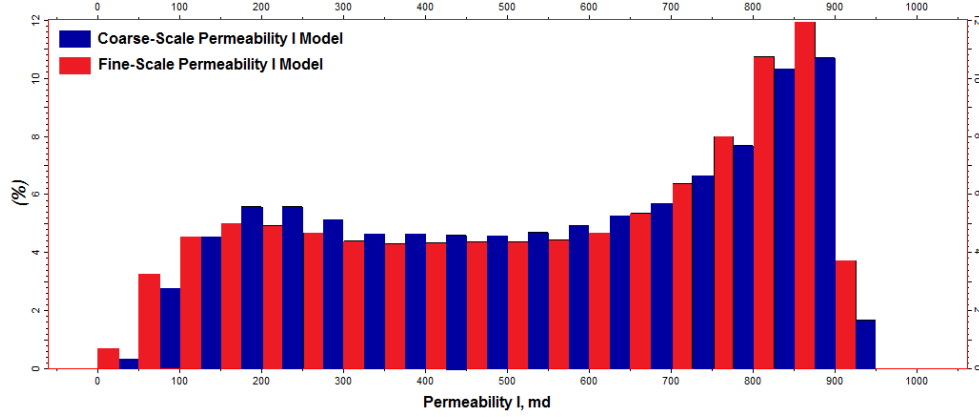


Figure 5.4: Histogram of Permeability I Distribution of Coarse and Fine Scale Models

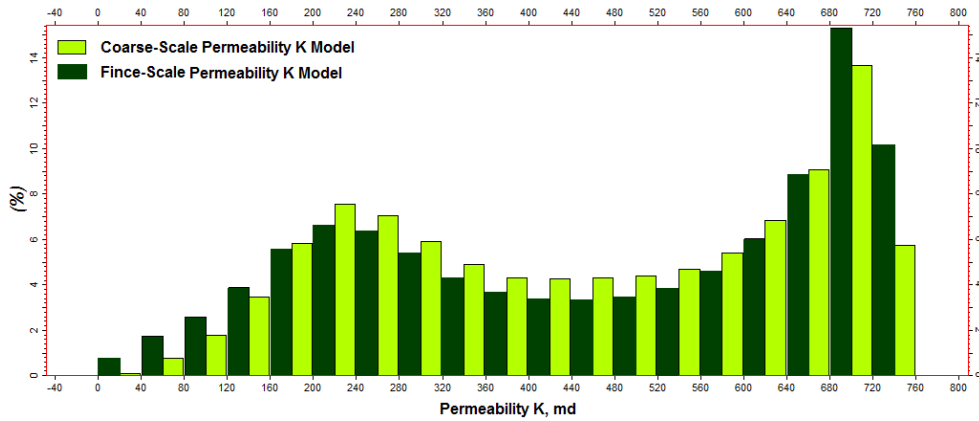


Figure 5.5: Histogram of Permeability K Distribution of Coarse and Fine Scale Models

in the range of 1-6 months. However, 1 month time steps were adopted because of the availability of measured field data on monthly production and injection rates.

### 5.2.6 Sensitivity Analysis

Screening the geological parameters for the sensitivity analysis, in the reservoir simulation, is an essential step for fast history matching. However, the history matching was manually achieved in this study. Therefore, the sensitivity analysis was performed for the purpose of identifying the most influential geological parameters on the GAGD process. Chapter 6 has a complete description of sensitivity analysis.

### 5.2.7 GAGD Process Simulation

The main pay reservoir in the South Rumaila oil field was selected for full detailed compositional reservoir simulation to enhance oil recovery through the GAGD process. The main pay does not contain any complex geological features such as faults or fractures. It has only three lithology types, which are sand, shaly sand, and shale, with distinct areal permeability distributions. The detailed reservoir description has been provided in Chapter 2.

The upscaled reservoir model was exported to build the compositional reservoir flow simulation. To apply the main concept of the GAGD process, several vertical injection wells and a series of horizontal producers with 3000 m length were placed through the reservoir at sand and shaly-sand lithology zones, which are far from shale channels. Initially, CO<sub>2</sub> is injected through the vertical injectors at the top two layers. At the same time, the 2nd three layers were left as a transition to allow a vertical depth interval for gas gravity drainage. Then, horizontal producers were set up through the sixth, seventh, and eighth layers where the oil saturation has the highest levels. Finally, the last four layers were left with no injection/production activity, as they are fully flooded with water from the infinite edge water aquifer. The total prediction period of CO<sub>2</sub>-GAGD process evaluation was set for 10 years (2016-2026).

Figure 5.6 shows the locations of injection and production wells that are installed for CO<sub>2</sub> flooding in the Gas Assisted Gravity Drainage process. The reservoir body is represented by the red color, which denotes to shale zones. However, the perforations of producers and injectors were mainly placed in sand zones (indicator number 2), which have high permeability ranges, as well as shaly-sand zones (indicator number 1). Figure 5.7 shows the locations of the CO<sub>2</sub> injection wells and horizontal producers in the high permeability zones, where the sand and shaly-sand facies were distributed. Red color refers to the high permeable zones of sand and light red color represents the shaly sand zones, as well. This slab figure was prepared to provide an imagination about the extension of each



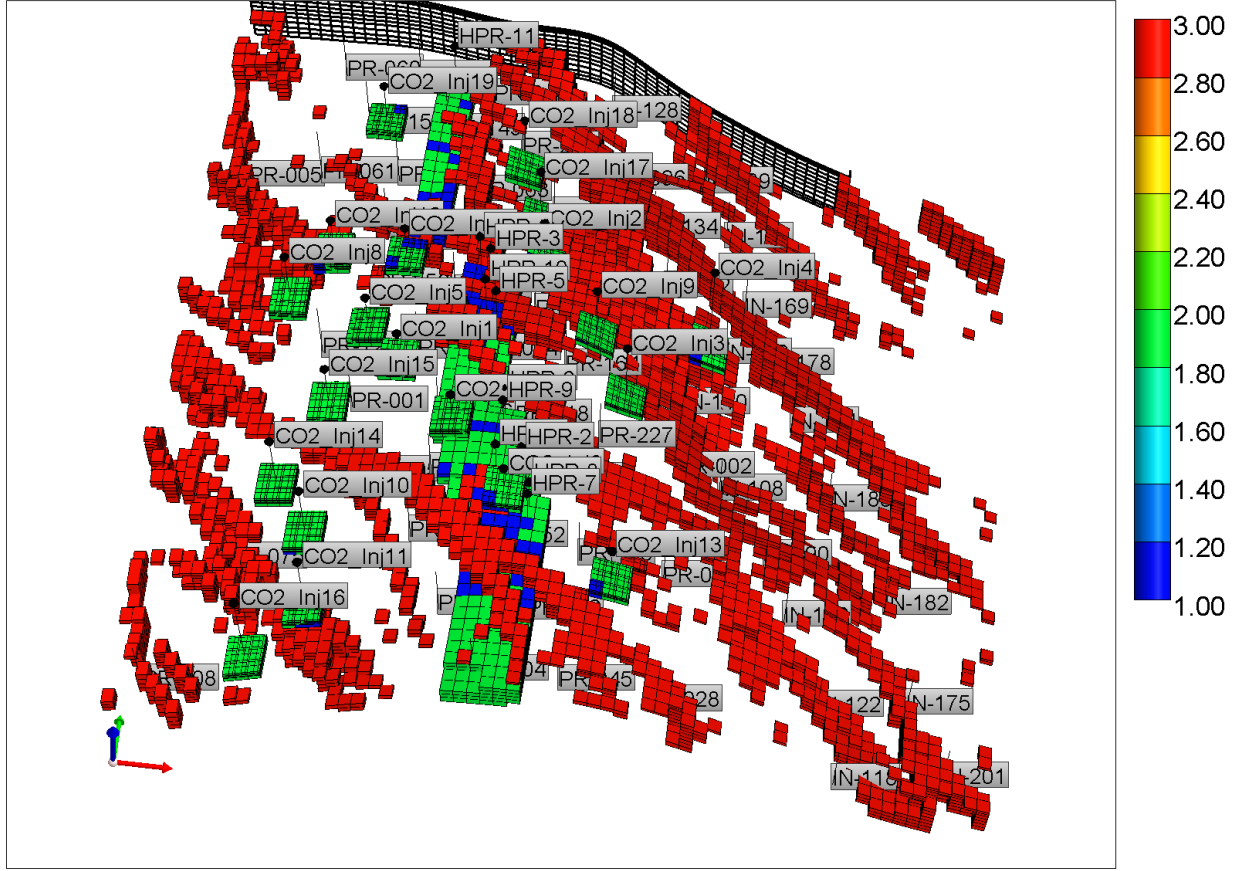


Figure 5.6: Production and Injection Well Locations in Sand and Shaly-Sand Zones

well perforations inside the reservoir.

The edge-water drive aquifer, which is located at the eastern and western boundaries, was activated in the reservoir model to support pressure maintenance and improve history matching procedure. The Carter-Tracy aquifer approach was adopted to simulate the infinite active aquifer.

In order to improve the description of fluid flow especially near wellbores, the orthogonal corner point grid structure was considered as a more appropriate approach in the reservoir model rather than the cartesian. Corner point geometry is more appropriate for complex full-field studies as it preserves the accuracy of fluid flow modeling and well treatment. Additionally, it acclimates to different geological characteristics such as faults, horizontal wells, and complex flow patterns (Ding and Lemonnier, 1995).

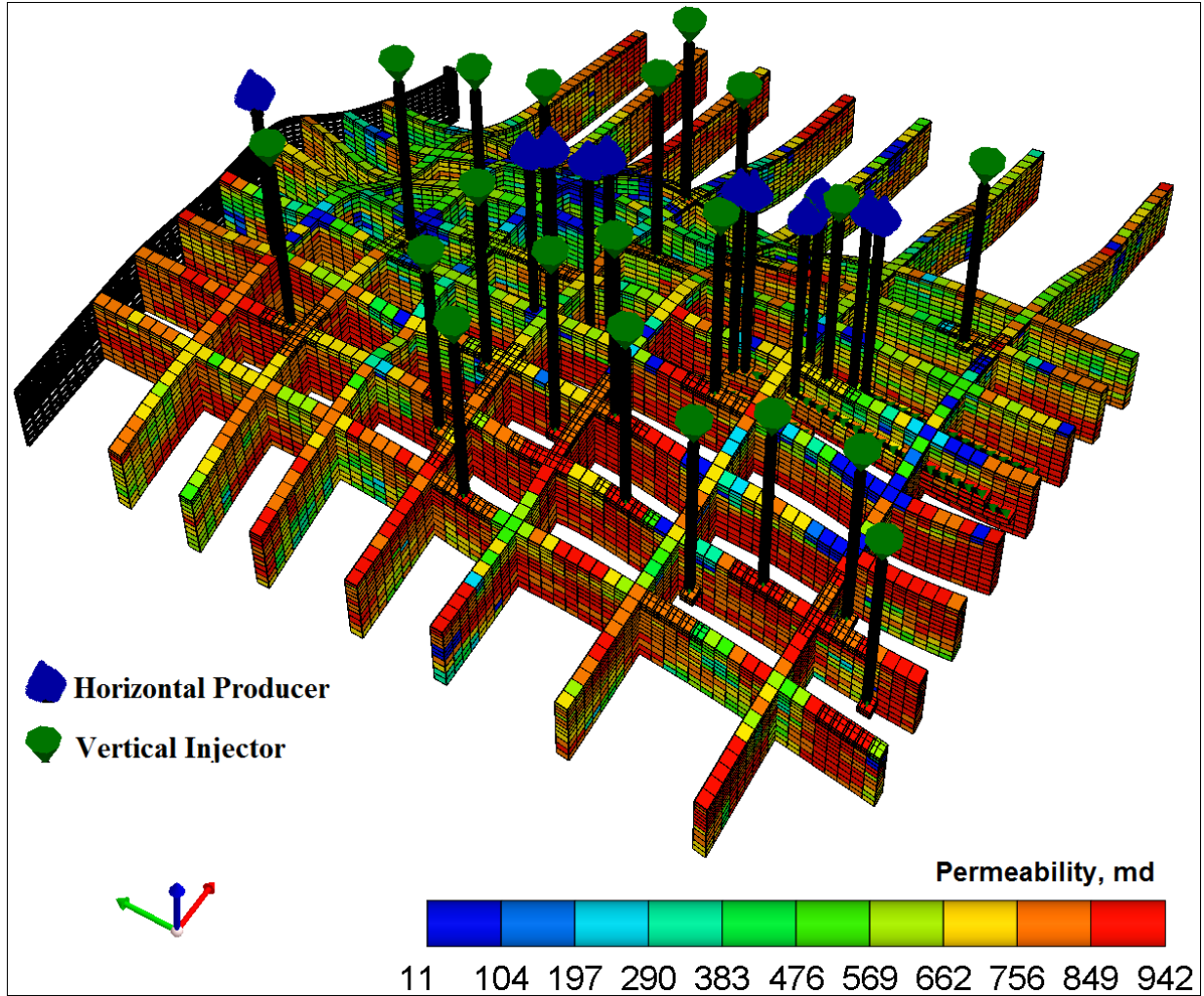


Figure 5.7: Locations of Horizontal Producers and Vertical Injectors through the GAGD Process

In addition, small grid blocks are needed in regions where fluid saturations have large variations. Therefore, non-uniform grid refinements were set in the vicinity of the CO<sub>2</sub> injection wells through the first two layers that have perforations. The grid refinement leads to eliminate sharp changes in the fluid property calculations. Moreover, it precludes the instability of the compositional transition between the injected CO<sub>2</sub> and crude oil (viscous fingering). Contrary, selecting coarse grid blocks in the compositional modeling of the CO<sub>2</sub> flooding gives rise to a significant effect of numerical dispersion, which leads to inaccurate recovery calculations.

Given these points, three horizontal and two vertical grid blocks were included for the refinement around well blocks. In I and J (horizontal) directions, the number of block divisions in the 1st, 2nd, and 3rd blocks were set to be 3, 2, and 1, respectively. Likewise, the number of block divisions in the 1st and 2nd vertical blocks was 2 and 1, respectively.

Figure 5.8 shows the non-uniform areal grid system that was employed around the CO<sub>2</sub> injection wells in the GAGD reservoir simulation.

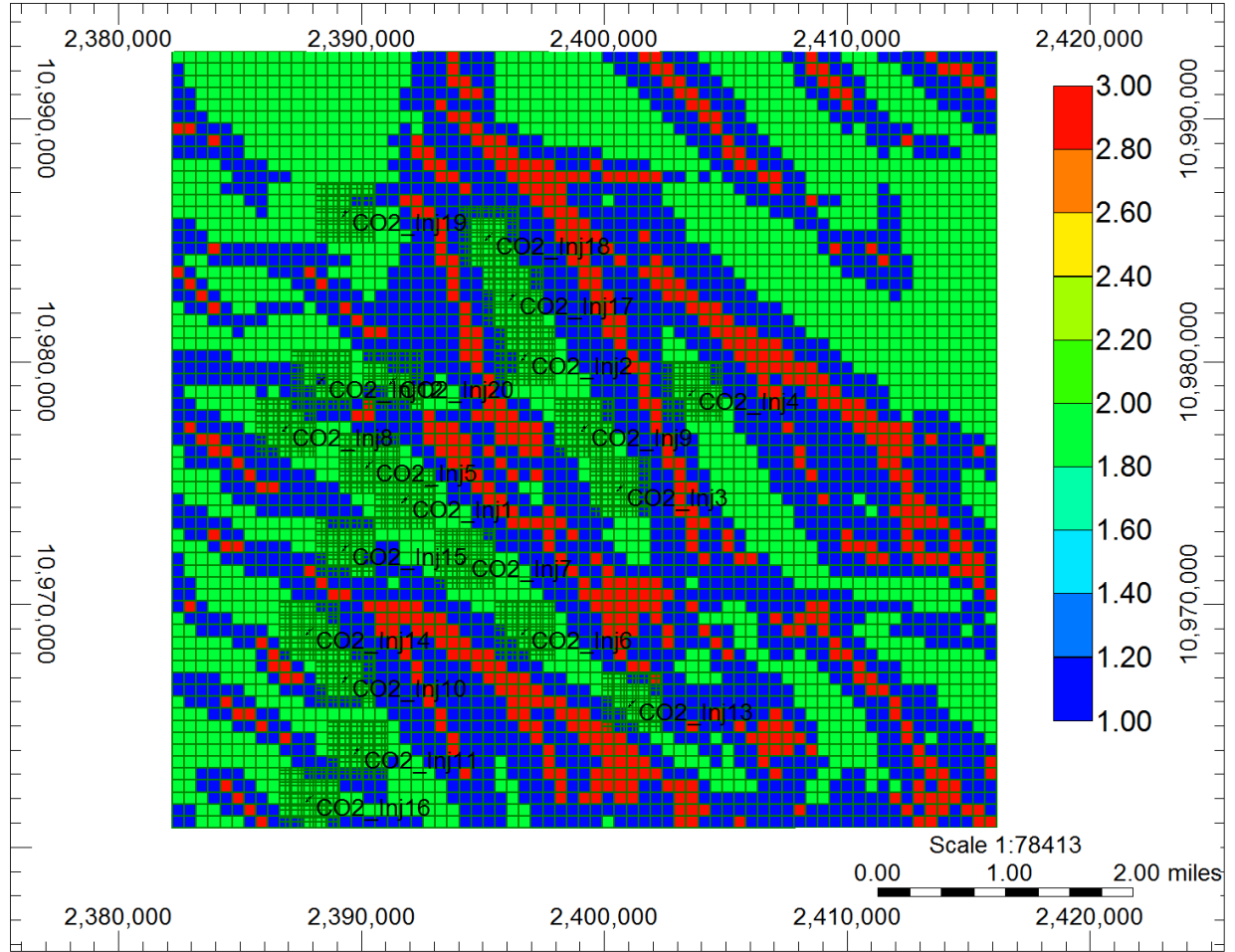


Figure 5.8: Non-Uniform Grid System for CO<sub>2</sub>-GAGD Injectors in Rock Types Map

## 5.3 Results and Discussion

### 5.3.1 Initialization Results

The simulation initialization results include original oil, water, and gas volumes in place, in addition to the pressure and fluid saturation distributions at the first time-step. Table 5.3 illustrates the reservoir properties as well as original and current fluid volumes in place for the main sector of the main pay/South Rumaila oil field. The table also has the current fluid volumes in place at the last time-step of the simulation period (1-1-2016).

Table 5.3: Reservoir Properties and Initial and Current fluid Volumes in Place

Initial Reservoir Properties	
Average Pressure, psia	= 5.27785e+03
Average Porosity, faction	= 0.16333
Average Water Saturation	= 0.133
Total Bulk Volume, res ft3	= 3.66631e+11
Total Pore Volume, res ft3	= 5.98814e+10
Total Hydrocarbon Pore Volume, res ft3	= 3.75195E+10
Originally in Place at $P_i = 5.277$ psia	
Oil in Place, STB	6.12305e+09
Dissolved Gas in Place, SCF	9.46021e+11
Water in Place, STB	4.04288e+09
Reservoir Water, MMM rbbl	7.62289
Current in Place at $P_f = 2,800$ psia	
Oil in Place, STB	2.02074e+09
Dissolved Gas in Place, SCF	3.04101e+11
Water in Place, STB	7.69565e+9
Reservoir Oil, rbbl	2.24386e+9
Reservoir Water, rbbl	7.62289e+9
Cum Water Influx, STB	1.04104e+10

### 5.3.2 History Matching

Validation of the reservoir model as relates to matching between the observed and calculated flow parameters is a crucial inverse step in the integrated reservoir studies (Kabir et al., 2003). The validated reservoir model truly represents the fluid flow behavior through

porous media and better captures the current pressure and saturation state. The reservoir model can be then used to predict the future reservoir performance through reservoir management and optimization, as shown in this study.

Performing an acceptable history matching is not an easy step as it requires many iterations of simulation runs to reduced the discrepancy errors between observed and calculated values. The difficulty of history matching process comes from the unknown uncertainties in reservoir and production parameters (Kabir et al., 2003).

The parameters considered for history matching are: production and pressure data, average reservoir pressure, well bottom hole pressure, water cut, initial time of breakthrough, and depth of oil-water contact. The most common parameters changed to achieve history matching are: permeability and porosity distributions, aquifer strength, relative permeability and capillary pressure, along with vertical to horizontal permeability ratios ( $Kv/Kh$ ). The mismatch error can be quantified by the weighted sums of squares of difference between the observed response  $Y^{obs}$  and the calculated model results  $Y^{cal}$  (Landa, 2001):

$$Q(x) = \sum_{i=1}^{n_{obs}} w_i (Y^{obs} - Y^{cal})^2 \quad (5.1)$$

where:

$Q(x)$ : the mismatch error between predicted and measured history matching parameters.

$w_i$ : weight coefficient of mismatch.

$Y^{obs}$ : the observed parameters.

$Y^{cal}$ : the calculated parameters.

In this study, an excellent history matching was obtained through trial and error process with respect to cumulative oil production and water injection as well as fluid flow rates in regard to the entire field together with all wells. The production and injection matching is a good indicator of reservoir and fluid behavior as it reflects the matching of water cut and saturation distributions. It was not necessary to adopt an optimization tool for

automatic history matching procedure because the matching, in this study, was performed through limited number of runs after setting the suitable aquifer size.

The entire production history for the simulation period in this study is approximately 56 years. The production and injection flow rates were available until February, 2010. Therefore, the history matching was achieved from 1954 until 2010. Figure 5.9 shows the matching between field production rates and cumulative oil production. Additionally, Figure 5.10 illustrates the production matching for only four producers. The production matching curves of all other producers, were provided in Appendix D. On the other hand, Figure 5.11 shows the matching between field injection rates and cumulative water injection. Similarly, Figure 5.12 depicts the water injection matching for only four injectors. However, the history matching of other injection wells were also included in Appendix D.

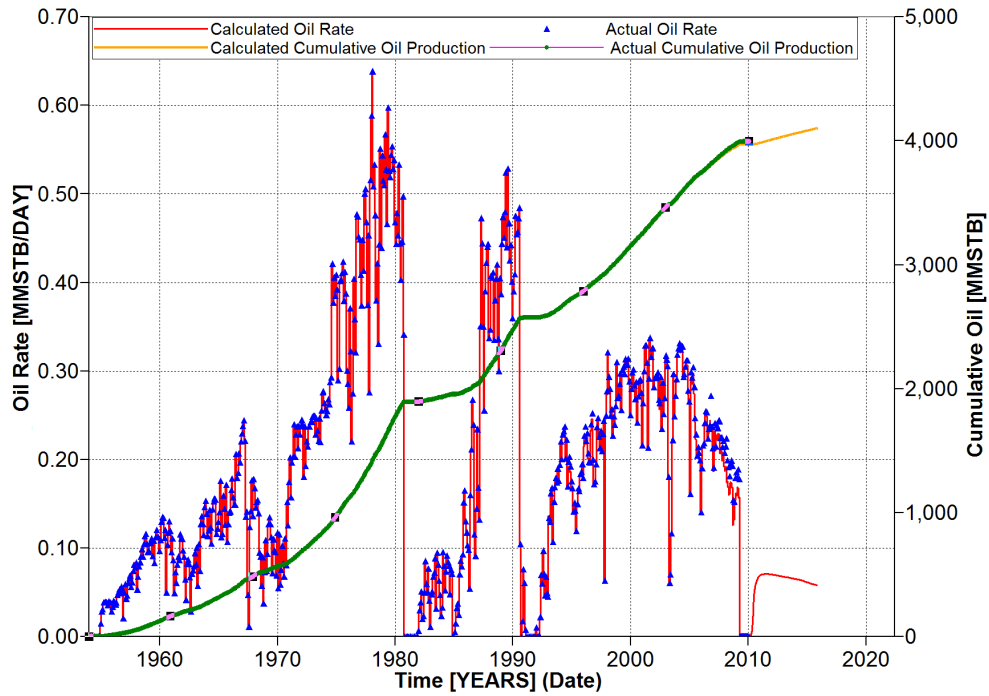


Figure 5.9: History Matching of Entire Field Production of South Rumaila Oil Field

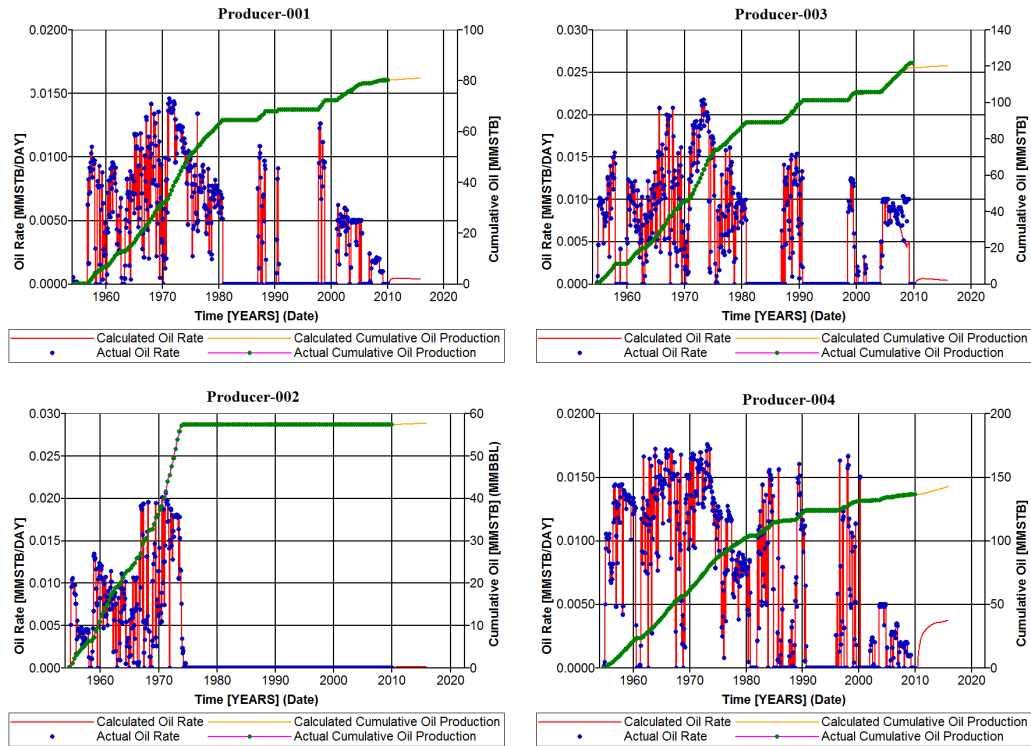


Figure 5.10: History Matching of Production Wells

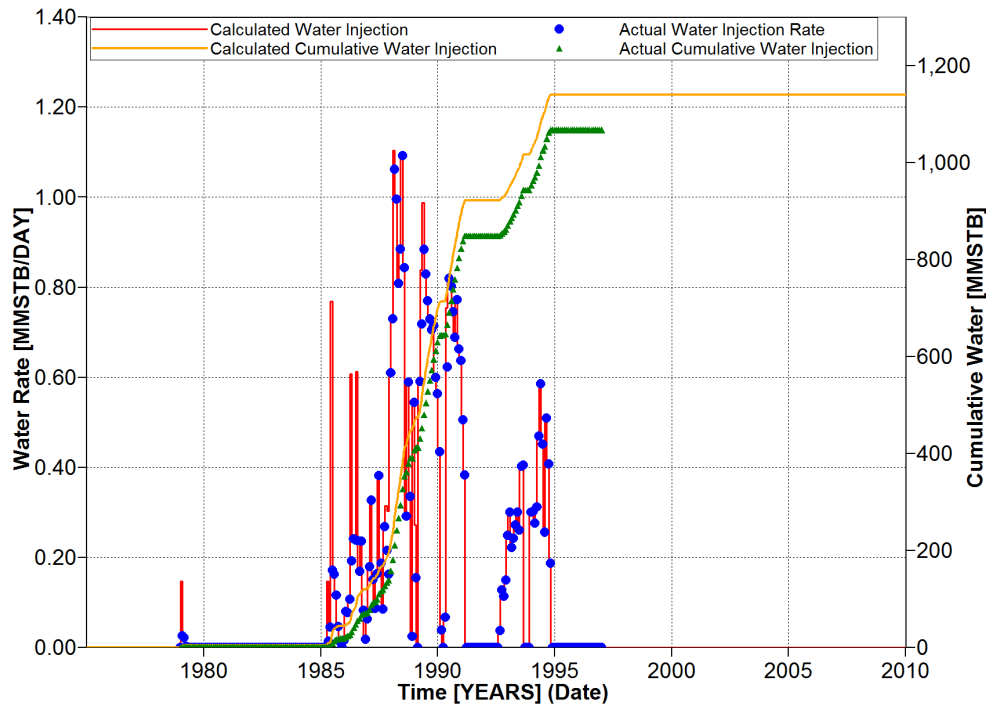


Figure 5.11: History Matching of Entire Field Injection of South Rumaila Oil Field

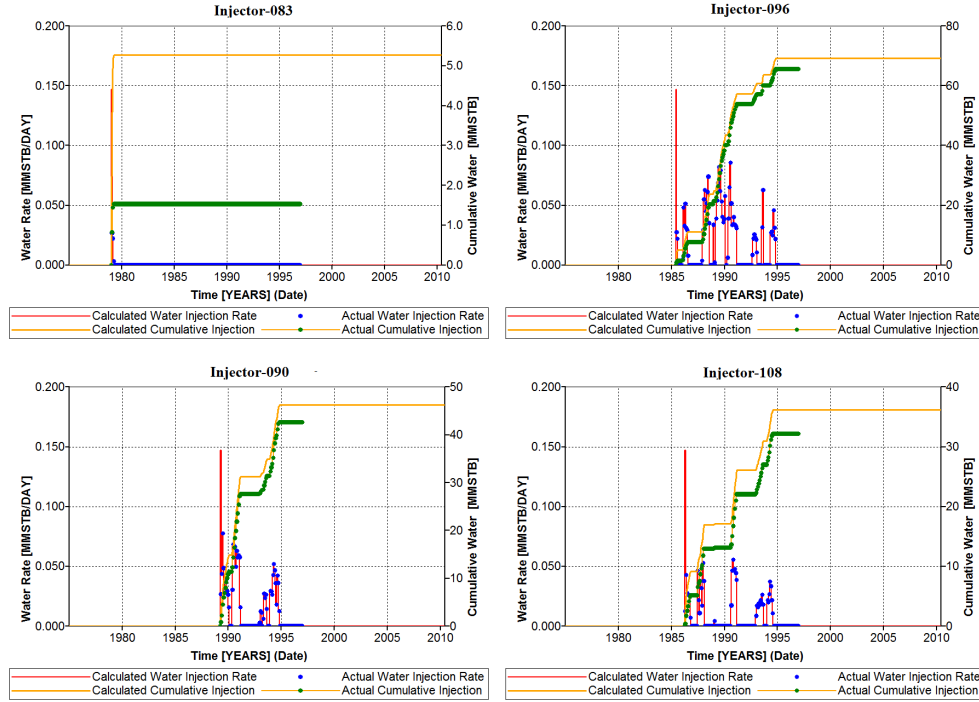


Figure 5.12: History Matching of Injection Wells

### 5.3.3 Base-Case GAGD Simulation Results

The period from 2010 until the end of December, 2015 was left under primary recovery without applying any IOR/EOR process. The GAGD reservoir simulation starts on January 1, 2016 and the production forecasting period extends for 10 years until January 1, 2026.

At the beginning of prediction period, the remaining oil in place is 2.02869 billion stock-tank barrels and the current oil in place is 6.12305 billion stock-tank barrels. The cumulative oil production for the period from 1954-2016 and current oil recovery factor are 4.09436 billion stock-tank barrels and 66.8%, respectively. Table 5.4 exhibits the cumulative fluid production and injection until the 1st time-step of prediction period (1-1-2016).

The immiscible base case of GAGD process was defined based on the default setting of operational design factors in injection and production wells. The factors (constrains) activated for CO<sub>2</sub> flooding through the GAGD process are maximum oil production rate (STO) and minimum bottom hole pressure in production wells (BHP). Additionally, max-



Table 5.4: Current Cumulative fluid Production and Injection (Prior to GAGD)

Cumulative Production		Cumulative Injection	
Oil, STB	4.10231e+09	Solvent, SCF	0
Dissolved Gas, SCF	6.4284e+09	Water, STB	1.14004e+09
Water, STB	7.91679e+09		
Oil Recovery Factor, %	66.8		

imum gas injection rate (BHG) and maximum bottom hole injection pressure (BHP) in injection wells were also specified in the CO<sub>2</sub> injection wells. Table 5.5 shows the parameters values set of the base GAGD process case. The forecasting period of the GAGD

Table 5.5: Base Case Setting of the GAGD Operational Design Parameters

Minimum BHP in production Wells, psi	2660
Maximum STO in production wells, STB/DAY	750000
Maximum BHP in Injection Wells, psi	3000
Maximum BHG in Injection wells, SCF/DAY	1e+007

process base case extends to January 1, 2026 for 10 years. The GAGD process led to a significant increment in oil recovery in comparison with the primary production case, as illustrated in Figure 5.13. The cumulative oil production at the beginning of the prediction period is 4.09436 billion stock-tank barrels and by the end of the prediction period through the primary production is 4.255 billion stock-tank barrels. However, the implementation of GAGD process resulted in obtaining approximately 4.4 billion bbls of cumulative oil in 10 years. This is an increase of 145 millions barrels over primary depletion. The field oil production rate through the GAGD process has higher rates than the primary production. According to Figure 5.13, the amount of cumulative oil production, by primary production at the end of prediction period, can be obtained in less than two years through the GAGD process. This demonstrates effectiveness of the GAGD process in improving oil rates and recovery. Due to the fast sweep efficiency in the GAGD process (Figure 5.13), the field oil rate jumped to a very high level for a limited time interval at the beginning of prediction period. The curve of gas production rate refers to the amounts of released gas during the

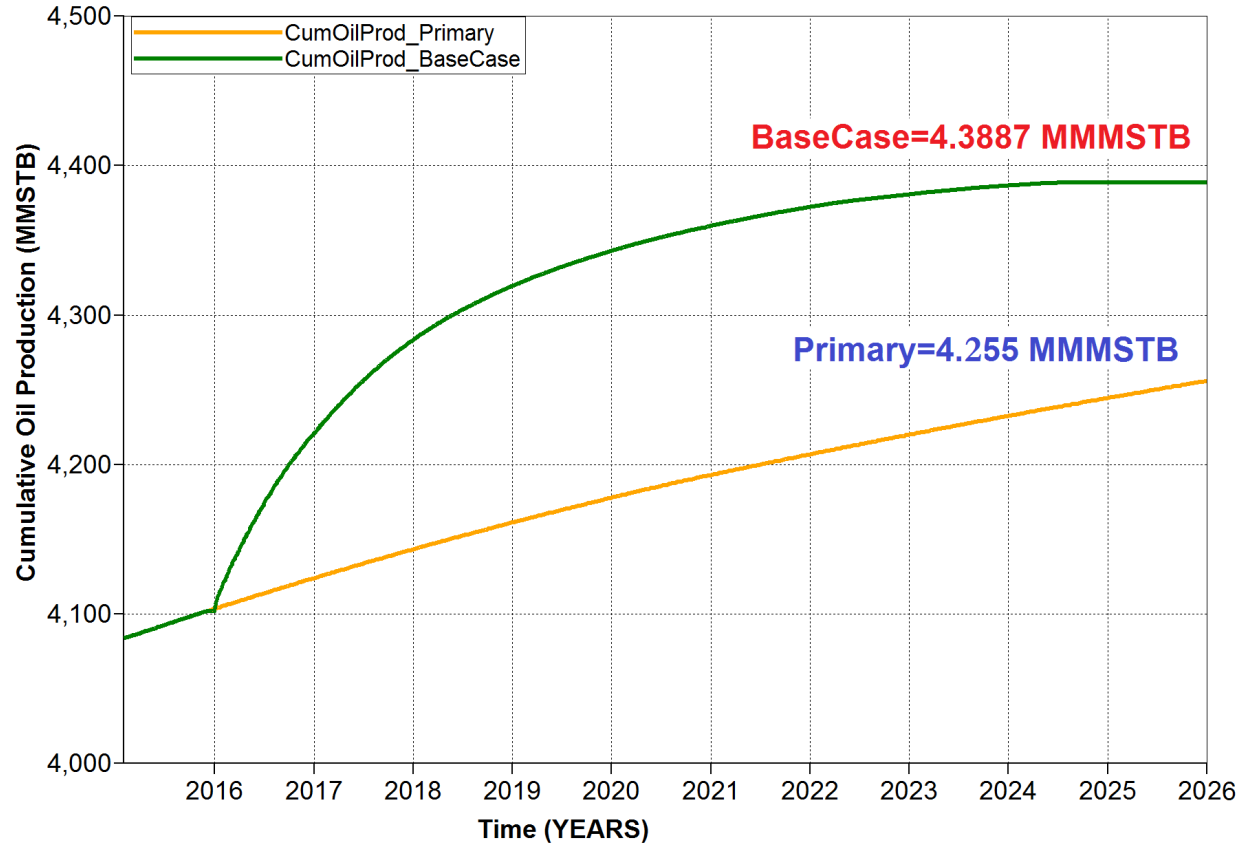


Figure 5.13: Comparison of Field Oil Response between Primary Production and BaseCase GAGD Process

production operations when the well flow pressure becomes less than the bubble-point pressure in the vertical tubing. The increasing in reservoir pressure that supports the well flow pressure causes the decreasing in gas production rates during the period from 2021-2026. The increasing in average reservoir pressure was increased due to accumulating the injected gas at the top of the reservoir.

The oil recovery factor through the immiscible GAGD process in a 10 years of prediction period is only 15%. In addition, the GAGD process has led to a faster oil production than the primary case. This fact can be easily identified in Figure 5.14 and 5.15. More specifically, Figure 5.14 illustrates the 3D pressure and ternary distributions at the initial and last time-steps of the primary production case's forecasting period. No significant changes in oil saturation between the period from 2016 to 2026 can be seen in the ternary distribution images of primary production, as illustrated in Figure 5.14. This fact reflects the

inefficiency of primary production to increase oil recovery. The same fact appears in the pressure distribution images as there is limited pressure decline through the same period. Figure 5.15 depicts the 3D pressure and fluid distributions at the last time-step of the GAGD base case's prediction period. No breakthrough of the injected gas can be identi-

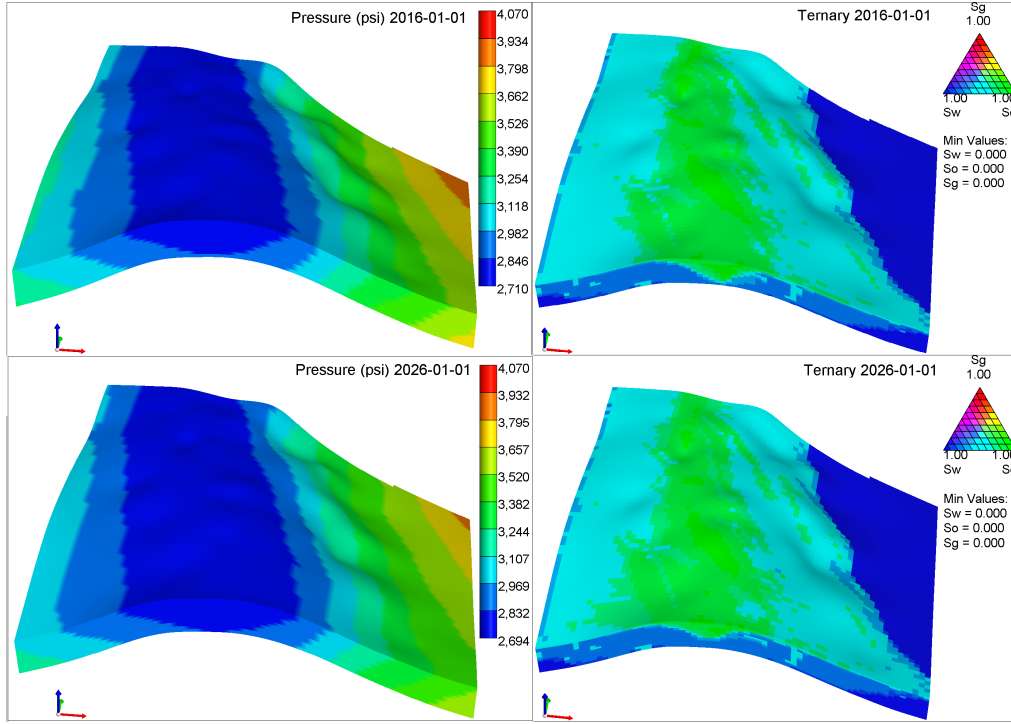


Figure 5.14: 3D Pressure and Ternary Fluid Distributions for the Primary Production Case at 2016 (top) and 2026 (bottom)

fied from the gas saturation image at the end of prediction period, as shown in Figure 5.15. All the injected gas through the 10 years accumulated at the top of the payzone and did not reach the production zone located exactly at the middle of the reservoir. This fact can be seen in Figure 5.16 that shows low Gas Oil Ratios (GOR) for the base case of the GAGD process simulation. The no gas breakthrough is also substantiated by the observation in Figure 5.16 that shows a continual increase in average reservoir pressure, as there is no increased gas production, which results in decreasing the reservoir pressure. In addition, the low GOR values also indicates there is no gas breakthrough in the oil production wells as all the injected CO<sub>2</sub> accumulates at the top of the payzone.

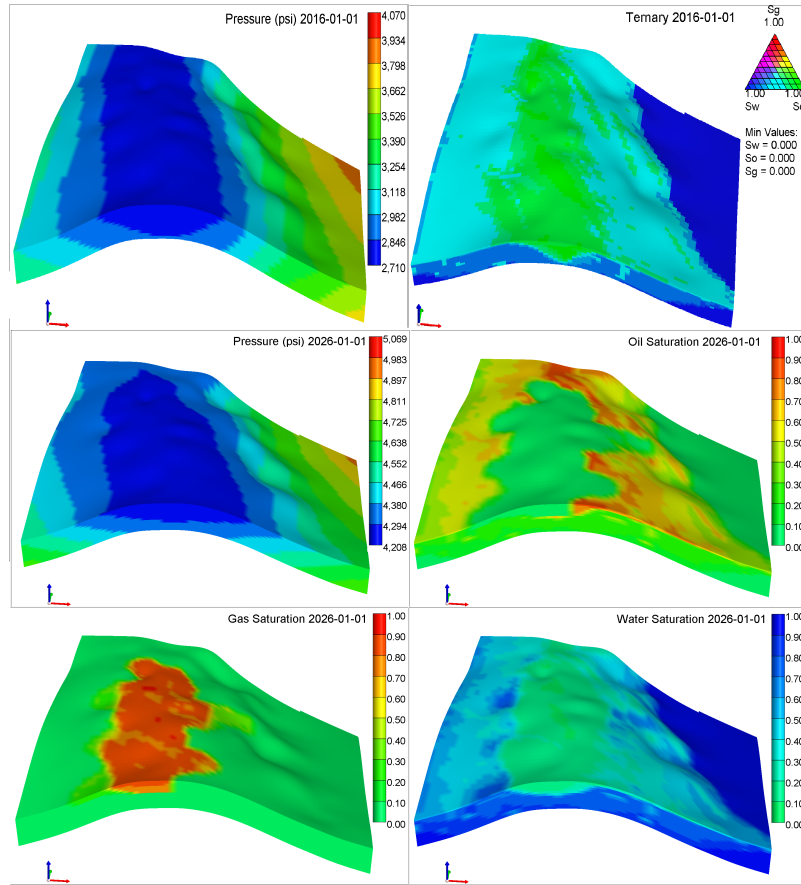


Figure 5.15: Effectiveness Comparison of the 3D Pressure and Fluid Distributions for the Base-Case GAGD Implementation at 2026

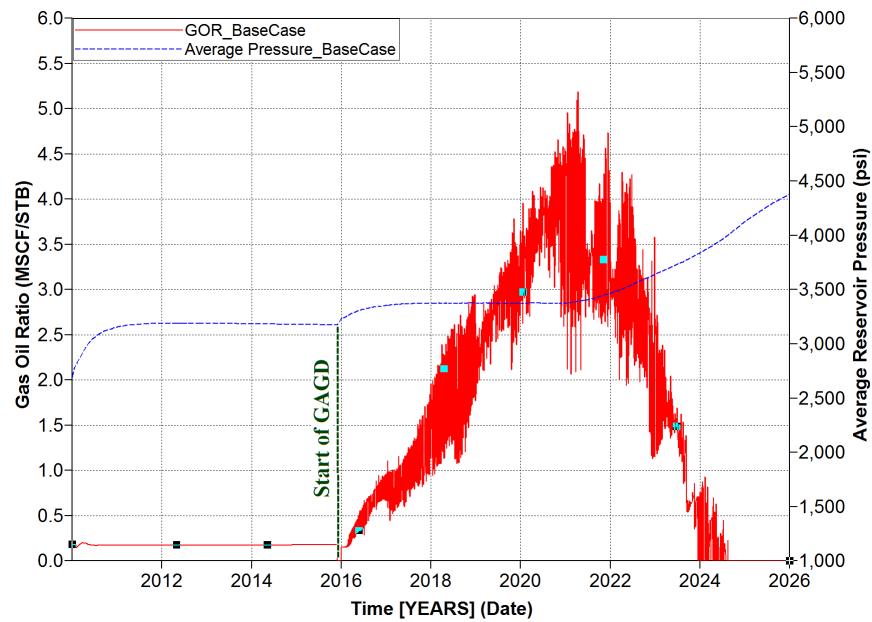


Figure 5.16: Gas Oil Ratio and Average Pressure for the Base-Case GAGD Process

In this immiscible case (Figure 5.16), the low GOR reflects the negligible gas breakthrough in the horizontal producers. More specifically, GOR keeps increasing from the beginning of the prediction period and continues until year 2021. This increase comes from the released gas through the vertical tubing when the well flowing pressure becomes less than bubble-point pressure. After year 2021, GOR starts decreasing because of average reservoir pressure increasing due to accumulating a large amount of injected gas at the top of the reservoir. The increasing reservoir pressure results in keeping the well flowing pressure above the bubble-point pressure through the production tubing.

It can also be identified from Figure 5.16, in addition to the pressure distribution map at the last time-step in Figure 5.15, the average reservoir pressure gradually increases and at the end of 10 years exceeds 4000 psia. The injected gas causes this increase in reservoir pressure leading to improving the oil recovery.

Oil recovery through 10 years of GAGD process prediction period is only 15%. In order to show the effectiveness of GAGD process to reach promising levels of oil recovery, two special cases, immiscible and miscible, were implemented for a 25 year of prediction period. When running the reservoir simulator for 25 years rather than 10 years, the recovery factor through immiscible and miscible special cases reach to 30% and 42%, respectively, as shown in Figure 5.17. In the same figure, the production rates during the entire prediction period of the miscible GAGD process can also be identified as higher than the immiscible case.

Since the reservoir pressure at the beginning of the GAGD process simulation was 2700 psi, the maximum injection pressure of CO<sub>2</sub> was set to 3000 psi to provide immiscible conditions. The minimum bottom hole pressure in the production wells was set equal to the bubble point pressure in order to avoid producing the dissolved gas in reservoir oil. The continuous CO<sub>2</sub> injection in high rates through miscible mode for a long time can cause early gas breakthrough, therefore, increase the Gas Oil Ratio. Additionally, the miscible continuous CO<sub>2</sub> injection leads to a much higher GOR than the immiscible case, as illustrated in Figure 5.18, which shows the GOR and average reservoir pressure for the special

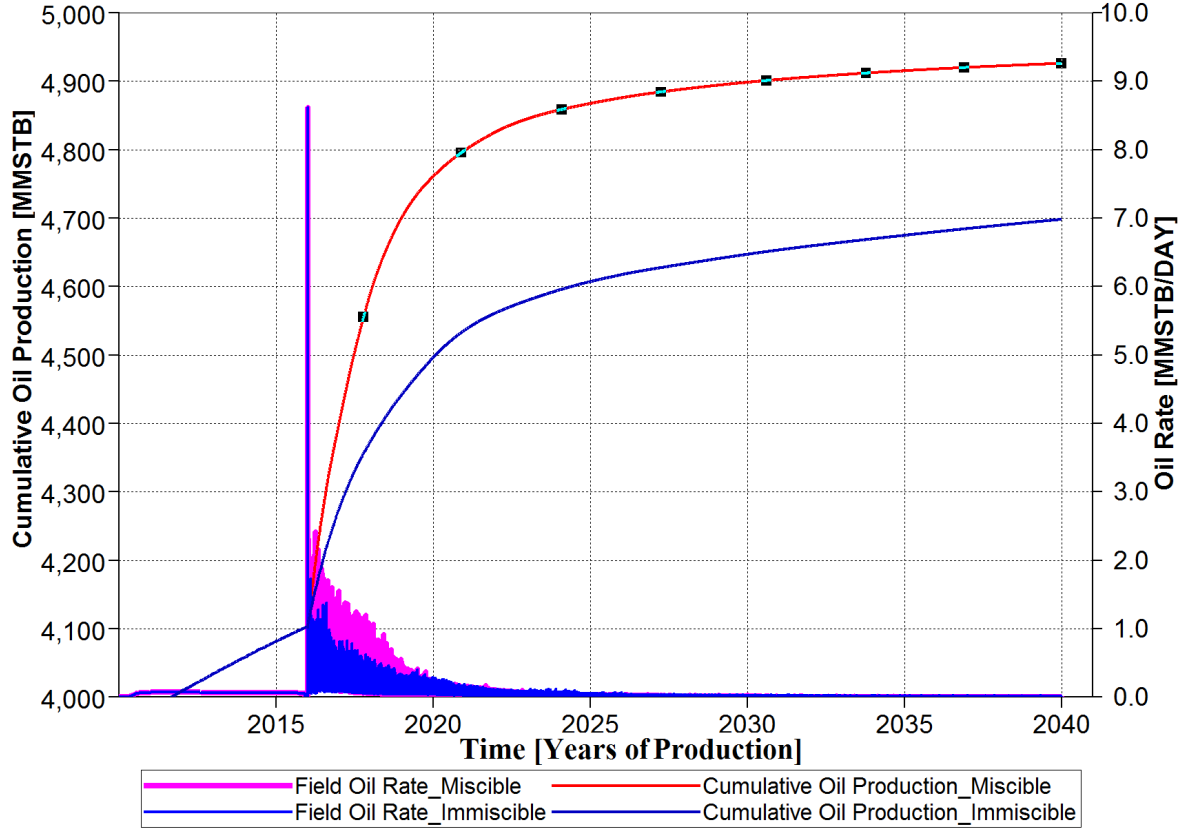


Figure 5.17: Comparison of Field Flow Response between Immiscible and Miscible GAGD Process for 25 Years of Production

Immiscible and miscible cases. In the miscible case, the fluctuations in the GOR curve come from the different times of reaching the production zones due to the heterogeneous reservoir. The comparison of gas breakthrough between the immiscible and miscible special cases of the GAGD process are illustrated in the side view images of gas saturation advance in years 2021, 2026, 2035, and 2040, respectively. Figures 5.19, 5.20, 5.21, and 5.22 show the slow immiscible injected gas advancement towards the middle of the payzone where the horizontal production wells are installed. Furthermore, Figures 5.23, 5.24, 5.25, and 5.26 reveal how the CO<sub>2</sub> moves faster towards the production zone in miscible mode than the immiscible mode. This fact can be also identified from Figure 5.18 that shows higher GOR rates for the miscible case at the beginning of the prediction period.

The absence gas breakthrough in immiscible flooding and early breakthrough in miscible mode after 5, 10, 20 and 25 years of injection can be also observed in the 3D gas and

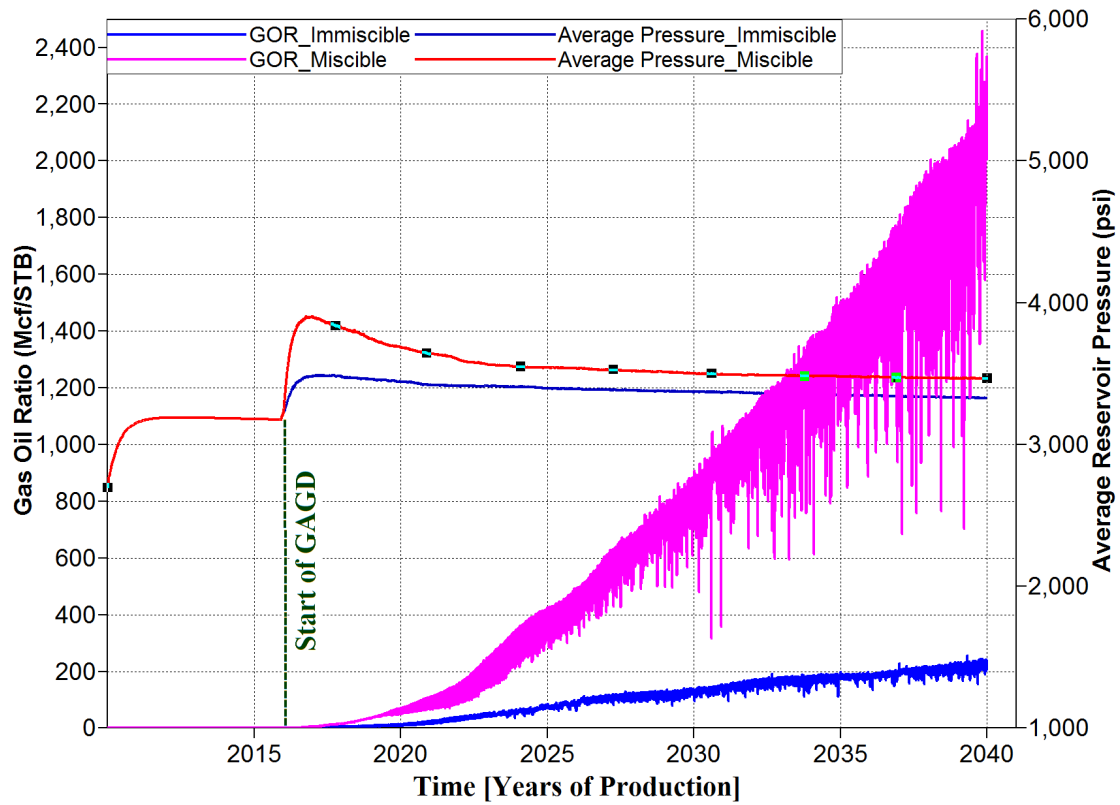


Figure 5.18: Gas Oil Ratio and Average Reservoir Pressure for Immiscible and Miscible Special Cases of GAGD Process

oil saturation distributions in Figures 5.27, 5.28 along with Figure 5.29, respectively.

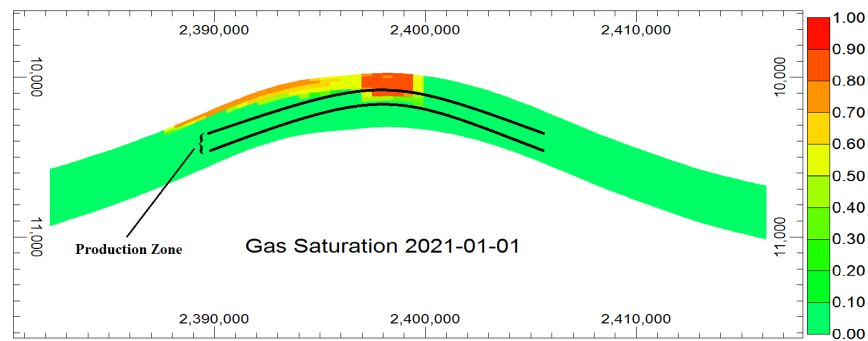


Figure 5.19: Side View of Immiscible Injected Gas Advancement in year 2021

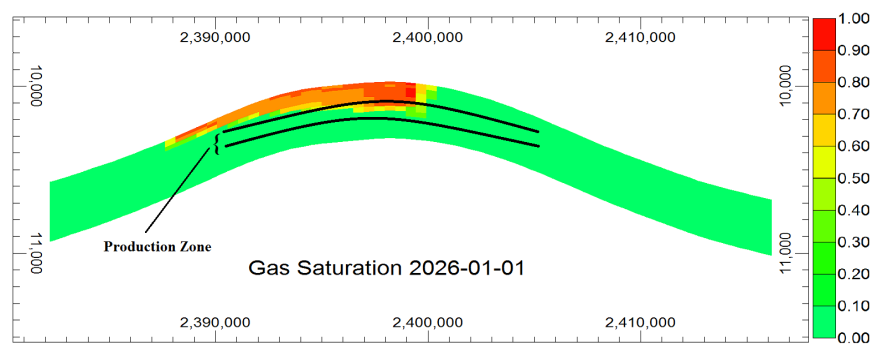


Figure 5.20: Side View of Immiscible Injected Gas Advancement in year 2026

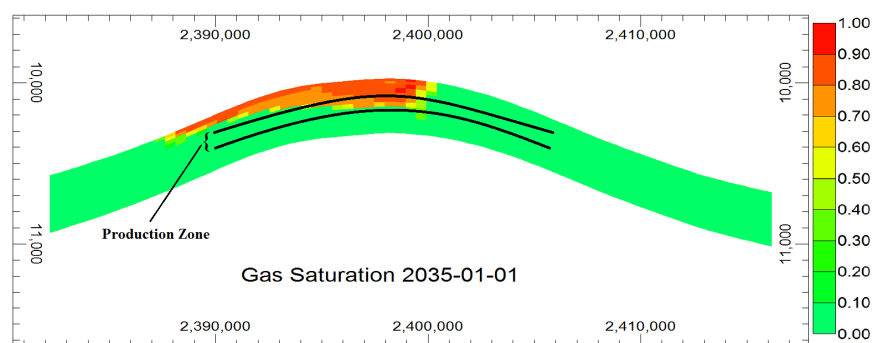


Figure 5.21: Side View of Immiscible Injected Gas Advancement in year 2035

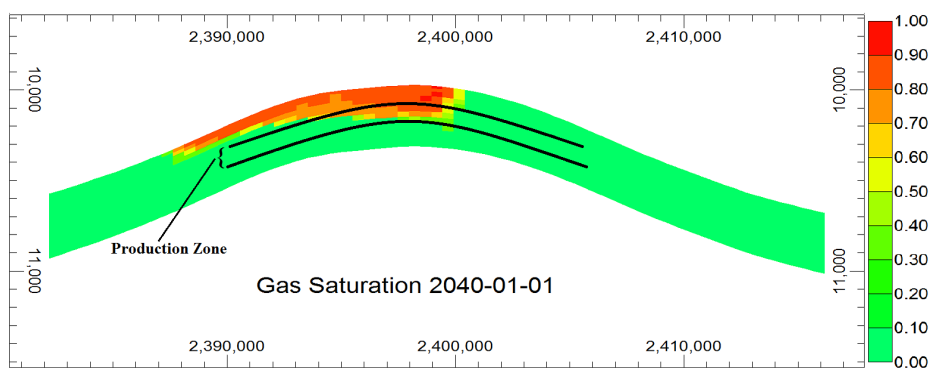


Figure 5.22: Side View of Immiscible Injected Gas Advancement in year 2040



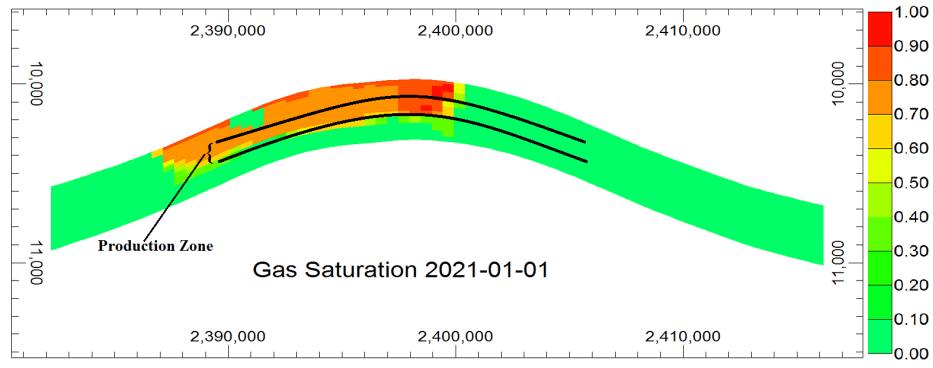


Figure 5.23: Side View of Miscible Injected Gas Advancement in year 2021

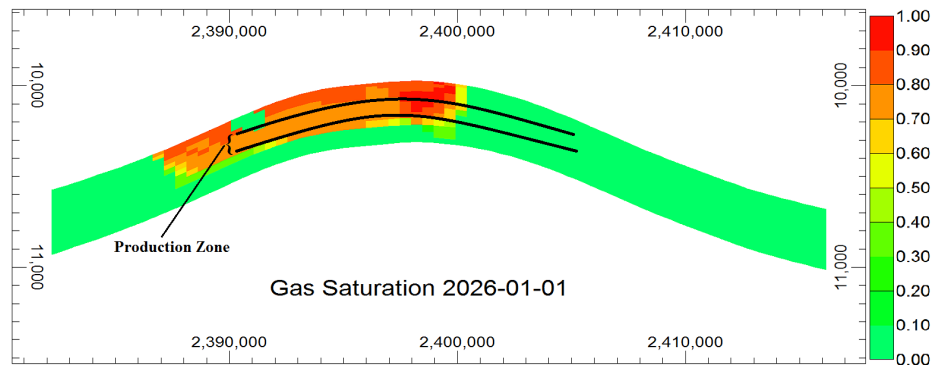


Figure 5.24: Side View of Miscible Injected Gas Advancement in year 2026

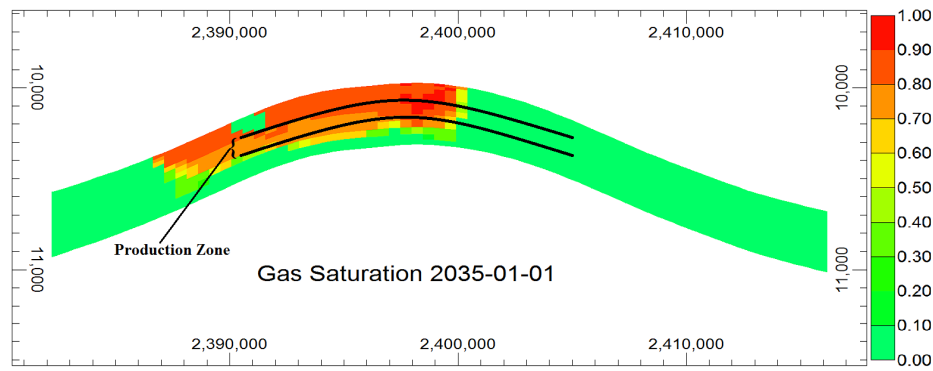


Figure 5.25: Side View of Miscible Injected Gas Advancement in year 2035

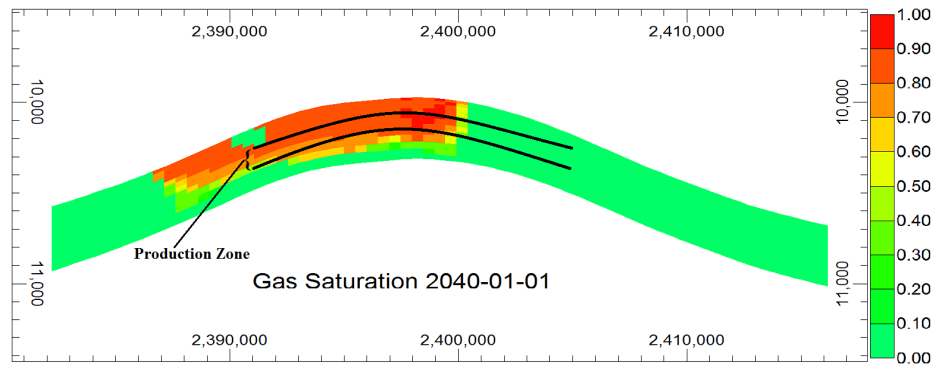


Figure 5.26: Side View of Miscible Injected Gas Advancement in year 2040

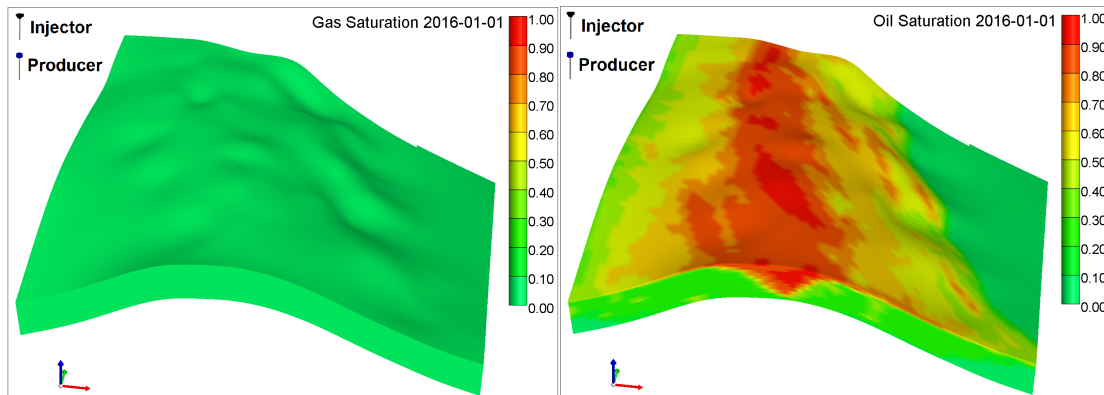


Figure 5.27: Gas and Oil Saturation Distributions in 2016 (Before GAGD)

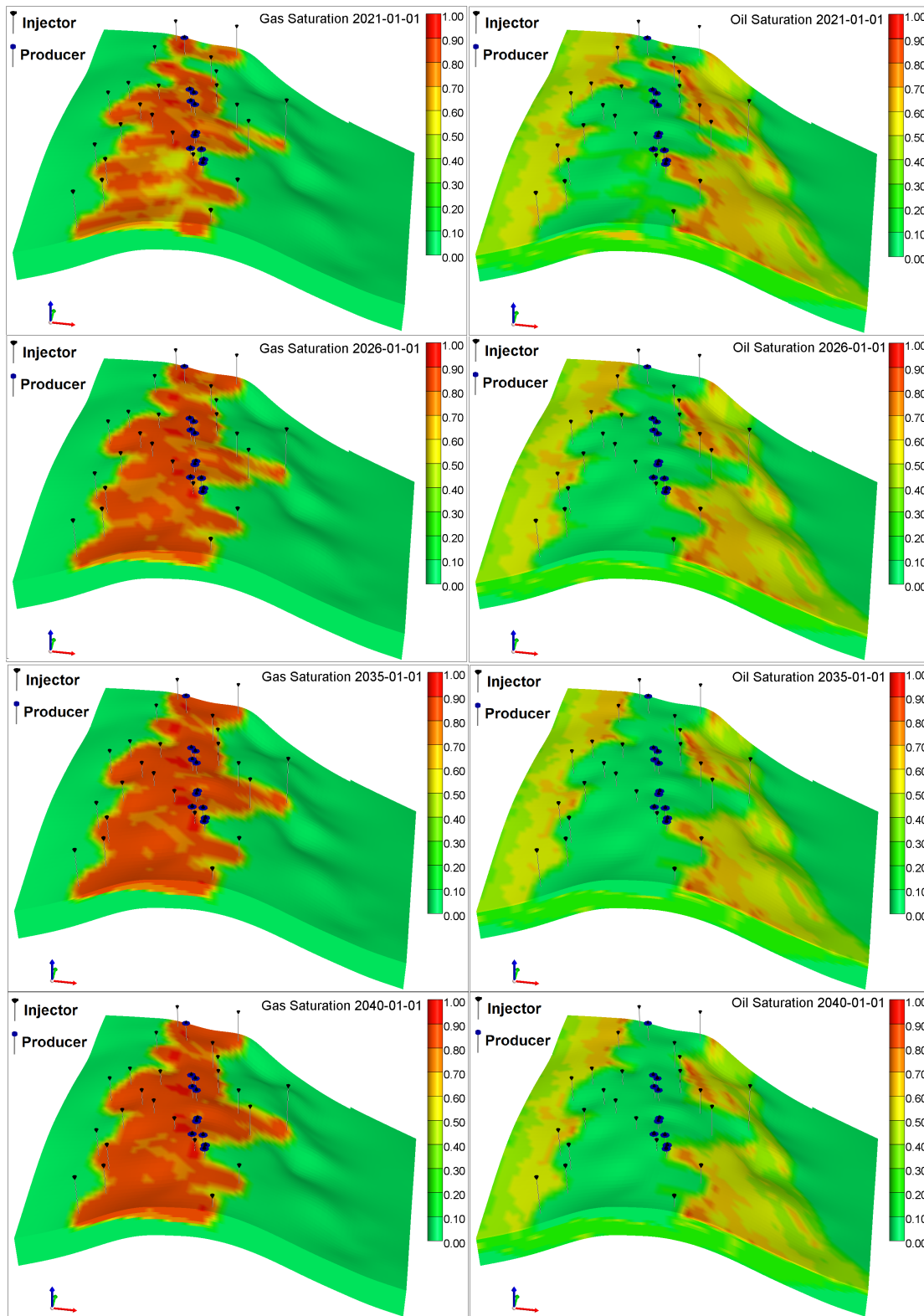


Figure 5.28: Gas and Oil Saturation Distributions in Immiscible Flooding (2021-2040)

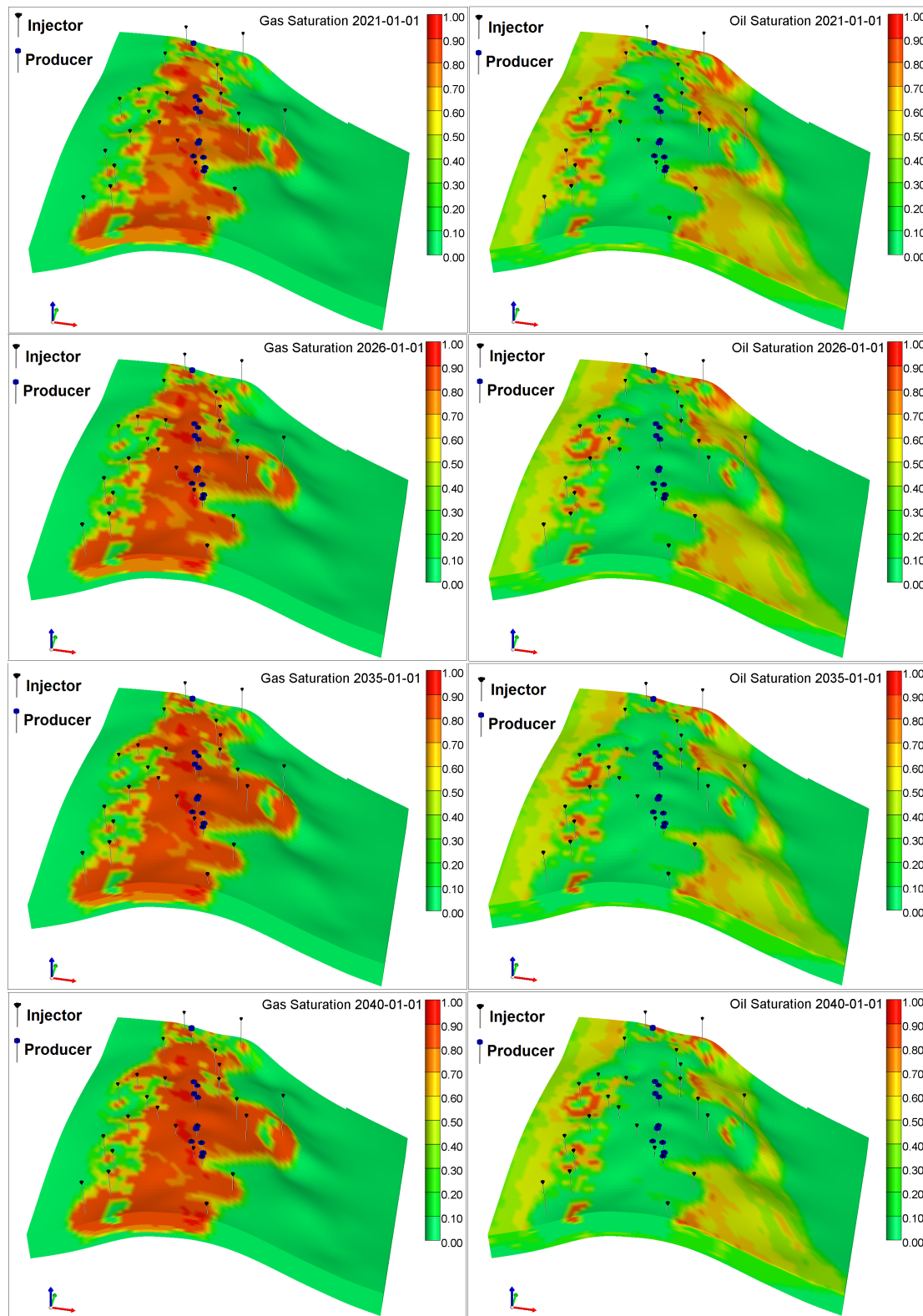


Figure 5.29: Gas and Oil Saturation Distributions in Miscible Flooding (2021-2040)

### 5.3.4 CGI, WAG, and GAGD Comparisons

A special case of the GAGD process has been set for comparison with the Continuous Gas Injection (CGI) and Water-Alternative-Gas (WAG) processes. The same reservoir with the same number of injection wells were considered in the setting of the three processes. Since the original base case of GAGD process has 11 horizontal producers, the number of vertical producers in the CGI and WAG was set to be 22 wells. That number was set because the production rate in horizontal wells is approximately twice the rate in vertical wells. The operational parameters in the three processes are outlined in Table 5.6. The

Table 5.6: CGI, WAG, and GAGD Operational Design Parameters

GAGD Process Minimum BHP in Horizontal Producers, psi	2660
Maximum STO in Horizontal Producers, STB/DAY	750000
Maximum Water Cut in Horizontal Producers	0.95
Maximum BHP in Injection Injectors, psi	3500
Maximum BHG in Injection Injectors, SCF/DAY	7500000
CGI Process Minimum BHP in Vertical Producers, psi	2660
Maximum STO in Vertical Producers, STB/DAY	100000
Maximum Water Cut in Vertical Producers	0.95
Maximum BHP in Injection Injectors, psi	3500
Maximum BHG in Injection Injectors, SCF/DAY	7500000
WAG Process Minimum BHP in Vertical Producers, psi	2660
Maximum STO in Vertical Producers, STB/DAY	100000
Maximum Water Cut in Vertical Producers	0.95
Maximum STW in Vertical Injectors, STB/DAY	10000
Maximum BHP in Injection Injectors, psi	3500
Maximum BHG in Injection Injectors, SCF/DAY	7500000

GAGD process has led to obtain higher cumulative oil production in 10 years prediction than the CGI and WAG processes. Figure 5.30 depicts the comparison between the CGI, WAG, and GAGD processes in terms of the 10 years of cumulative oil production.

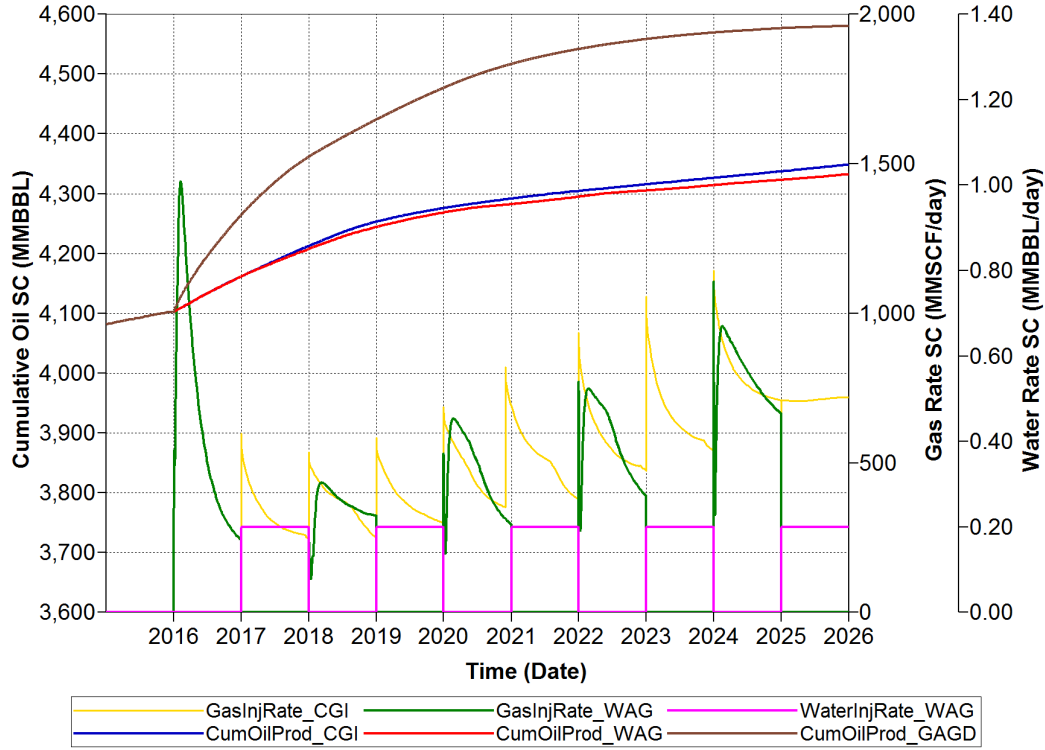


Figure 5.30: Comparison of Cumulative Oil Production between the CGI, WAG, and GAGD processes 10 Years Prediction Period

The cumulative oil production by the end of the 10th year of prediction period is 4.58 billions barrels of oil through the GAGD process. However, the CGI and WAG have resulted to obtain 4.35 and 4.33 billions of oil barrels, respectively. The WAG process is less effective than the CGI because water injection leads to increase the water saturation and impede the effect of gas injection to enhance oil recovery. The GAGD process has higher oil recovery than the other two processes as the horizontal producers lead to produce oil in much higher rates than the vertical ones. Consequently, the GAGD process is higher effectiveness to enhance oil recovery than the CGI and WAG processes.

In order to determine the most influential reservoir parameters on the GAGD process, a sensitivity analysis was conducted by incorporating different levels of permeability, anisotropy ratio ( $Kv/Kh$ ), and porosity into design of experiments to be evaluated in the compositional reservoir flow model. This is discussed in details in Chapter 6.

## 5.4 Summary and Significance

The fine-scale geostatistical lithofacies and petrophysical property model was upscaled and incorporated into the compositional reservoir model for the GAGD process simulation. The upscaling model has 69 grids in I-direction, 66 grids in J-direction, and 12 grids (layers) in K-direction. To validate the reservoir model, a history matching was conducted in terms of field and wells production and injection history. After obtaining an acceptable history match between the calculated and measured reservoir flow responses, the model was then used to evaluate the reservoir performance through the GAGD process and to predict the future reservoir performance.

Twenty vertical injectors were installed at the top two layers for CO<sub>2</sub> flooding. Then, the second, third, and fourth layers were left to provide a fluid movement transition and a gravity segregation domain between the injection and production zones. Next, the production layers are the fifth, sixth, seventh, and eighth layers. Through these four layers, twelve horizontal producers were placed with 3,000 m length. Finally, the four bottom layers were left with no injection/production activities as they approximately have full water flooding from the edge infinite aquifer.

The base case of the GAGD process with default setting of operational well design factors led to a significant increment in oil recovery in comparison with the primary production case. The cumulative oil production at the beginning of the prediction period is 4.1 billion bbls, while, by the end of this period through the primary production is 4.255 billions bbls. However, the GAGD process prediction resulted in obtaining approximately 4.4 billion bbls. So approximately a 155 million bbls of oil was obtained by the base GAGD case over the primary production case. The recovery factor in the base case is not more than 15%. In addition, the obtained amount of oil in 10 years primary production can be obtained in only one year by the GAGD base case. This outcome indicates the feasibility of the GAGD process to improve oil recovery in immiscible CO<sub>2</sub> flooding mode.

The immiscible base case of GAGD process showed there is no gas breakthrough in the

horizontal production wells as the injected gas did not reach the production zones at the middle of the reservoir. More specifically, the injected gas accumulated at the top of the reservoir leading to increase the average reservoir pressure, which supports the well flow pressure to be above bubble-point pressure. This eliminates the released gas from the dissolved gas through the vertical production tubing.

Two special cases were then implemented with a longer prediction period than the GAGD base case. These special cases represent the implementation of the GAGD process in immiscible and miscible for 25 prediction years. The recovery factor through immiscible and miscible special cases reach to 30% and 42%, correspondingly. More specifically, the cumulative oil production reach to 4.7 billion bbls and 4.93 billion bbls in immiscible and miscible injection modes, respectively. However, there is gas breakthrough seen during the early years of the prediction period in miscible mode as it was clearly identified by the GOR curve.

All three studied cases of the GAGD process evaluation discloses how the GAGD process increases oil recovery significantly in a real field-scale application.



## 6. Sensitivity Analysis

### 6.1 Introduction

Sensitivity analysis is an important step in the integrated reservoir simulation studies because it identifies the most influential geological parameters that impact the reservoir flow response. Identifying the influential geological parameters leads to fast automatic history matching by efficient manipulation the parameters that affect reducing the error between calculated and observed reservoir response.

In this chapter, Design of Experiments (DoE) was adopted for all sensitivity analysis cases to identify factors that influence the Gas Assisted Gravity Drainage (GAGD) Process performance in the main pay/South Rumaila oil field.

### 6.2 Design of Experiments

The Design of Experiments (DoE) is a systematic statistical tool that creates a proper set of experiments for simulation. DoE is used for the purpose of identifying the most sensitive factors that affect the response through the sensitivity analysis procedure. Furthermore, DoE helps an individual obtain the most-likely scenario that achieves the optimal response through a process. In addition, the DoE approach has the possibility to evaluate the interaction terms between some of all the factors in order to find out some combined influential roles in the process (Lazic, 2006). Since the designed experiments are typically faster, cheaper and more flexible than physical experiments, it is necessary to ensure achieving the most accurate model that mimics the physical model or process. To achieve that, the required numbers of factors and interactions should be analyzed to make the

interpretation and application of results correct and reliable (White and Royer, 2003).

The main terminologies in Design of Experiments are response variable that simply represents the outcome from an experiment or process. Factors are variable that affect the response variable. The values that a factor can assume is called levels. Primary and secondary factors are the variables that are most and less sensitive, respectively. The total number of designed experiments has an exponential formula. For instance, the number of experiments given  $k$  variables with 2 levels equal  $2^k$ . The sampling techniques, such as factorial design, one factor at a time, Box-Behnken design, etc, should be considered to combine multiple levels for each factor in a systematic procedure to create a population of observations. A common experimental design is to consider two levels for each factor. These levels are called 'low' and 'high' or '-1' and '+1', respectively. However, for three levels, they are called 'low', 'mid', and 'high' or '-1', '0', and '+1', respectively (Montgomery and Runger, 2003). Usually, full factorial design creates very large number of experiments for sensitivity analysis when there is more than five factors with three or more levels (Montgomery, 1997; Box et al., 2005; Yeten et al., 2005). In order to capture many levels of variation for each factor with minimum experiments, the sampling techniques provide limited data points through the design domain in a uniform distribution through the space-filling design (Bhat, 2001). The Latin Hypercube sampling is one of these efficient designs that produces uniform and low discrepancy observations (McKay et al., 1979).

Latin Hypercube sampling (LHS) was implemented in this study to determine the most influential reservoir factors affecting the Gas Assisted Gravity Drainage (GAGD) process performance. The response factor considered in the design of experiments is field cumulative oil production. Meanwhile, the factors that were included in DoE model are reservoir permeability, porosity, and vertical to horizontal permeability ratio ( $K_v/K_h$ ). The three reservoir factors are set into different approaches: general and specific. In the general approach, reservoir permeability, porosity, and ( $K_v/K_h$ ) were handled for the entire reservoir with three levels. While, the second approach includes setting the same factors in terms of

reservoir layers and the flow processes at these layers, for each property. Thus, two factors at the 1st and 2nd injection layers and one factor in the 3rd to 5th transition layers (where there is no injection or production wells). The 6th to 8th production layers were handled as three factors and the four bottom layers, which are completely flooded with water, was treated as one factor. The total number of factors for the three reservoir properties is 21.

### 6.2.1 Latin Hypercube Sampling Approach

Latin Hypercube Sampling (LHS) is a statistical sampling tool that is used to create samples from the input factors in order to construct many computer experiments from a multidimensional distribution (McKay, 1979). Selecting the random levels of each factor is based on the space-filling design. Therefore, it is able to simulate a process with minimum number of experiments to achieve a reasonably accurate random distribution. Figure 6.1 shows the space-filling design by LHS for two variables. In this figure, the sampled data are allocated within the entire space randomly. All the points are uniformly distributed to capture the entire variation of the process being studied, and that is the strength of Latin Hypercube Sampling approach.

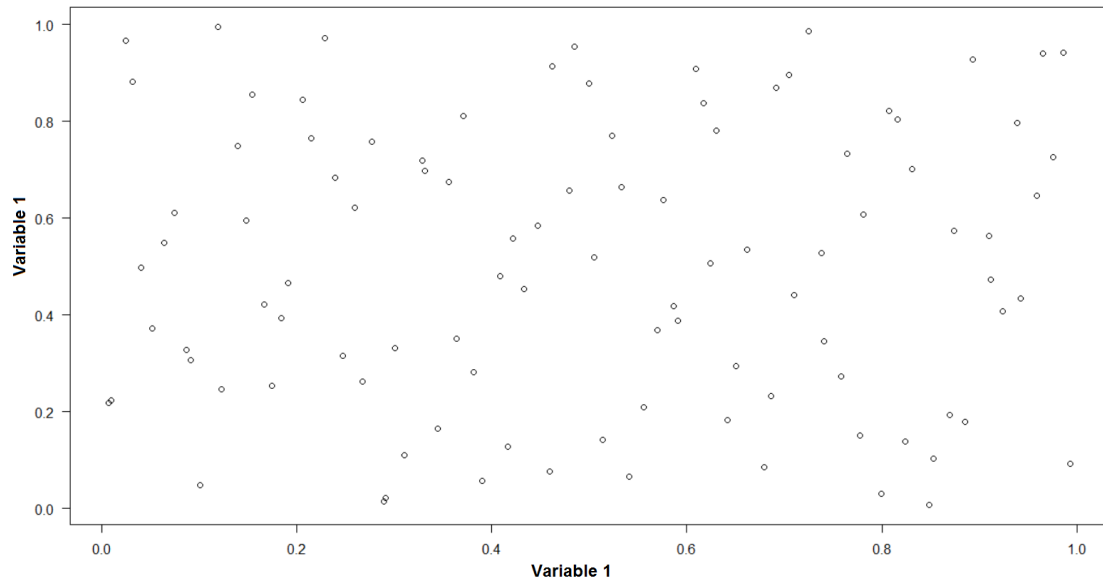


Figure 6.1: Latin Hypercube Sampling Design given Two Variables(R output)

Latin Hypercube sampling generates more efficient experiments for K parameters than simple Monte Carlo sampling. More specifically, LHS provides a regular spread points design because it keeps the maximum distance between each design point to all the other points (Stocki, 2005). Sampling K variables in LHS is performed by dividing each factor into many equal partitions. LHS is also an augmentation procedure that generates a new set of experiments in random manner if the original dataset does not represent the problem. There is no exact procedure to determine the number of experiments that can be created (Stein, 1987).

The computer experiments, which were generated for sensitivity analysis, optimization, and uncertainty assessments, were accomplished in R statistical language by `[lhs]` package (Carnell, 2015).

### 6.3 Model Selection

After evaluating all the designed experiments through obtaining the response factor values, the model is then fitted by linear regression analysis. The linear regression combines all the factors in a linear equation, as shown below:

$$y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \dots + \beta_p X_{pi} + \epsilon_i \quad (6.1)$$

where:

$y_i$ : the response factor.  $\beta_0, \beta_1, \beta_2, \dots, \beta_n$ : the coefficients of parameters  $n$ .  $X_1, X_2, \dots, X_n$ : the parameters  $X$ .  $\epsilon_i$ : the residual. Identifying the most influential factors is accomplished through eliminating the non-influential parameters from the model based on Analysis of Variance (ANOVA) test and probability of the null hypothesis. ANOVA is a statistical tool that is used to model the relationship between response and independent factors. It also used to determine the effect of each parameter on the response factor by computing the variance, which represents the division of sum of square on the degrees of freedom (Box et al., 2005). By ANOVA test, the regression model is statistically significant

if the probability of no-effect variables (p-value) is less than a significant level that justifies the acceptance or rejection of the null hypothesis. The F-value is the ratio of the explained variability ( $R^2$ ) by the unexplained variability ( $1 - R^2$ ), each divided by the corresponding degrees of freedom. The larger F statistic, the more useful the model. If a p-value is more than 0.05, then the probability tends to accept the null hypothesis and it does not perform the 95% confidence interval (Lazic, 2006).

The simplest way of modeling a subset selection in the regression model is to adopt the stepwise elimination which depends on the probability of null hypotheses. However, we adopt a new stochastic approach for modeling a subset selection, which is called Bayesian Model Averaging (BMA), that selects the most appropriate model for a given outcome variable based on Bayes factors. Based on its Bayesian procedure, BMA addresses the modeling uncertainty in the linear regression models.

After removing all non-influential parameters through obtaining the reduced linear model, The General Linear hypothesis Test (GLHT) is considered to compare between full (F) and reduced (R) model in order to decide if the reduced linear model is sufficient to represent the problem. Otherwise, the full model should be considered for analyzing the sensitivity analysis. The reduced linear model is sufficient when failing to reject the null hypotheses of the two models is achieved. Simply, the GLHT is implemented by calculating the F-value between the two models in order to obtain the p-value, as shown in Equation 6.2.

$$F = \frac{SSE(R) - SSE(F)}{df_R - df_F} / MSE(F) \sim F_{df_R - df_F, df_F} \quad (6.2)$$

where SSE is sum of squared error, MSE is mean square error, and df is degrees of freedom.

### 6.3.1 Stepwise Elimination

Backward and forward elimination are common procedures to eliminate the non-influential factors that have no effect on the response. Backward selection starts with the full model

that has all the factors and removes the non influential factors sequentially according to partial t-test when probability of null hypothesis when it exceeds 0.05. Forward selection starts with one variable that has the most significant effect based on partial t-test after fitting all variables independently. Other variables are then sequentially added to the model based on their significance and each remains in the model as long as it meets the criteria (Montgomery and Runger, 2003). Stepwise elimination is a forward selection with a backward glance. The variables are added to the model when they meet the partial t-test criteria. The model is then checked at each step to test if all the added variables are still meet the criteria. Otherwise, the sequentially added variables will be removed from the model (Montgomery and Runger, 2003).

### **6.3.2 Bayesian Model Averaging**

Bayesian Model Averaging (BMA) has been derived from bayesian model selection concept that considers the posterior probability distribution to choose the best subset model among a set of candidate models (Kass and Raftery, 1995; Raftery, 1995; Wasserman, 2000; Hoeting et al., 1999). BMA has been used in many different disciplines such as ranking the semivariogram through the spatial permeability process (Neuman, 2003; Ye et al., 2004). BMA has the advantage of modeling the data especially when the number of predictors is large ( $> 50$ ).

The model selection process in BMA considers the model's posterior probability and Bayesian Information Criterion (BIC). BMA produces a posterior distribution of the outcome factor that represents the weighted average of the posterior distributions of that factor for each likely model. Only five models are selected based on their posterior probabilities summation equals one. The best model selection has the maximum posterior probability and minimum Bayesian Information Criterion. Nevertheless, the model subset is determined when the probability of the predictor's coefficient is not zero, equal to 100% for all the sampled models. That means, this predictor has an effect on the outcome and should be included in a regression model. All these variables are shown in the computed Occam's

window. Let  $y$  is the dependent variable and  $X$  represents the explanatory variables, then:

$$y = \alpha_0 + \alpha_i X_i + \epsilon \quad \epsilon \sim N(0, \sigma^2 I). \quad (6.3)$$

BMA looks over the problem by estimating models for all possible combinations of predictors and constructing a weighted average over all of them. The model weights for this averaging stem from posterior model probabilities that arise from Bayes' theorem. Let the available data  $D$  along with multiple models  $M_k$  and for each model, there is a vector of parameters  $\theta_k$ . The posterior probability of a predicted quantity of interest  $\Delta$  can be obtained through the Bayesian Model Averaging based on the maximum probability criterion (Brown and Vannucci, 2002; Raftery, 1996) from Occam's window or MCMC model combination (MC3):

$$Pr(\Delta|D) = \sum_{k=1}^K Pr(\Delta|M_k, D) \times Pr(M_k|D) \quad (6.4)$$

where:

$M_k$ : multiple BMA models.  $\theta_k$ : vector of parameters in BMA.  $\Delta$ : quantity of interest in BMA.  $Pr(\Delta|D)$ : The posterior distribution of a quantity given the observation data,  $D$ .

$Pr(\Delta|M_k, D)$ : The conditional probability of the predicted quantity given the data,  $D$ .

$Pr(M_k|D)$ : The posterior distribution for each model  $M_k$  conditioning to the data  $D$  and it is given by:

$$Pr(M_k|D) = \frac{Pr(D|M_k)Pr(M_k)}{\sum_{k=1}^K Pr(D|M_k)Pr(M_k)} \quad (6.5)$$

where the likelihood of the data  $Pr(D|M_k)$  is given by:

$$Pr(D|M_k) = \int Pr(D|\theta_k, M_k)Pr(\theta_k|M_k)d\theta_k. \quad (6.6)$$

The maximum likelihood for the data is  $Pr(D|\theta_k, M_k)$  and the prior distribution for the parameters in the model is  $Pr(\theta_k|M_k)$ . The prior probability of the model is  $Pr(M_k)$ . Occam's window, is considered in this study, or Markov chain Monte Carlo model composition (MC3) are used as practical ways to implement the BMA procedure. In Occam's window, the model is neglected if there is other models have better prediction:

$$A' = \left\{ M_k : \frac{\max_k \{Pr(M_k|D)\}}{Pr(M_k|D)} \right\}. \quad (6.7)$$

where  $A'$  refers to the neglected model indication.

The best candidate model is selected based on the model's posterior probability distribution and Bayesian Information Criterion (BIC) that are computed and shown in Occam's window or Markov Chain Monte Carlo model composition (MC3) (Hoeting, 1999). The BIC was designed to find the most probable model and it is a rough estimate of logarithmic function of Bayes Factor:

$$BIC = -\log \pi(y|\hat{\theta}) + p \log(n) \quad (6.8)$$

where  $BIC$  represents the Bayesian Information Criterion,  $\hat{\theta}$  is the maximum likelihood estimate and  $\pi(y|\hat{\theta})$ , the Bayes Factor, gives a measure of evidence for models comparison.

Combining the predicted posterior probability and BIC is necessary to handle the likelihood of over-fitting problem because BIC works as penalty term for the number of parameters in the model (Schwarz, 1978). After ranking all the sampled models, only the five best models are selected when their posterior probabilities summation equals one. The optimal model has the maximum posterior probability, and minimum BIC. In this study, BMA is accomplished in the R statistical computation programming language through *BMA* package (Raftery et al., 2010). Figure 6.2 illustrates the posterior distribution for the coefficient given the data. However, the spike corresponds to probability of equal zero coefficient (Hoeting et al., 1999).



Finally, the novelty of Bayesian Model Averaging comes from its stochastic design to

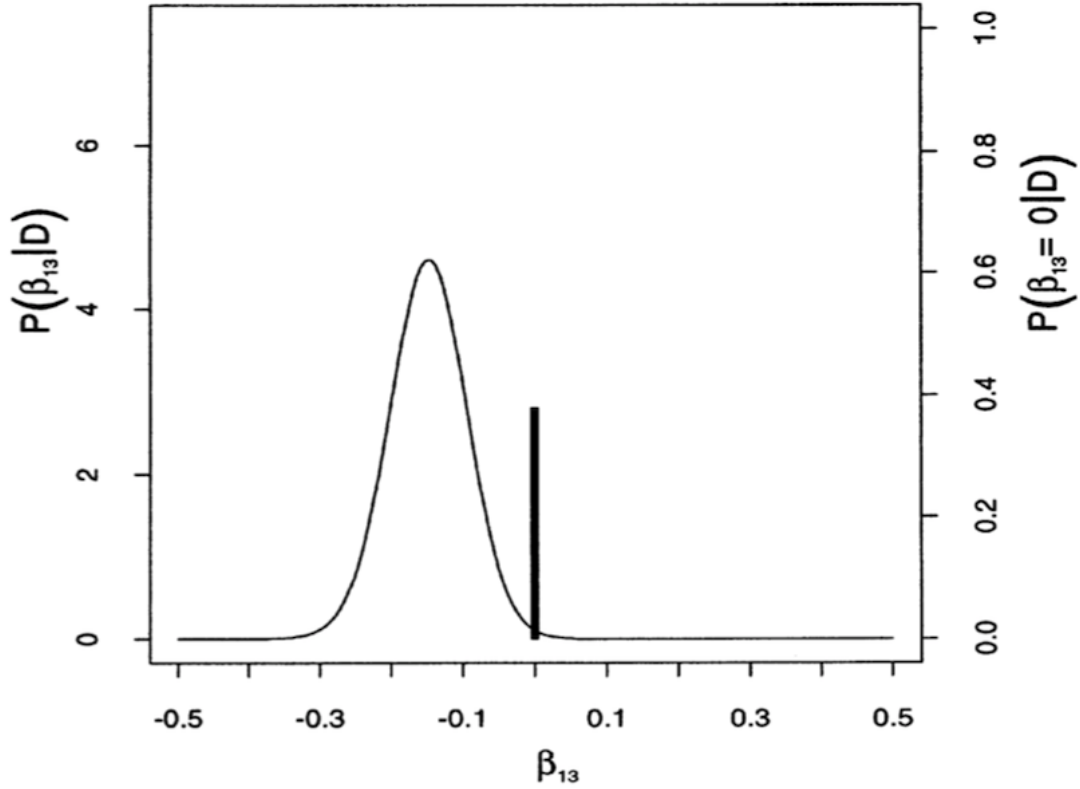


Figure 6.2: The left vertical axis corresponds to the posterior distribution for  $\beta_{13}$  and the right vertical axis represents the posterior distribution for  $\beta$  equal to 0. The spike corresponds to  $P(\beta = 0|D) = 0.38$  (Hoeting et al., 1999).

generate a better fit between the dependent and independent variables by providing less uncertain space of solution via multiple models, rather than one deterministic solution in linear regression (Raftery et al., 1997).

## 6.4 Results and Discussion

There are many uncertain parameters in a reservoir simulation study that need to be quantified such that the uncertainty is decreased. Identifying the most influential factors is an important step in integrated reservoir studies to facilitate history matching and optimization process. The Design of Experiments (DoE) that generates many computer experimental runs through combining the response and expected sensitive factors is the common way to perform sensitivity analysis. In DoE, Analysis of Variance (ANOVA) and partial t-test consider the null hypothesis to eliminate the non influential factors that are not affecting the response (White and Royer, 2003).

The sensitivity analysis was implemented based on two different cases of reservoir parameters setting. The 1st one is a general case that represents a screening for the most sensitive geological factors impacting the GAGD process. The general case has only three main factors: permeability,  $K_v/K_h$  and porosity for all the layers of the reservoir. On the other hand, the second sensitivity analysis case is more specific than the 1st one as it handles the sensitivity analysis in the matter of the separate layers not the entire reservoir.

The Latin Hypercube Sampling was adopted as an efficient Experimental Design approach to create more uniform simulation observations for the two sensitivity analysis cases to capture the maximum variety of the sampled points.

### 6.4.1 General Reservoir Sensitivity Case

In order to determine the most influential geological parameters affecting the Gas Assisted Gravity Drainage process, sensitivity analysis was conducted through Design of experiments (DoE). DoE combines multi-level of each parameter to create many computer experiments evaluated by the compositional reservoir simulation. The factors that were adopted in this case are horizontal permeability, anisotropy ratio  $K_v/K_h$ , and porosity, all given for the entire reservoir. The most influential factors are diagnosed through building a statistical linear modeling, partial t-test, and Analysis of Variance (ANOVA).

Firstly, more than 80 computer experiments (simulation job) were created by the Latin Hypercube Sampling approach. The simulation jobs were then evaluated through the reservoir simulator to calculate the reservoir flow responses, as shown in Figure 6.3. The figure shows the field cumulative oil production (top) and oil production rates (bottom) for the forecasting period (2016-2026).

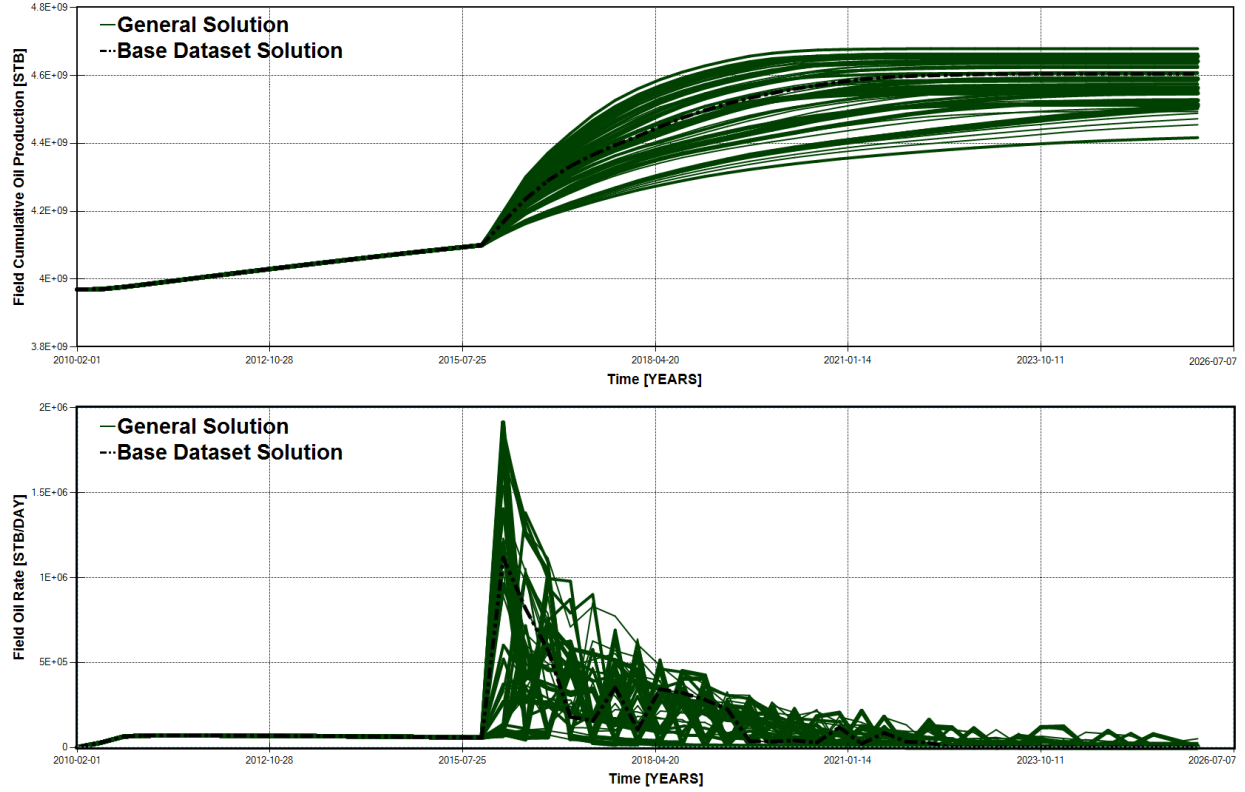


Figure 6.3: Sensitivity Analysis of Reservoir Parameters on Field Flow Response

This fact can be likewise identified in Figure 6.4, which depicts how the reservoir performance, with reference to field cumulative oil production, responds to the different computer experiments. All simulation jobs of different factors' setting led to distinct flow responses.

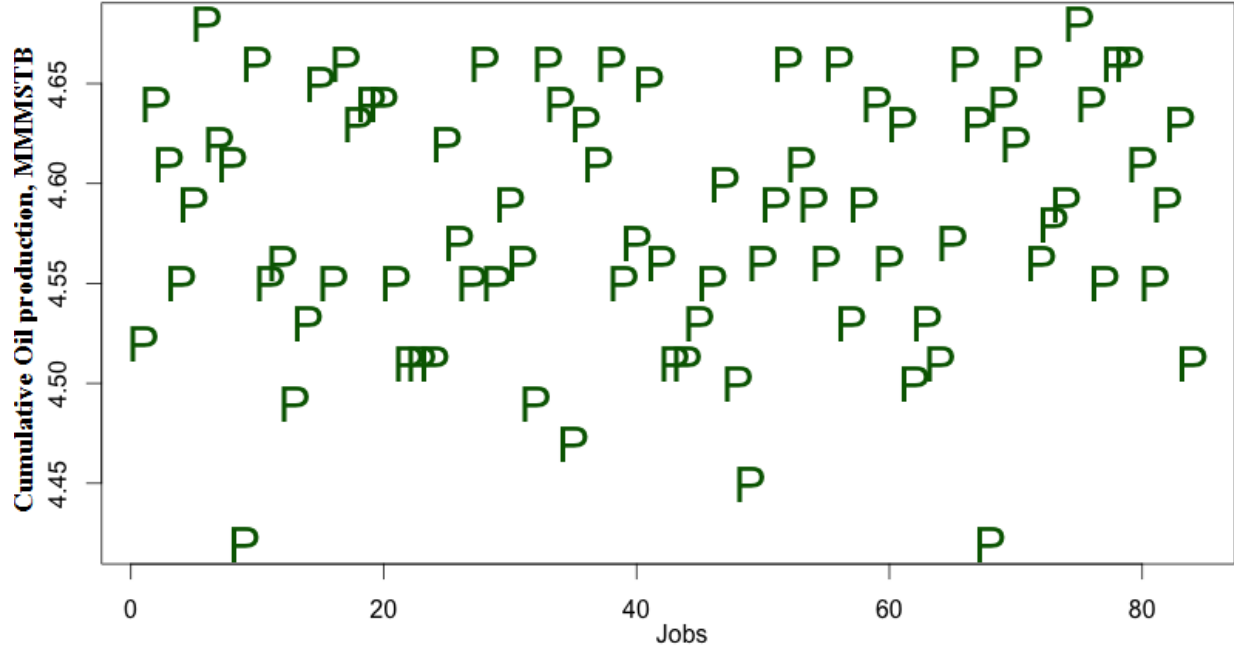


Figure 6.4: Field Cumulative Oil Production of the Designed Experiments

Given these generated experiments, the full linear model fits all the aforementioned three parameters to the response factor, field cumulative oil production. To determine the non-influential factors, partial t-test and AVOVA tables were constructed as shown in Table 6.1 and Table 6.2, respectively.

From these two tables, the porosity should not be kept in the linear model as it failed

Table 6.1: Partial t-Test of Full Linear Model

	Estimate	Std. Error	t value	Pr(> t )	
(Intercept)	4.415505	0.009004	490.38	$< 2e - 16$	***
Kv/Kh	0.058832	0.004429	13.28	$< 2e - 16$	***
Permeability_I	0.083872	0.004309	19.46	$< 2e - 16$	***
Porosity	0.005343	0.004268	1.25	0.2143	
Adjusted R-squared=	0.8615	Model p-value=	$< 2.2e - 16$		

to reject the null hypotheses of no impact assumption ( $p\text{-value} > 0.05$ ). That means the horizontal permeability and anisotropy ratio are the only influential factors affecting the reservoir performance through the GAGD process. This fact can be easily observed in

Table 6.2: ANOVA Test of Full Linear Model						
	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
Kv/Kh	1	0.073635	0.073635	139.12	$< 2e - 16$	***
Permeability_I	1	0.200491	0.200491	378.80	$< 2e - 16$	***
Porosity	1	0.000829	0.000829	1.57	0.2143	
Residuals	80	0.042343	0.000529			

Figure 6.5 that decorates the component-residual plots for each factor. If a variable with horizontal straight line, it has no effect on the response and should be removed form the linear model, specifically the porosity variable.

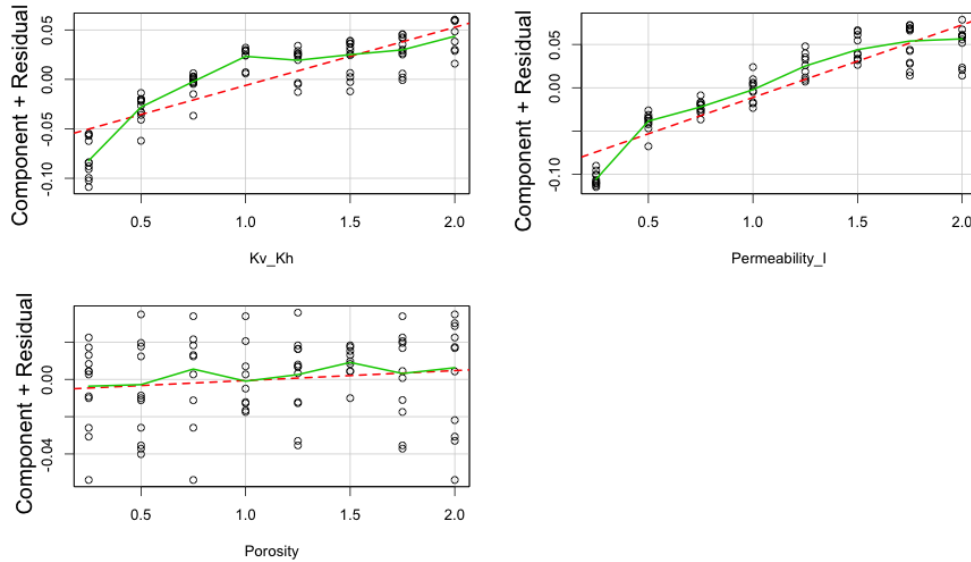


Figure 6.5: Component-Residual Plots of Influential Factors

As shown above, the reservoir porosity factor has no effect on the flow response, even though the full linear model, which includes all the factors, fit looks suitable for this case because there is approximate matching between theoretical and standardized residual curves, as illustrated in Figure 6.6 that shows the basic diagnostic plots for the full linear model.

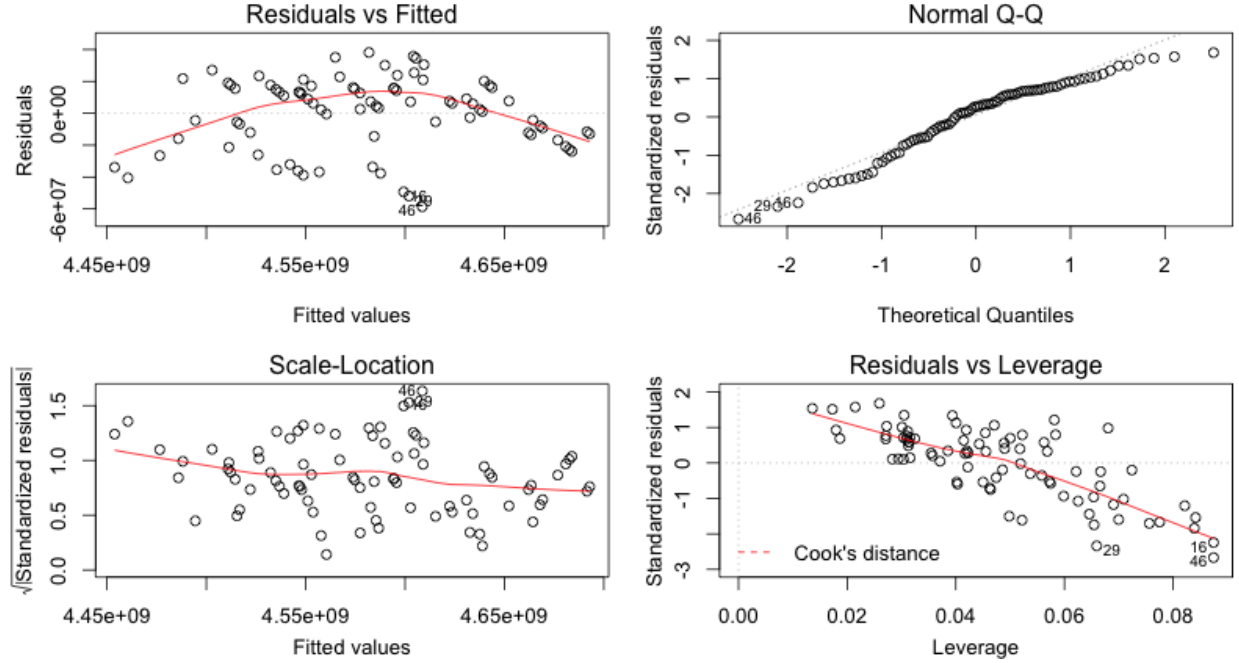


Figure 6.6: Basic Diagnostic Plots of Full Linear Model

To statistically remove the non-influential factor (s), stepwise elimination was employed to eliminate the non-influential variable(s). As expected, stepwise elimination led to remove the porosity parameter from the model as it has no effect on the flow response and it rejects the null hypothesis. Table 6.3 and 6.4 illustrate the partial t-test and ANOVA tables for the reduced linear model after removing the porosity variable, respectively. All the variables in the reduced linear model reject the null hypothesis of no impact assumption, therefore, they should not be removed from the model.

To verify this result, the component-residual plots for the reduced linear model were decorated in Figure 6.7 that shows how the variation of the remained two factors increased as they split the variation of the entire model.

Table 6.3: Partial t-Test of Reduced Linear Model

	Estimate	Std. Error	t value	Pr(> t )	
(Intercept)	4421543	0.007629	579.53	$< 2e - 16$	***
Kv/Kh	0.058786	0.004444	13.23	$< 2e - 16$	***
Permeability_I	0.083863	0.004324	19.39	$< 2e - 16$	***
Adjusted R-squared= 0.8606    Model p-value= $< 2.2e - 16$					

Table 6.4: ANOVA Test of Reduced Linear Model

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
Kv_Kh	1	0.073635	0.073635	138.16	$< 2e - 16$	***
Permeability_I	1	0.200491	0.200491	376.17	$< 2e - 16$	***
Residuals	81	0.043172	0.000533			

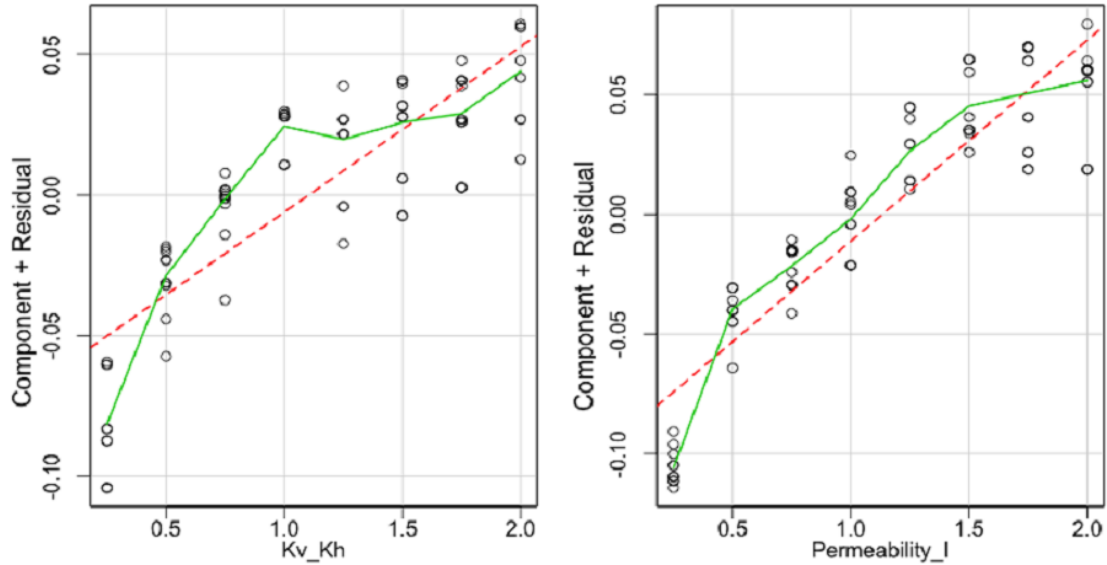


Figure 6.7: Component-Residual Plots of Reduced Influential Factors

The basic diagnostic plots for reduced linear model were shown in Figure 6.8. There is good matching between theoretical and standardized residual curves in the normal Q-Q plot. That reflects the eligibility of the linear model to simulate the given data.

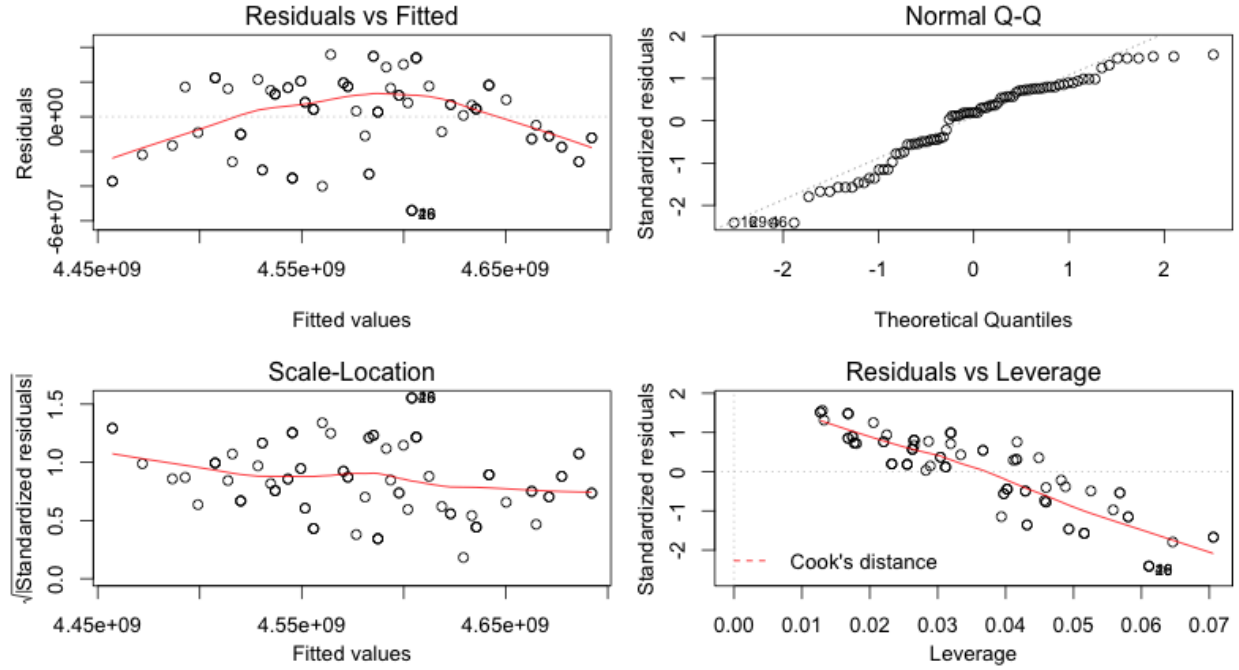


Figure 6.8: Basic Diagnostic Plots of Reduced Linear Model

To show that the reduced linear model is sufficient to represent the problem without need to consider the full model, General Linear hypothesis Test (GLHT) was computed to compare between the two models. The null hypothesis in GLHT declares the full and linear model are similar and no need to select the full model ( $p\text{-value} > 0.05$ ). In Table 6.5, the  $p\text{-value}$  is 0.2143 and there is failing to reject null hypothesis. Accordingly, the reduced model is sufficient to be considered for final analysis. There is a hypothesis that declares

Table 6.5: GLHT of the Full and Reduced Models						
	Res.Df	RSS	Df	Sum of Sq	F	Pr(>F)
1	80	0.042343				
2	81	0.043172	-1	-0.00082939	1.567	0.2143

porosity should have an effect on the reservoir flow response. However, this hypothesis was rejected as the reservoir porosity does concern with the fluid storage not flow through porous media.



In order to ensure that there is no doubt about the validity and accuracy of the fitted models, Bootstrapping can be a very useful tool to approve that. Bootstrapping is a nonparametric statistical method that computes the estimated standard errors, confidence intervals and hypothesis testing. It simply re-samples the data into a specific number of times and recalculate the variable given each sample. The confidence intervals can be then computed from the distribution of that variable. The bootstrapping was conducted in this case for Adjusted R-squared and the coefficients of predictors. Figure 6.9 and Figure 6.10 decorate the histograms for Adjusted R-squared and coefficients of factors for the full model, receptively. On the other hand, Figure 6.11 shows the histogram for Adjusted R-squared of the reduced linear model.

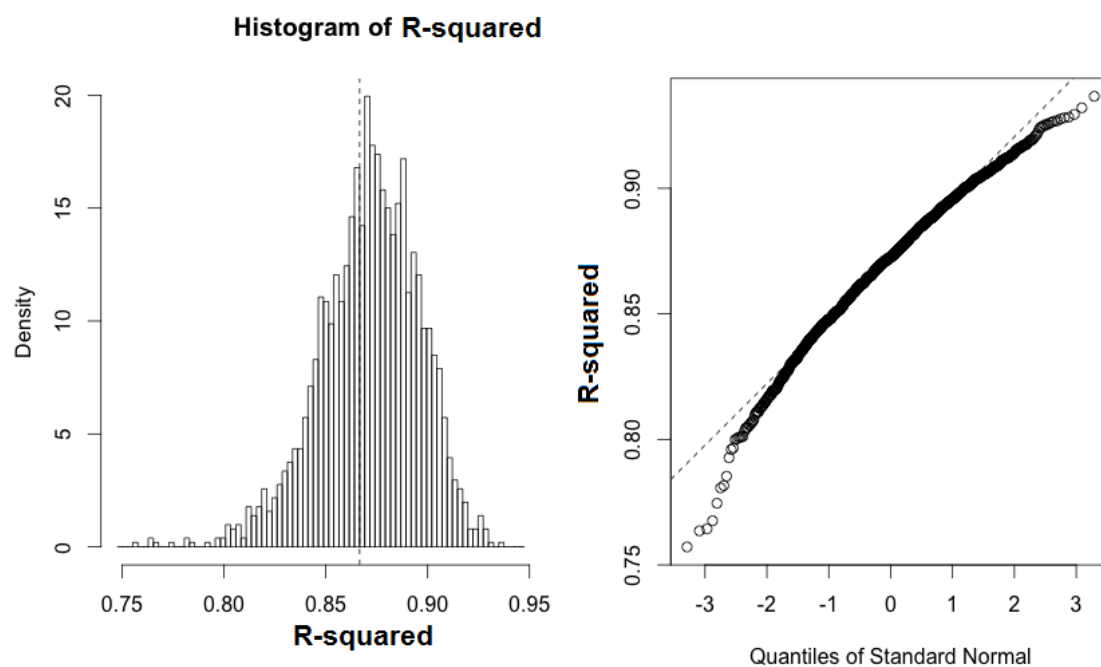


Figure 6.9: Nonparametric Bootstrapping of R-squared in Full Linear Model

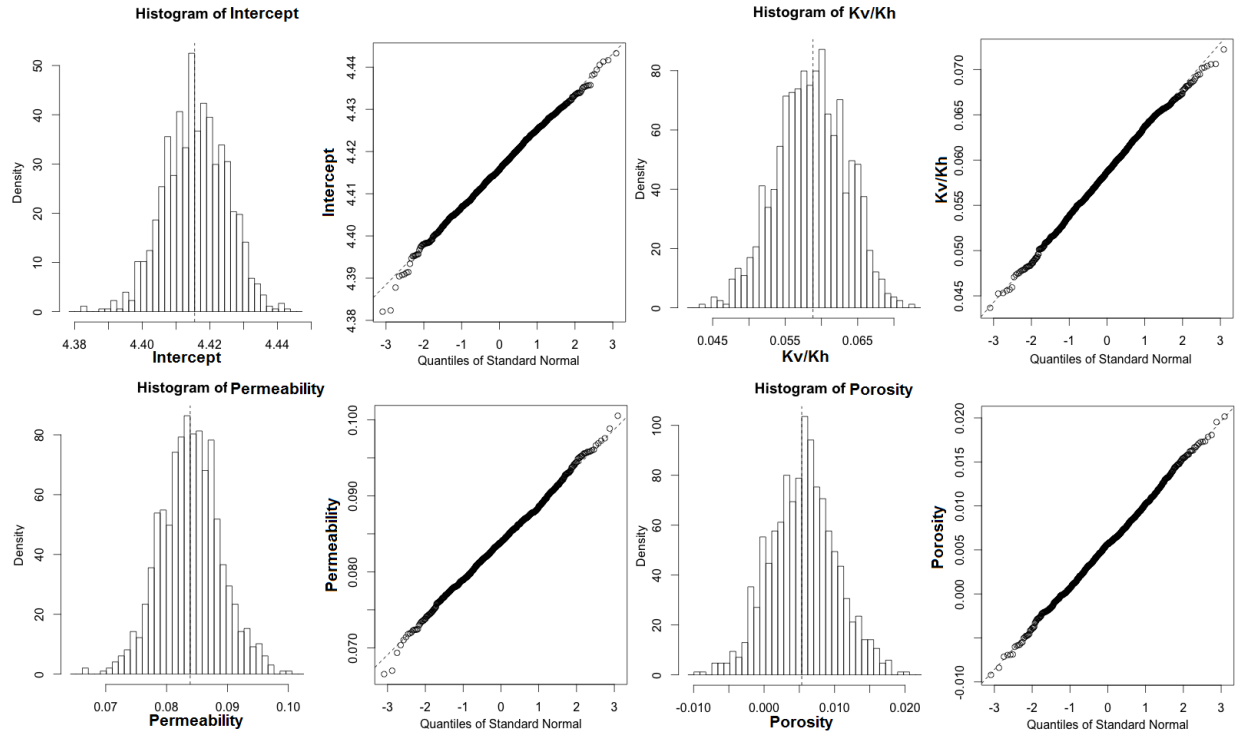


Figure 6.10: Nonparametric Bootstrapping of Coefficients in Full Linear Model

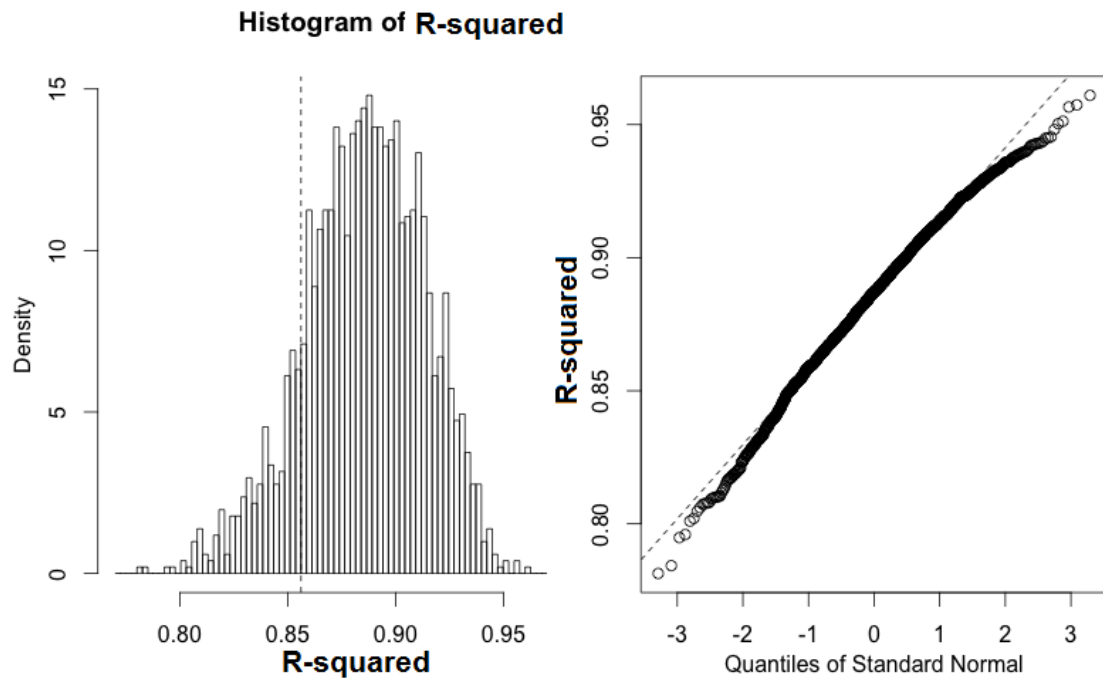


Figure 6.11: Nonparametric Bootstrapping of R-squared in Reduced Linear Model

From these two figures, approximate full normal distribution for all the quantities along with the excellent matching between the theoretical and standardized residual curves in the related normal Q-Q plot. That reflects the accuracy of the full and reduced linear models to quantify the given problem of GAGD process evaluation.

Finally, Figure 6.12 illustrates the effect diagrams of the reservoir parameters in full and reduced linear model.

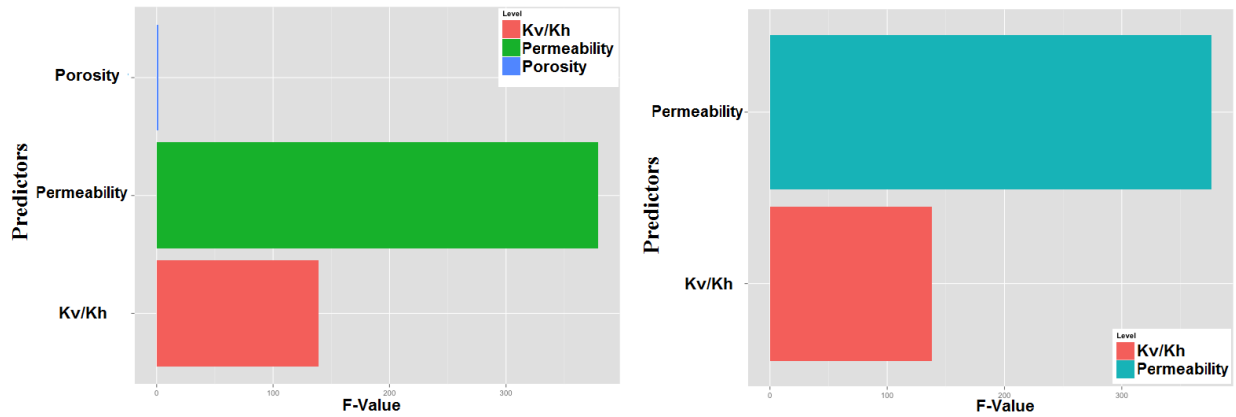


Figure 6.12: Influence of Reservoir Parameters on the GAGD Process Performance

### Validation of Influential Factors-General Case

To demonstrate that permeability and anisotropy ratio have large effect on the flow response through the GAGD process with no effect for the reservoir porosity, many levels, ranged from less to most likely, were set for each of the three parameters to regenerate multiple observations (simulation jobs). These observations were then evaluated through the reservoir simulator to check their effects on the reservoir flow response, particularly field cumulative oil production, oil rate, and water cut. In Figures 6.13 and 6.14, a significant difference can be seen in the three reservoir flow responses between the generated experiments for permeability and anisotropy ratio.

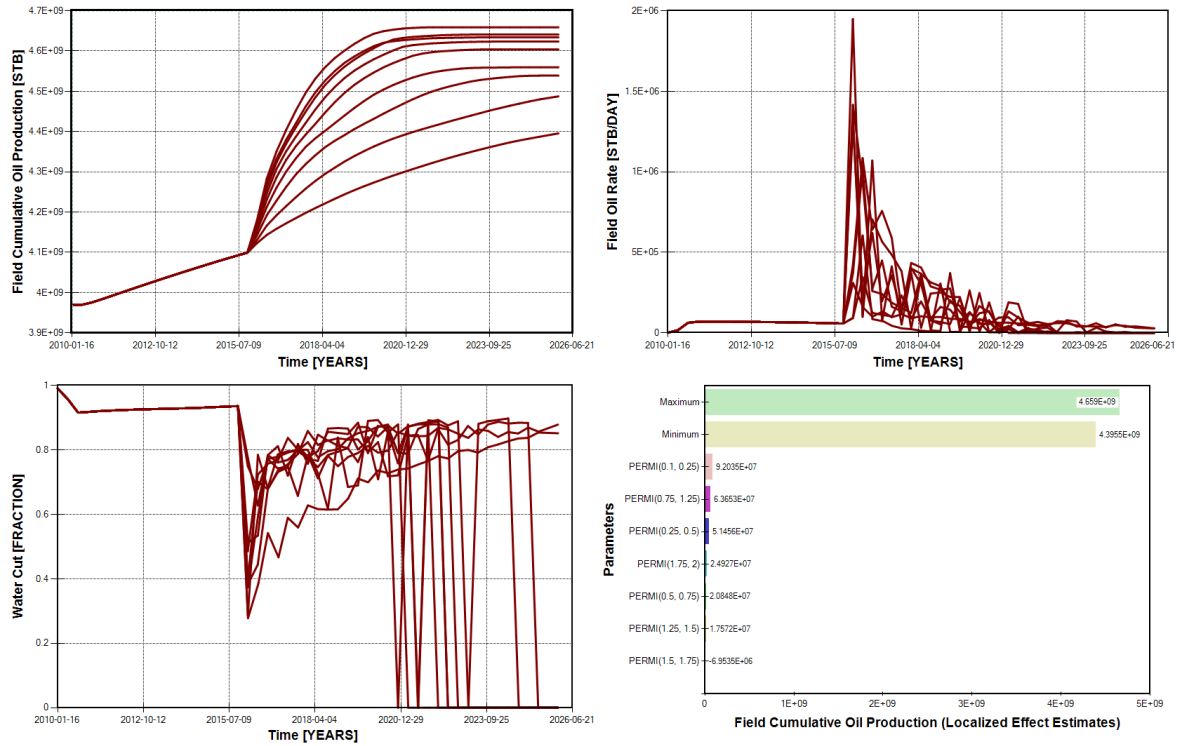


Figure 6.13: Effect of Reservoir Permeability on the Field Flow Response

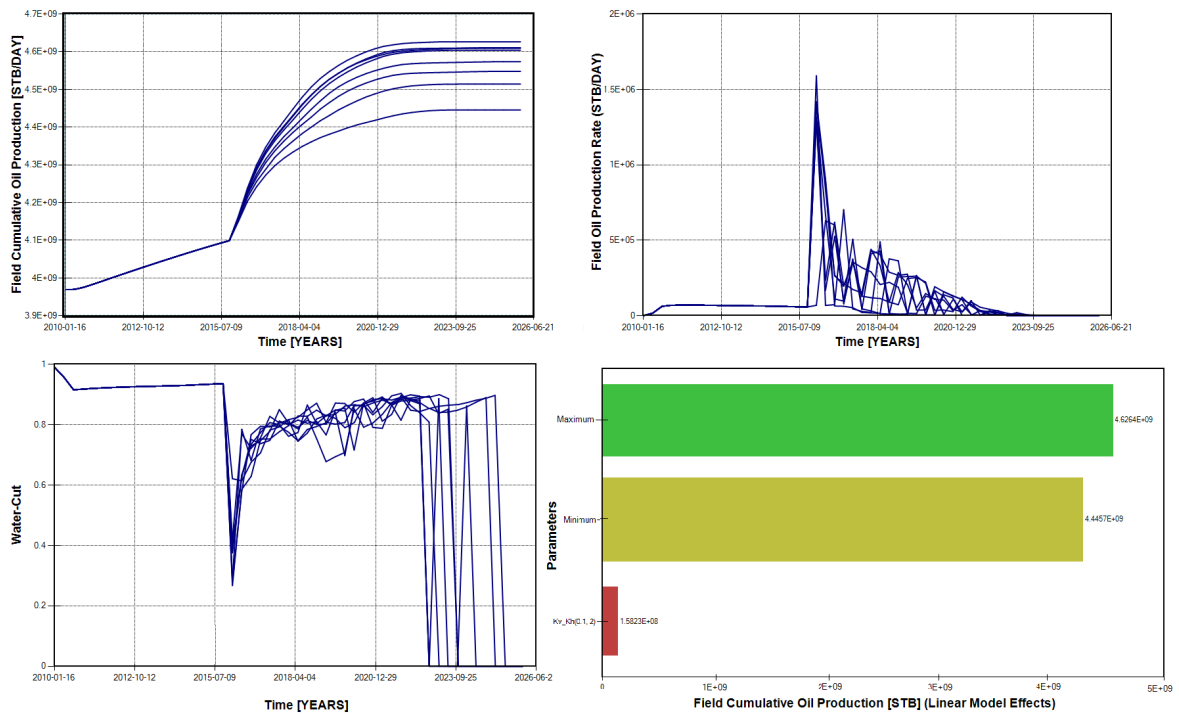


Figure 6.14: Effect of  $K_v/K_h$  on the Field Flow Response

In a different manner, the reservoir porosity is seen to have no effect on the responses as all the experiments have the same values of field cumulative oil production, oil rate, and water cut, as depicted in Figure 6.15.

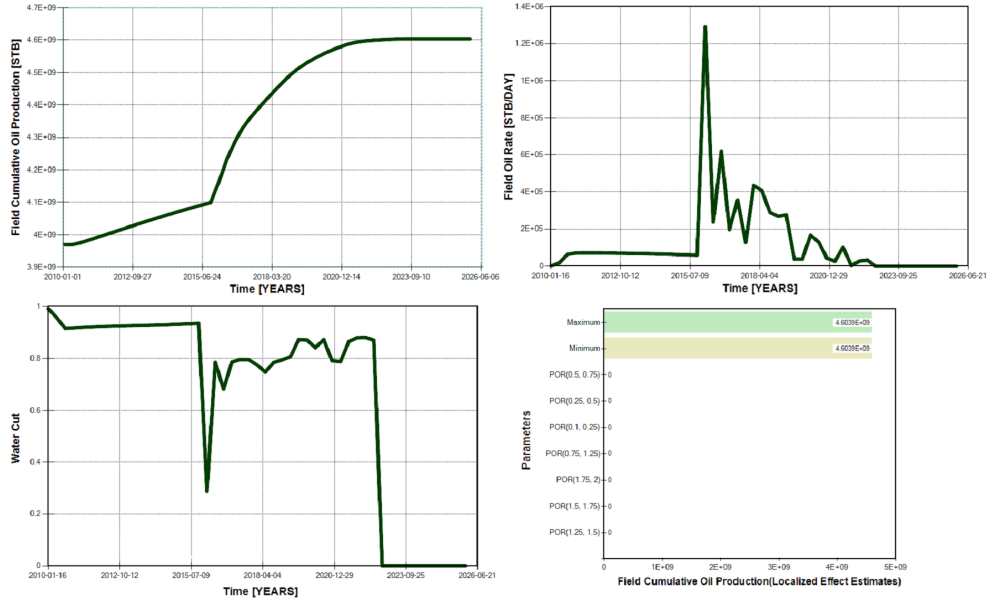


Figure 6.15: Effect of Reservoir Porosity on the Field Flow Response

#### 6.4.2 Specific Reservoir Sensitivity Case

The sensitivity analysis in this case was conducted through investigating the effects of reservoir properties given the respected layers and not for the entire reservoir. The reservoir parameters considered for sensitivity analysis are horizontal permeability, anisotropy ratio ( $K_v/K_h$ ), and porosity. There are 12 reservoir layers classified into four groups: injection (two layers), transition (three layers), production (three layers), and the bottom below-production layers (four layers). Each layer in the injection and production zones were handled separately in the analysis. However, the transition and bottom layers were treated as single zone each. Thus, the total number of factors that were checked for their effect on the GAGD process is 21. Table 6.6 illustrates the 21 factors that were considered for the specific sensitivity analysis case through stepwise elimination and bayesian model averaging approaches. Comparison determining the most sensitive factors impacting the GAGD pro-

Table 6.6: The 21 Factors Considered in Specific Sensitivity Analysis Case

Permeability in the Injection Layer 1	PERMIINJ1
Permeability in the Injection Layer 2	PERMIINJ2
Permeability in the Transition Layers	PERMITRAN
Permeability in the Production Layer 6	PERMIPROD6
Permeability in the Production Layer 7	PERMIPROD7
Permeability in the Production Layer 8	PERMIPROD8
Permeability in the Below-Production Layers	PERMIBOTM
Anisotropy Ratio in the Injection Layer 1	Kv/KhINJ1
Anisotropy Ratio in the Injection Layer 2	Kv/KhINJ2
Anisotropy Ratio in the Transition Layers	Kv/KhTRAN
Anisotropy Ratio in the Production Layer 6	Kv/KhPROD6
Anisotropy Ratio in the Production Layer 7	Kv/KhPROD7
Anisotropy Ratio in the Production Layer 8	Kv/KhPROD8
Anisotropy Ratio in the Below-Production Layers	Kv/KhBOTM
Porosity in the Injection Layer 1	PORINJ1
Porosity in the Injection Layer 2	PORINJ2
Porosity in the Transition Layers	PORTRAN
Porosity in the Production Layer 6	PORPROD6
Porosity in the Production Layer 7	PORPROD7
Porosity in the Production Layer 8	PORPROD8
Porosity in the Below-Production Layers	PORBOTM

cess was accomplished on this case through stepwise elimination as a deterministic method and Bayesian model Averaging as a stochastic approach. In nonlinear geological system, it is difficult to have a deterministic linear model to generate a linear relationship between a response factor and a high number of independent variables. Accordingly, Bayesian Model Averaging (BMA) was adopted as a stochastic framework to reduce the uncertainty in the influential factors' selection. Bayesian Model Averaging transforms every independent variable (predictor) into a normal distribution to provide a basis for sampling and to generate several models through generalized linear modeling. BMA adopts posterior probability to rank the generated models and to select the best model, which attains the lowest Bayesian Information Criterion (BIC) and largest posterior probability. Low BIC values are preferred because it provides a penalty term for the number of parameters in the model in order not to fall into over-fitting (Schwarz, 1978).

First, the full linear model fits all the 21 reservoir factors to the reservoir flow response, field cumulative oil production. Then, sufficient experiments were generated by combining the levels of all factors through Latin Hypercube Sampling approach. Next, the generated experiments were evaluated through the compositional reservoir simulator to calculate field oil rate and cumulative oil production for the prediction period (2016-2026), as decorated in Figure 6.16. In this figure, the different experiments that combined various levels of factors have led to assorted values of the last-mentioned reservoir flow responses.

In particular, this fact can further be seen in Figure 6.17 that illuminates the ranges of field cumulative oil production values given all these experiments (simulation jobs).

Subsequently, the partial t-test, and Analysis of Variance (ANOVA) were adopted to dispose the non-influential factors through the stepwise elimination approach. Table 6.7 and Table 6.8 expose the partial t-test and AVOVA tables for the full linear model that have all factors, appropriately.

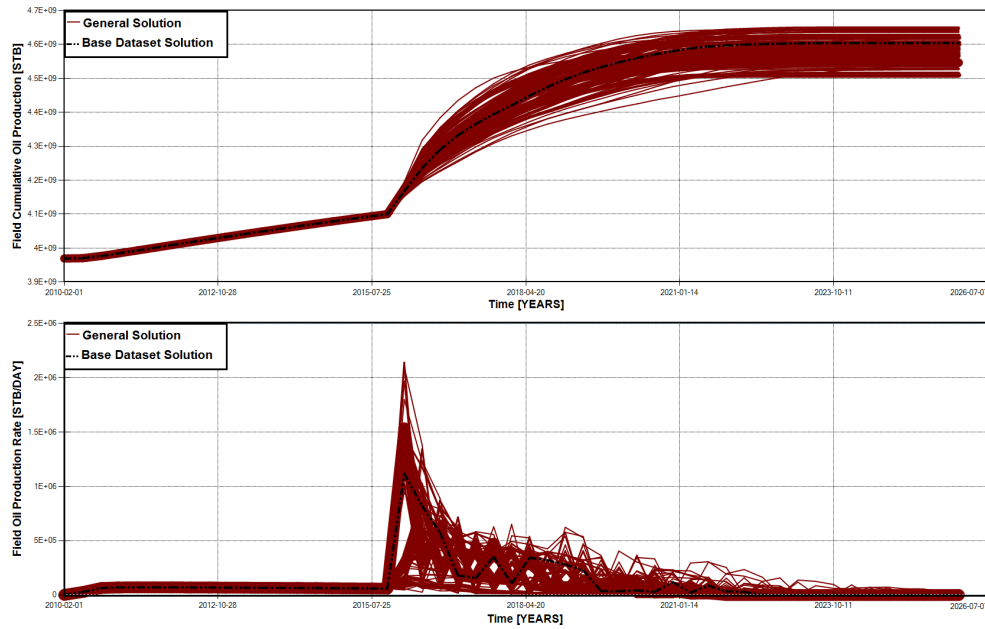


Figure 6.16: Sensitivity Analysis of Reservoir Parameters on Field Flow Response

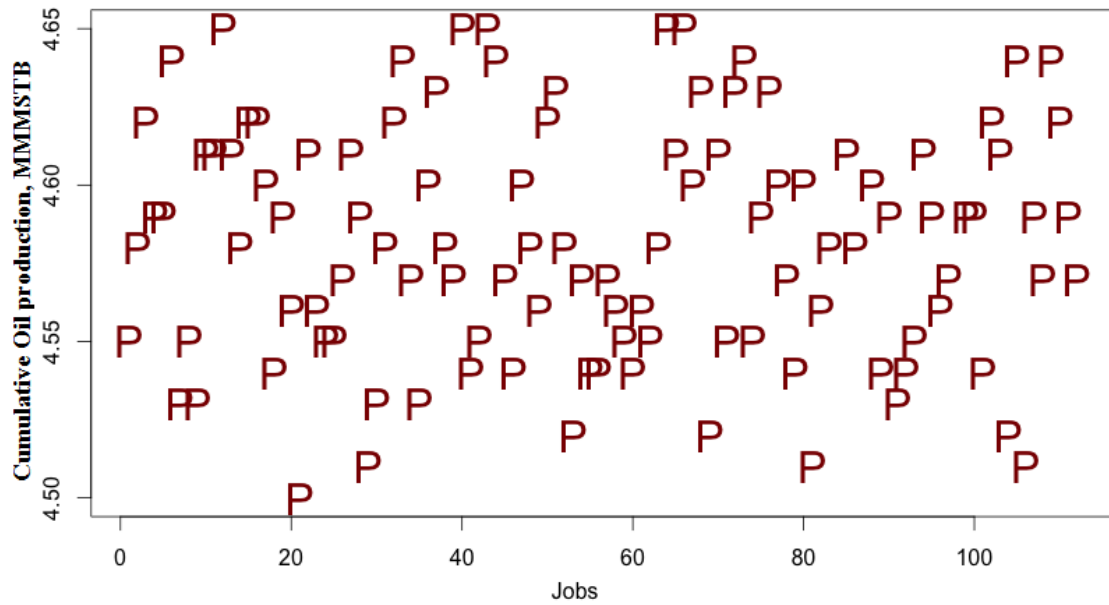


Figure 6.17: Field Cumulative Oil Production of the Designed Experiments

From first glance on the aforementioned two tables, one can notice there are many factors failed to reject null hypotheses of non-effect assumption. More specifically, many non-influential factors with p-value greater than 0.05 need to be removed from the full model.

That fact was even more clarified in Figure 6.18 that shows the tornado plot of all parameters' effect, in full linear model, on the response.



Table 6.7: Partial t-Test of Full Linear Model

	Estimate	Std. Error	t value	Pr(> t )	
(Intercept)	4.4499	0.0136	327.50	$< 2e - 16$	***
PERMIINJ1	0.0255	0.0026	9.93	4.02e-16	***
PERMIINJ2	0.0193	0.0027	7.14	9.85e-10	***
PERMITRAN	-0.0233	0.0027	-8.53	7.82e-13	***
PERMIPROD6	0.0106	0.0026	4.08	4.242-5	***
PERMIPROD7	0.0103	0.0026	3.97	0.000179	***
PERMIPROD8	0.0105	0.0027	3.88	0.000382	***
PERMIBOTM	0.0267	0.0028	9.66	8.22e-16	***
Kv/KhINJ1	-0.0005	0.0027	-0.20	0.8424	
Kv/KhINJ2	-0.0014	0.0027	-0.52	0.6066	
Kv/KhTRAN	0.0039	0.0028	1.43	0.1552	
Kv/KhPROD6	0.0123	0.0026	4.75	4.29e-6	***
Kv/KhPROD7	0.0095	0.0029	3.35	0.0012	
Kv/KhPROD8	0.0109	0.0029	3.74	0.0003	
Kv/KhBOTM	0.0066	0.0026	2.53	0.0133	
PORINJ1	0.0002	0.0027	0.06	0.9526	
PORINJ2	-0.0021	0.0026	-0.80	0.4278	
PORTRAN	0.0002	0.0026	0.08	0.9385	
PORPROD6	0.0015	0.0026	0.58	0.5665	
PORPROD7	-0.0069	0.0027	-2.59	0.0113	*
PORPROD8	0.0007	0.0026	0.25	0.8052*	
PORBOTM	-0.0023	0.0026	-0.88	0.3837*	
Adjusted R-squared= 0.8562    Model p-value= $< 2.2e - 16$					

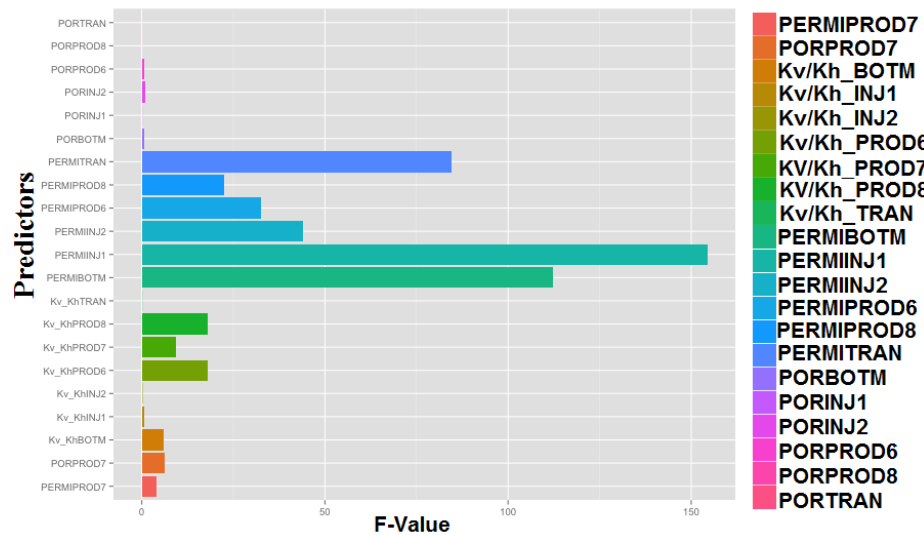


Figure 6.18: Effects of Full Reservoir Parameters on the GAGD Process Performance

Table 6.8: ANOVA Test of Full Linear Model

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
PERMIINJ1	1	0.04	0.04	154.42	$< 2.2e - 16$	***
PERMIINJ2	1	0.01	0.01	44.05	7.5e-9	
PERMITRAN	1	0.02	0.02	84.67	3.087e-14	
PERMIPROD6	1	0.01	0.01	32.65	4.413e-8	
PERMIPROD7	1	0.00	0.00	4.00	0.0487	
PERMIPROD8	1	0.01	0.01	22.44	2.165e-5	
PERMIBOTM	1	0.03	0.03	112.22	$< 2.2e - 16$	***
Kv/KhINJ1	1	0.00	0.00	0.77	0.3812	
Kv/KhINJ2	1	0.00	0.00	0.12	0.7295	
Kv/KhTRAN	1	0.00	0.00	0.06	0.8139	
Kv/KhPROD6	1	0.00	0.00	17.94	2.168e-5	***
Kv/KhPROD7	1	0.00	0.00	9.46	0.0028	
Kv/KhPROD8	1	0.00	0.00	18.04	8.526e-5	***
Kv/KhBOTM	1	0.00	0.00	6.03	0.0161	
PORINJ1	1	0.00	0.00	0.04	0.8484	
PORINJ2	1	0.00	0.00	0.92	0.3394	
PORTRAN	1	0.00	0.00	0.01	0.9192	
PORPROD6	1	0.00	0.00	0.59	0.4435	
PORPROD7	1	0.00	0.00	6.22	0.0145	
PORPROD8	1	0.00	0.00	0.01	0.9113	
PORBOTM	1	0.00	0.00	0.77	0.3837	
Residuals	87	0.02	0.00			

In order to inspect the validity of this full linear model to really fit the problem and to capture the factors' variation, an approximate matching between the theoretical and standardized residual curves should be performed in the normal Q-Q plot. Figure 6.19 depicts the basic diagnostic plots for this model and the normal Q-Q plot shows a non-approximate matching between the theoretical and standardized residual curves.

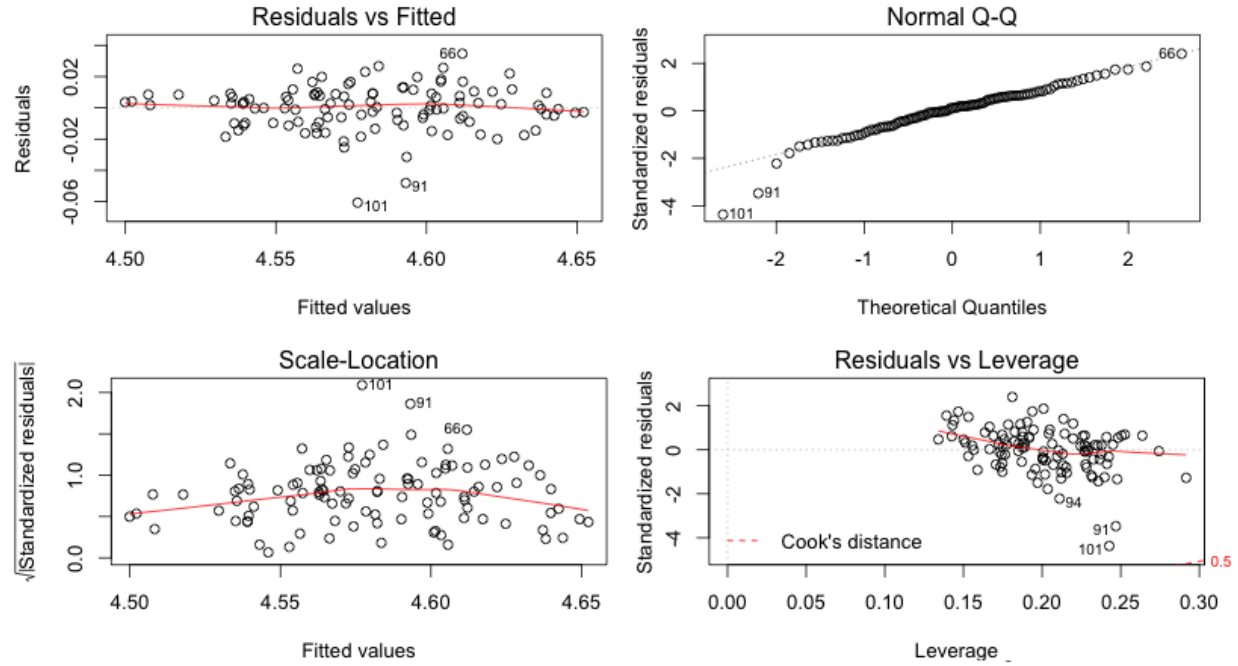


Figure 6.19: Basic Diagnostic Plots of Full Linear Model

As a consequence, the stepwise selection was employed to eliminate these non-influential parameters and to produce the reduced linear model. Table 6.9 and Table 6.10 represent the partial t-test and AVOVA tables for the reduced linear model, receptively.

Not all of the non-influential factors were eliminated by the stepwise selection method. For instant, (Kv/KhTRAN) is still in the model, even though it has no effect on the response because the p-value has exceeded 0.05. That fact was easily emphasized through the normal Q-Q plot in the basic diagnostic plots in Figure 6.20 which shows a non full matching between the theoretical and standardized residual curves. However, the Generalized Linear Hypothesis Test indicated that the reduced linear model is sufficient to represent the problem. The p-value was 0.9041 and it indicates that null hypothesis of similarity between full and reduced linear models failed to be rejected, as shown in Table 6.11.

As noted from the reduced linear model, the most influential factors affecting the GAGD process in the specific sensitivity analysis case are shown in the tornado plot in

Table 6.9: Partial t-Test of Reduced Linear Model

	Estimate	Std. Error	t value	Pr(> t )	
(Intercept)	4.4450	0.0111	402.12	$< 2e - 16$	***
PERMIINJ1	0.0257	0.0025	10.25	$< 2e - 16$	***
PERMIINJ2	0.0185	0.0025	7.40	5.51e-11	***
PERMITRAN	-0.0237	0.0026	-9.25	6.5e-15	***
PERMIPROD6	0.0116	0.0024	4.74	7.45e-6	***
PERMIPROD7	0.0099	0.0025	3.96	0.000147	***
PERMIPROD8	0.0099	0.0025	3.92	0.000166	***
PERMIBOTM	0.0272	0.0026	10.31	$< 2e - 16$	***
Kv/KhTRAN	0.0043	0.0026	1.65	0.1023	
Kv/KhPROD6	0.0130	0.0025	5.24	9.77e-7	***
Kv/KhPROD7	0.0093	0.0025	3.67	0.000407	***
Kv/KhPROD8	0.0105	0.0027	3.84	0.000225	***
Kv/KhBOTM	0.0065	0.0025	2.60	0.0108	*
PORPROD7	-0.0062	0.0025	-2.44	0.0164	*
Adjusted R-squared= 0.8524    Model p-value= $< 2.2e - 16$					

Figure 6.21.

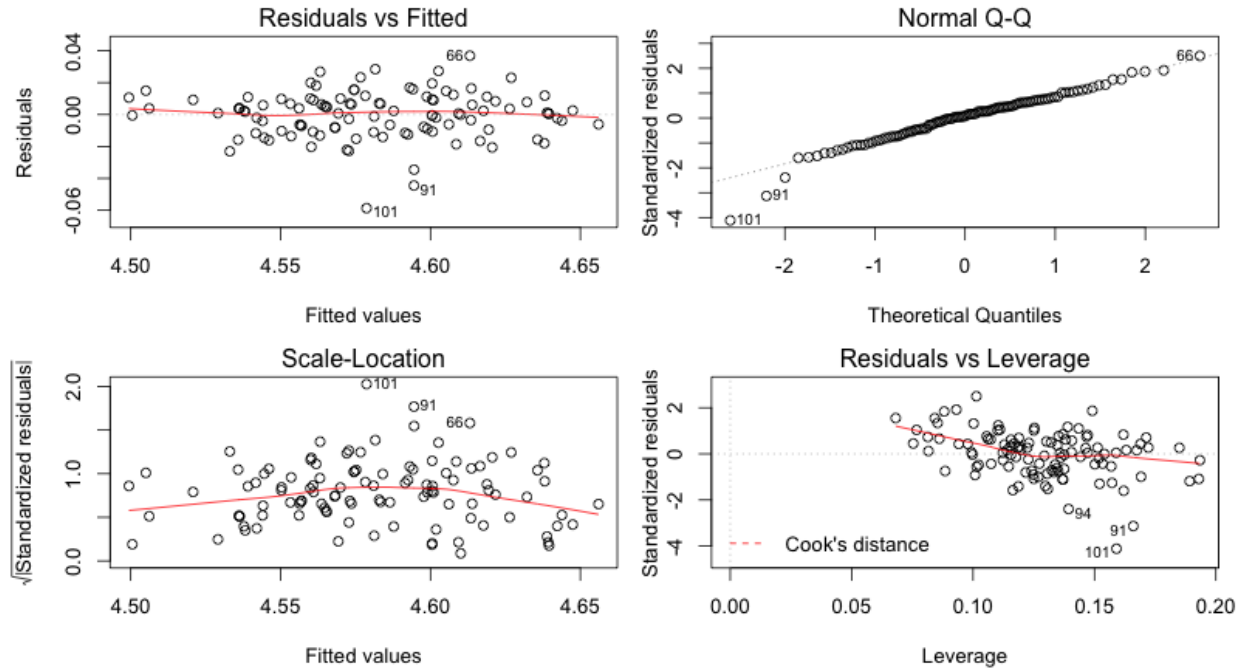


Figure 6.20: Basic Diagnostic Plots of Reduced Linear Model

Table 6.10: ANOVA Test of Reduced Linear Model

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
PERMIINJ1	1	0.04	0.04	165.88	$< 2.2e - 16$	***
PERMIINJ2	1	0.01	0.01	43.62	2.283e-9	***
PERMITRAN	1	0.02	0.02	87.70	3.765e-15	***
PERMIPROD6	1	0.01	0.01	38.30	1.531e-8	***
PERMIPROD7	1	0.00	0.00	3.90	0.0511	*
PERMIPROD8	1	0.01	0.01	21.46	1.148e-5	***
PERMIBOTM	1	0.03	0.03	124.97	$< 2.2e - 16$	***
Kv/KhTRAN	1	0.00	0.00	0.01	0.9199	
Kv/KhPROD6	1	0.00	0.00	20.37	1.831e-5	***
Kv/KhPROD7	1	0.00	0.00	11.14	0.001208	**
Kv/KhPROD8	1	0.00	0.00	18.75	3.699e-5	***
Kv/KhBOTM	1	0.00	0.00	6.46	0.0127	*
PORPROD7	1	0.00	0.00	5.97	0.0164	*
Residuals	95	0.02	0.00			

Table 6.11: GLHT of the Full and Reduced Models

	Res.Df	RSS	Df	Sum of Sq	F	Pr(>F)
1	87	0.02				
2	95	0.02	-8	-0.00	0.42	0.9041

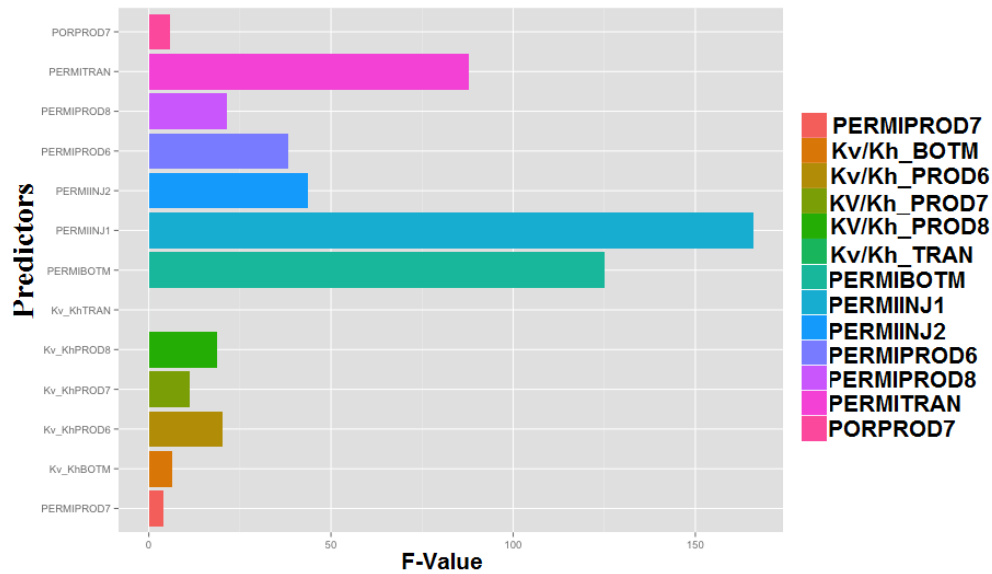


Figure 6.21: Effects of Reduced Reservoir Parameters on the GAGD Process Performance

The best subset model (reduced model) from the stepwise elimination approach has indicated that horizontal permeability and anisotropy ratio ( $K_v/K_h$ ) are more sensitive in the injection, bottom, and transition layers than the production layers. In addition, the effect of permeability on the GAGD process in these zones is larger than  $K_v/K_h$ . There was a negligible porosity effect on all layers as well.

In general, the permeability and anisotropy ratio have a direct impact on the fluid flow through the reservoir media. This effect barely appears with respect to the reservoir porosity, which affects the reservoir storage and not the fluid flow. As previously stated, the permeability and the anisotropy ratio has less effect on the GAGD process in the production layers because most of the fluid flow was through the injection and transition zones with regards to the injected CO<sub>2</sub> and the displaced oil.

Although the reduced linear model was quite enough to replace the full one, it has not accurate fit for the problem. Therefore, Bayesian Model Averaging was adopted to produce many models, ranking them, and select the best one with the highest posterior probability and lower BIC. The BMA procedure starts with obtaining normal transformation for all factors prior to fit the models, as shown in Figure 6.22. In this figure, the vertical axis on the right represents the posterior distribution given the data. In some factors' plots, the vertical spikes correspond to the posterior probability of a factor's coefficient equals to 0.

In this stochastic case, 26 models were generated as the sum of their posterior probability equals one. Only, five models were shown in the final Occam's windows that represent the highest ranges of posterior values, as included in Table 6.12.

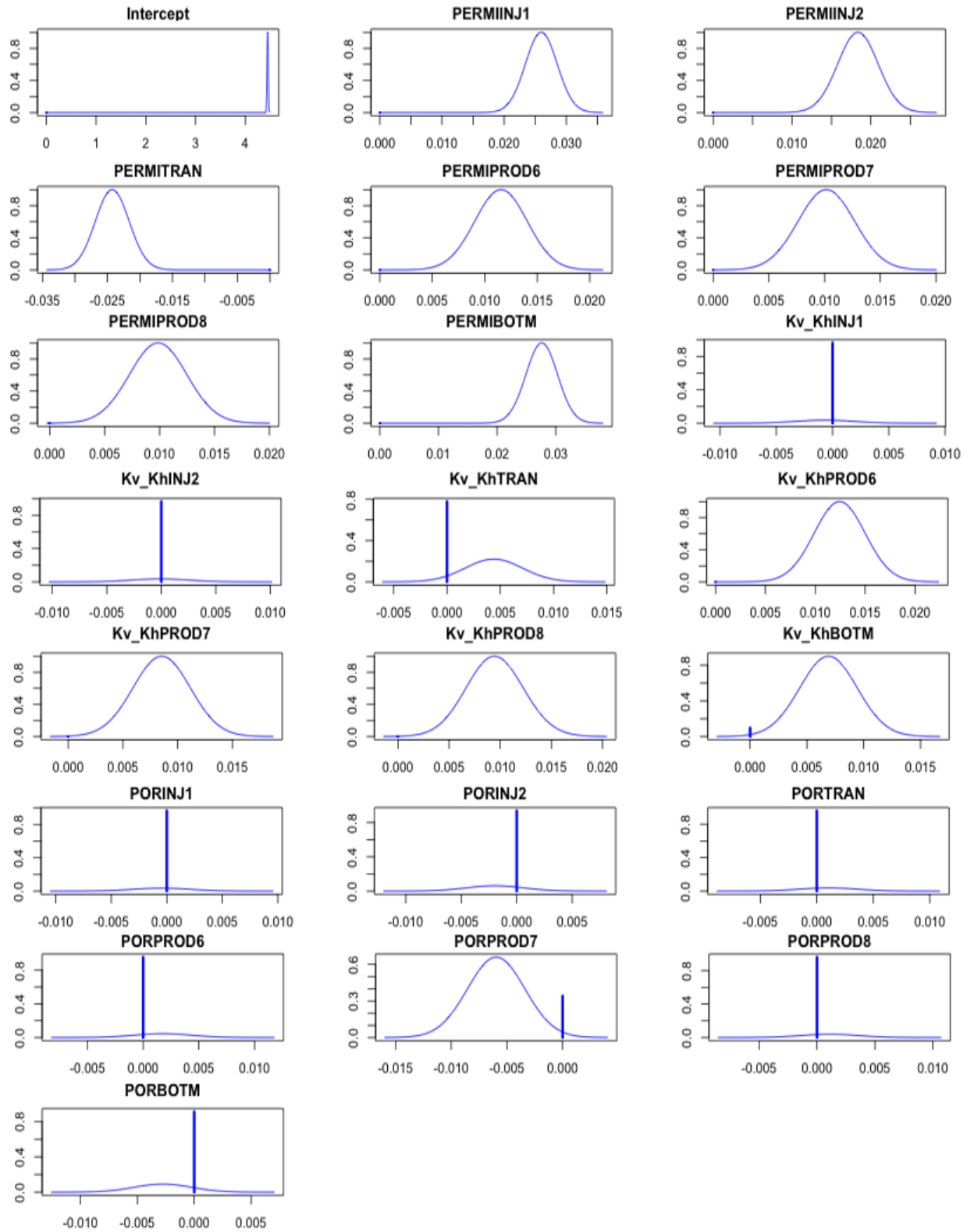


Figure 6.22: Normal Transformation for all Factors in the Linear Model

Table 6.12: The First Best Fitted Models through Bayesian Model Averaging

	p!=0	EV	SD	model 1	model 2	model 3	model 4	model 5
Intercept	100	4.45e+00	0.0112	4.45e+00	4.45e+00	4.44e+00	4.45e+00	4.44e+00
PERMIINJ1	100	2.59e-02	0.0025	2.58e-02	2.61e-02	2.56e-02	2.59e-02	2.59e-02
PERMIINJ2	100	1.83e-02	0.0025	1.85e-02	1.81e-02	1.84e-02	1.83e-02	1.80e-02
PERMITRAN	100	-2.42e-02	0.0026	-2.46e-02	-2.45e-02	-2.36e-02	-2.40e-02	-2.36e-02
PERMIPROD6	100	1.15e-02	0.0025	1.16e-02	1.14e-02	1.16e-02	1.18e-02	1.14e-02
PERMIPROD7	100	1.01e-02	0.0025	1.02e-02	9.99e-03	9.85e-03	1.04e-02	9.66e-03
PERMIPROD8	100	9.86e-03	0.0026	9.61e-03	1.04e-02	9.91e-03	9.86e-03	1.07e-02
PERMIBOTM	100	2.75e-02	0.0026	2.80e-02	2.72e-02	2.72e-02	2.81e-02	2.65e-02
Kv_KhINJ1	3.7	-2.49e-05	0.0005	.	.	.	.	.
Kv_KhINJ2	3.6	-2.76e-06	0.0004	.	.	.	.	.
Kv_KhTRAN	22	9.70e-04	0.0022	.	.	4.34e-03	.	3.78e-03
Kv_KhPROD6	100	1.24e-02	0.0025	1.24e-02	1.18e-02	1.30e-02	1.23e-02	1.23e-02
Kv_KhPROD7	100	8.53e-03	0.0026	8.68e-03	8.11e-03	9.32e-03	8.27e-03	8.64e-03
Kv_KhPROD8	100	9.49e-03	0.0028	9.20e-03	8.75e-03	1.04e-02	9.39e-03	9.84e-03
Kv_KhBOTM	90	6.22e-03	0.0031	7.07e-03	6.86e-03	6.49e-03	7.08e-03	6.33e-03
PORINJ1	3.6	-1.69e-05	0.0005	.	.	.	.	.
PORINJ2	6.3	-1.23e-04	0.0008	.	.	.	.	.
PORTRAN	3.9	4.04e-05	0.0005	.	.	.	.	.
PORPROD6	4.6	8.12e-05	0.0006	.	.	.	.	.
PORPROD7	65.9	-3.93e-03	0.0035	-5.85e-03	.	-6.22e-03	-6.23e-03	.
PORPROD8	4.0	4.31e-05	0.0005	.	.	.	.	.
PORBOTM	9.1	-2.50e-04	0.0011	.	.	.	-2.95e-03	.
nVar				12	11	13	13	12
r2				0.848	0.840	0.852	0.85	0.84
BIC				-149	-148	-147	-146	-145
post prob				0.234	0.136	0.105	0.05	0.04

In table 6.12, the selected best model, which has the highest posterior probability and lowest BIC, is the first one. On other hand, the other four models have less posterior probability than the first one. The anisotropy ratio in the injection and transition zones along with the porosity in most zones are indicated to be non-influential on the GAGD process based on the outcome of the 1st model. Beside that, most of these non-influential factors are also explained by the other four models in Table 6.12.



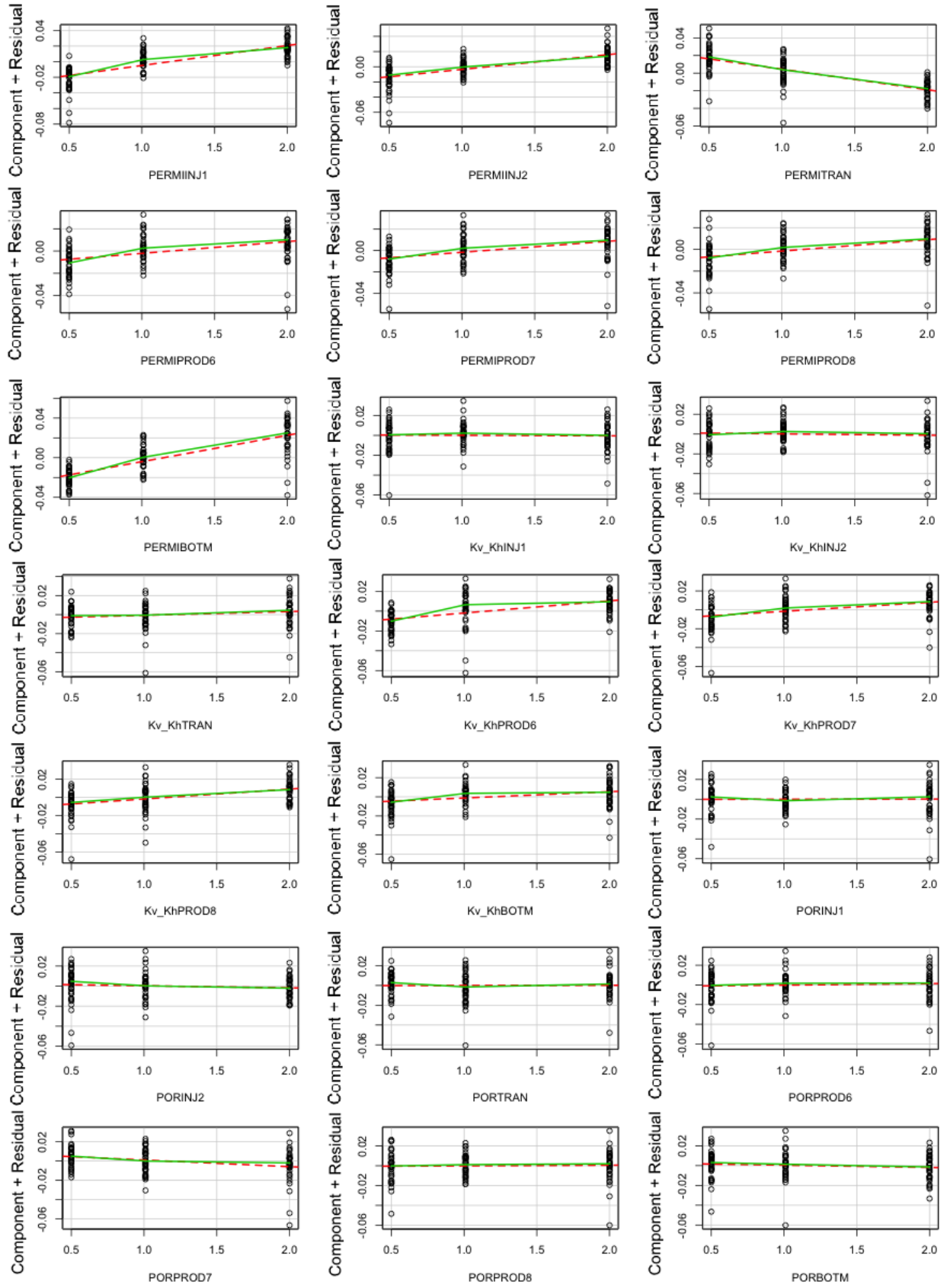


Figure 6.23: Component-Residual Plots for the Full BMA Linear Model

Furthermore, the effects of all factors under study is illustrated in the component-residual plots (Figure 6.23). It is noticeable that many factors have no effect on the response as they have horizontal straight line.

These non-influential factors are again identified in Figure 6.24 that shows the five models with their subset selections. In this figure, the background has light yellow color and the sensitives factors are represented by red and blue colors. The blue bar refers to the negative effects for factors on the flow response.

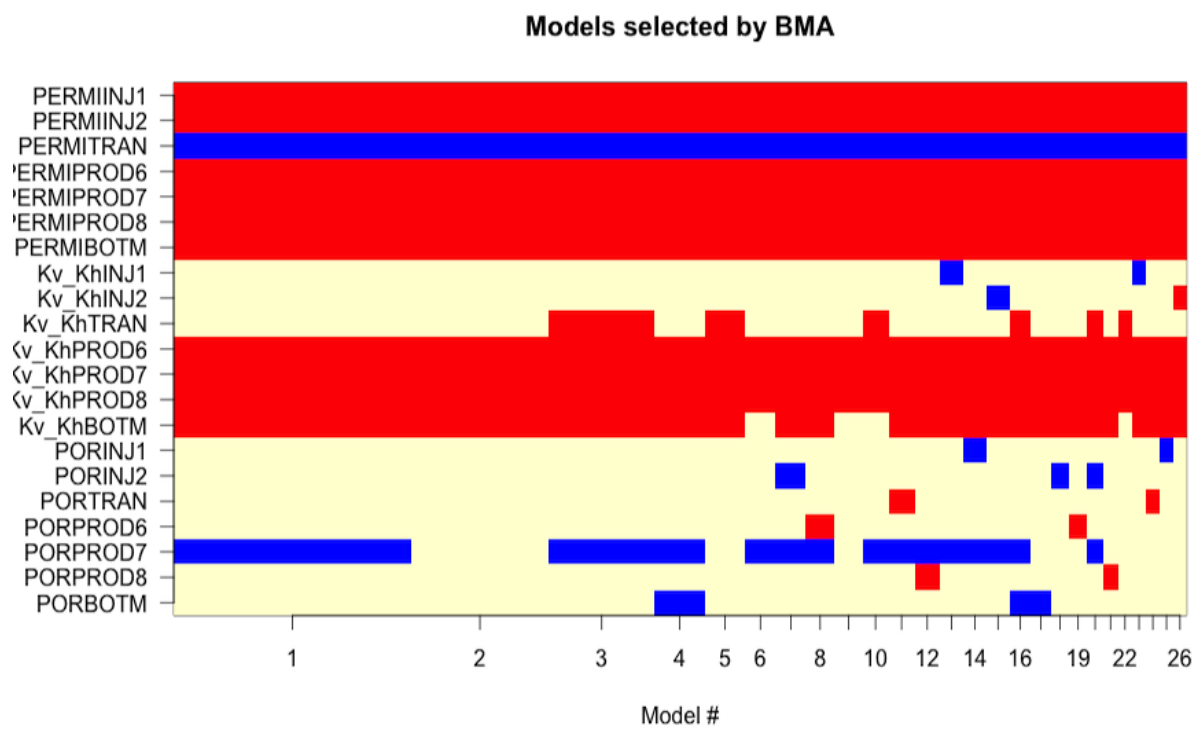


Figure 6.24: Best Selected Linear Model By Bayesian Model Averaging

As a result, the most influential parameters on the GAGD process performance through the Bayesian Model Averaging approach, after eliminating the non-influential factors, are shown in Figure 6.25 that exposes the component-residual plots. Figure 6.26 shows the tornado plot of the most influential factors for the reduced BMA model.

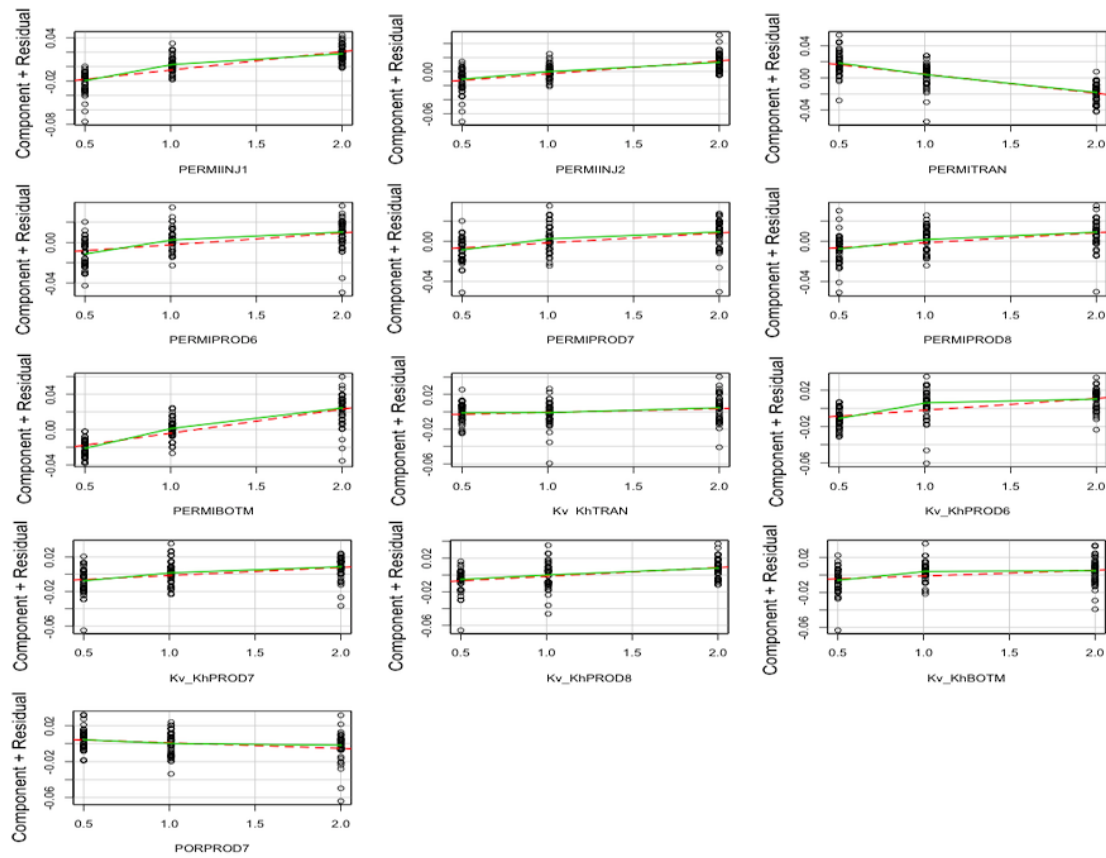


Figure 6.25: Component-Residual Plots for the Reduced BMA Linear Model

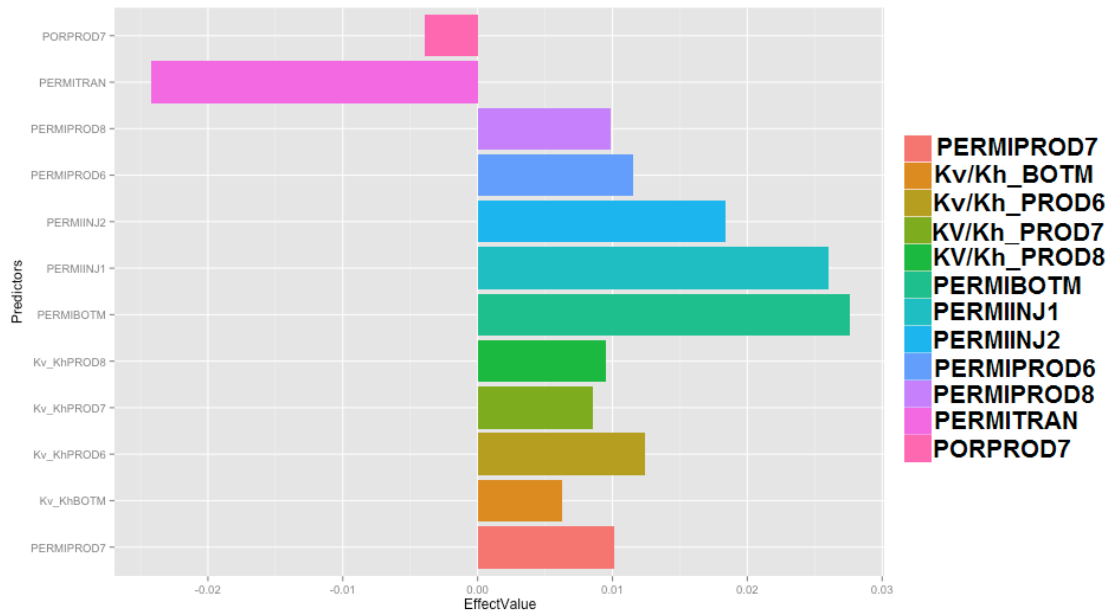


Figure 6.26: Effects of BMA-Based Reservoir Parameters on the GAGD Process

To check the validity and accuracy of the fitted models in stepwise and BMA approaches, Bootstrapping was conducted with respect to Adjusted R-squared to compute the confidence intervals. Figure 6.27 and Figure 6.28 illustrate the histograms for Adjusted R-squared, intercept, and three most sensitive factors for the stepwise-based reduced linear model.

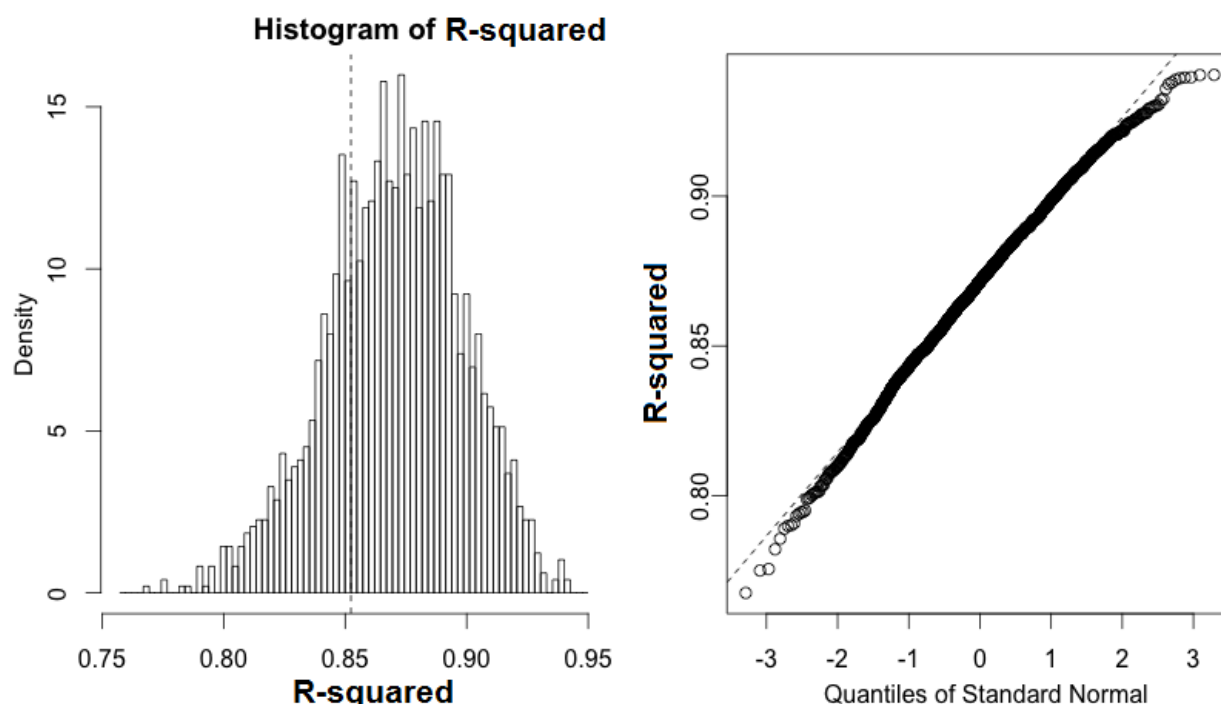


Figure 6.27: Nonparametric Bootstrapping of  $R^2$  in Stepwise-Reduced Linear Model

On the other hand, Figure 6.29 and Figure 6.30 decorate the histograms for Adjusted R-squared, intercept, and three most sensitive factors for the BMA-based reduced linear model. All the bootstrapped parameters have normal distributions with significant confidence intervals.

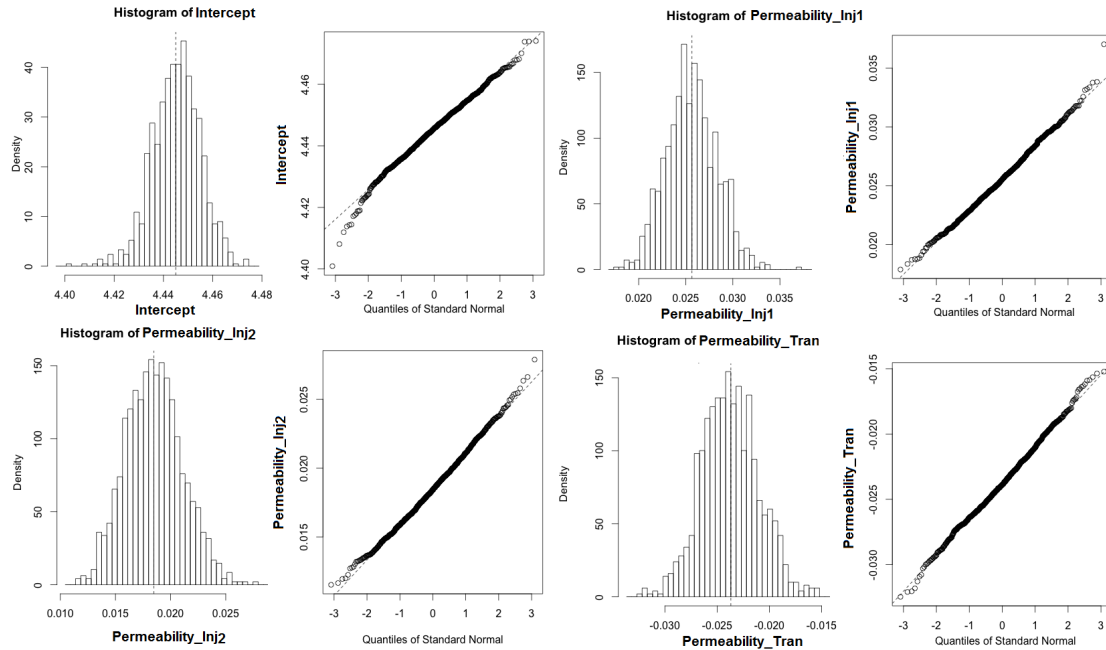


Figure 6.28: Nonparametric Bootstrapping of Four Coefficients in Stepwise-Based Reduced Linear Model

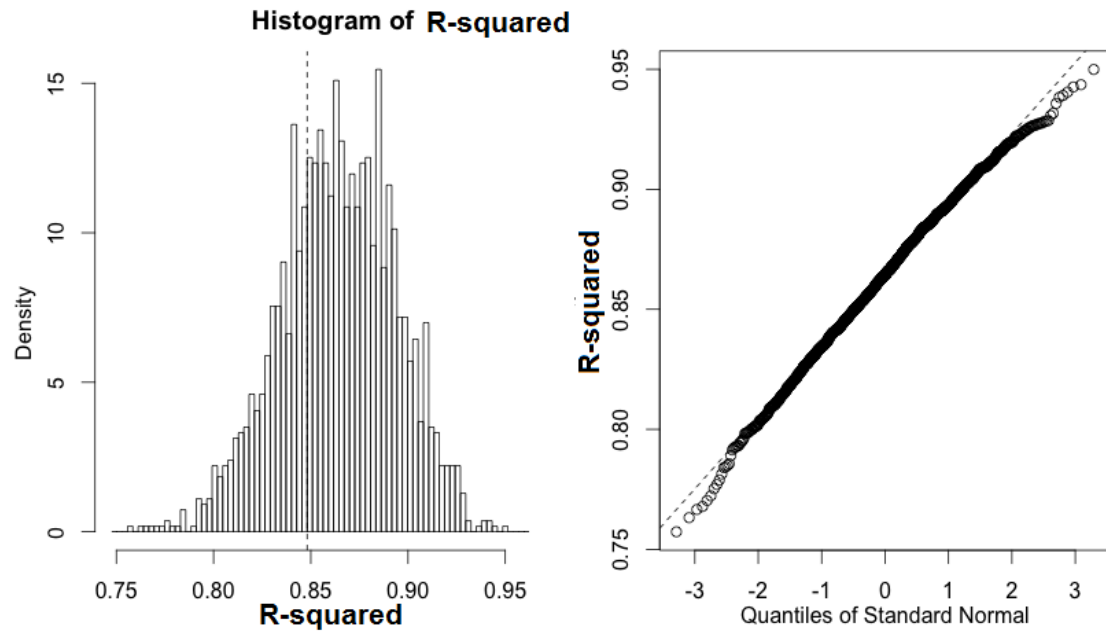


Figure 6.29: Nonparametric Bootstrapping of  $R^2$  in BMA-Reduced Linear Model

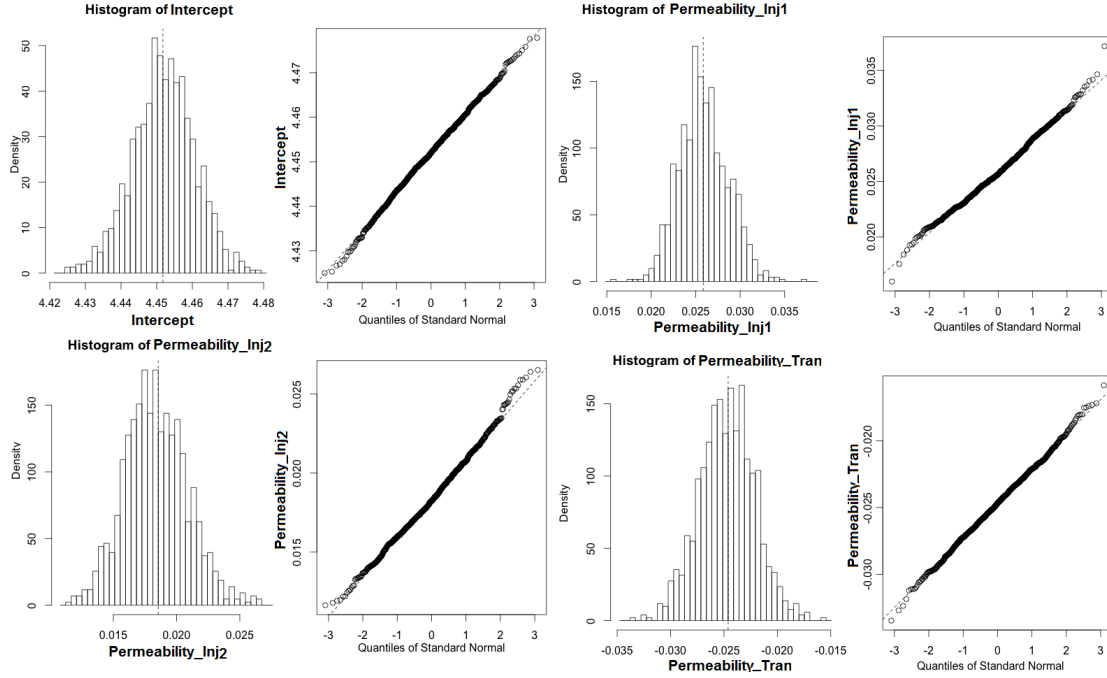


Figure 6.30: Nonparametric Bootstrapping of Four Coefficients in BMA-Reduced Linear Model

### Validation of Influential Factors-Specific Case

Many other observations (computer experiments) were generated to investigate the effect of each factor independently on the flow response. The generated observations were then evaluated with their impacts on the reservoir flow responses with regards to field cumulative oil production, oil rate, and water cut. Multiple runs were carried out for permeability in the injection, transition, production, below-production zones to prove their larger effects, as illustrated in Figures 6.31, 6.32, 6.33, and 6.34. The figures again have a tornado plot, which shows the calculated minimum and maximum field cumulative oil production from the generated simulation jobs.

The discrepancy between the flow responses of various observations reflects the impacts of these factors on the reservoir performance through the GAGD process. The same number of experiments were created and evaluated for anisotropy ratio for production layers, as illustrated in Figure 6.35.

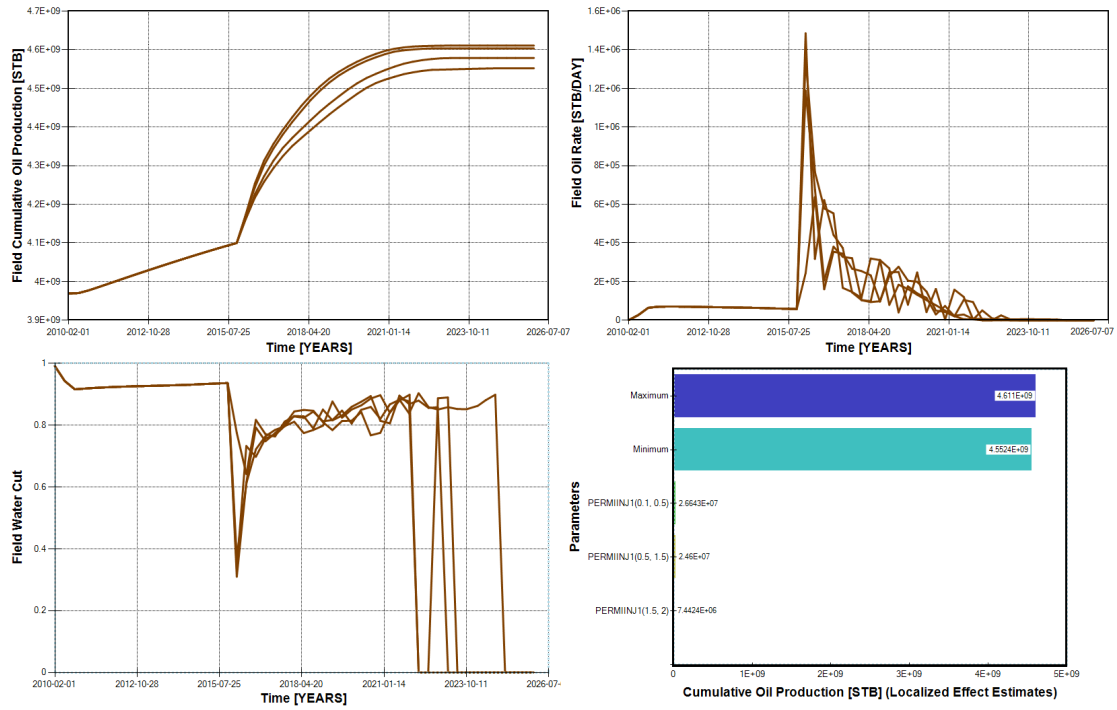


Figure 6.31: Validation Effect of Injection Layer Permeability on the Field Flow Response

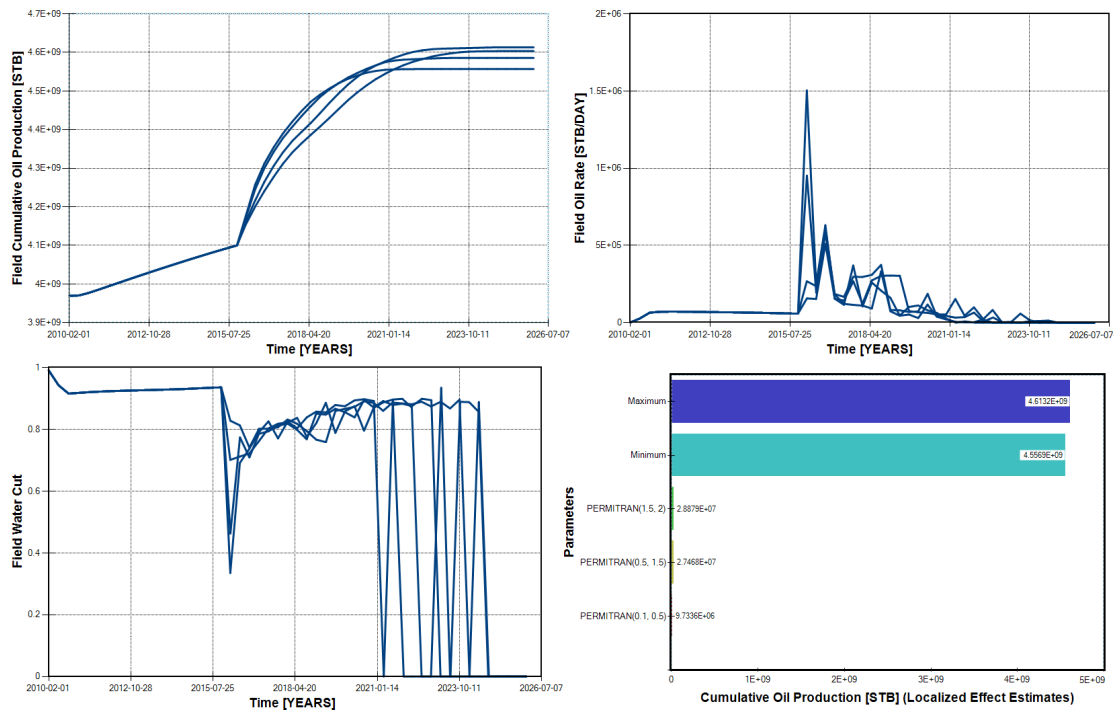


Figure 6.32: Validation Effect of Permeability of Transition Layers on the Flow Response

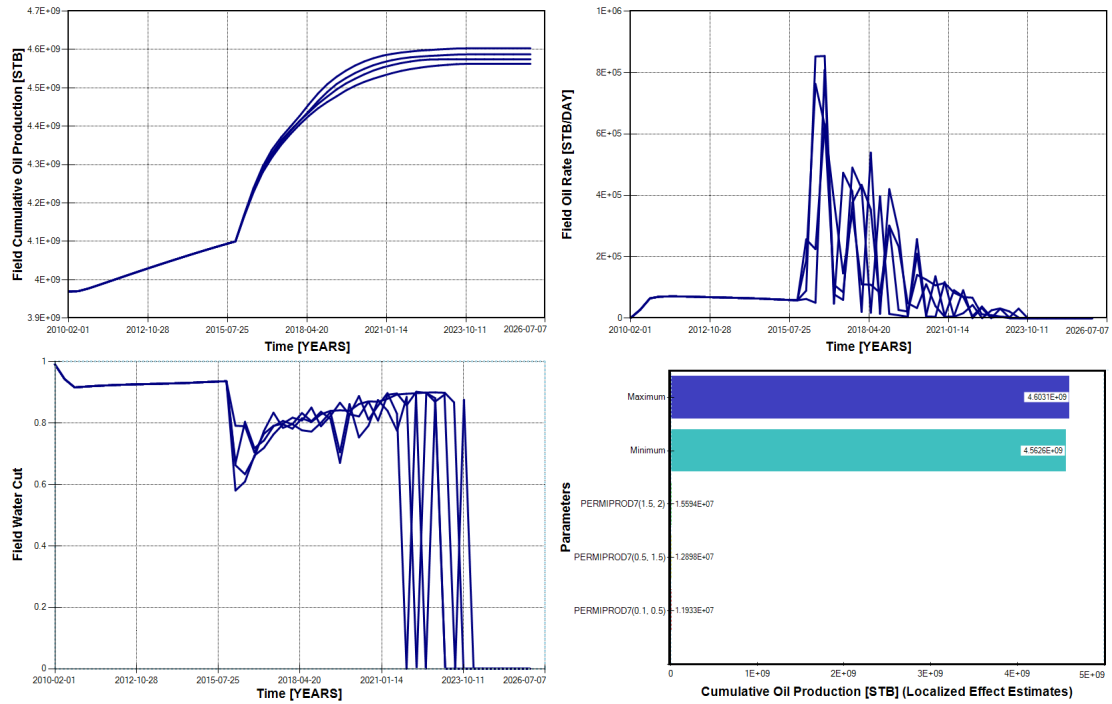


Figure 6.33: Validation Effect of Permeability of Production Layer on the Flow Response

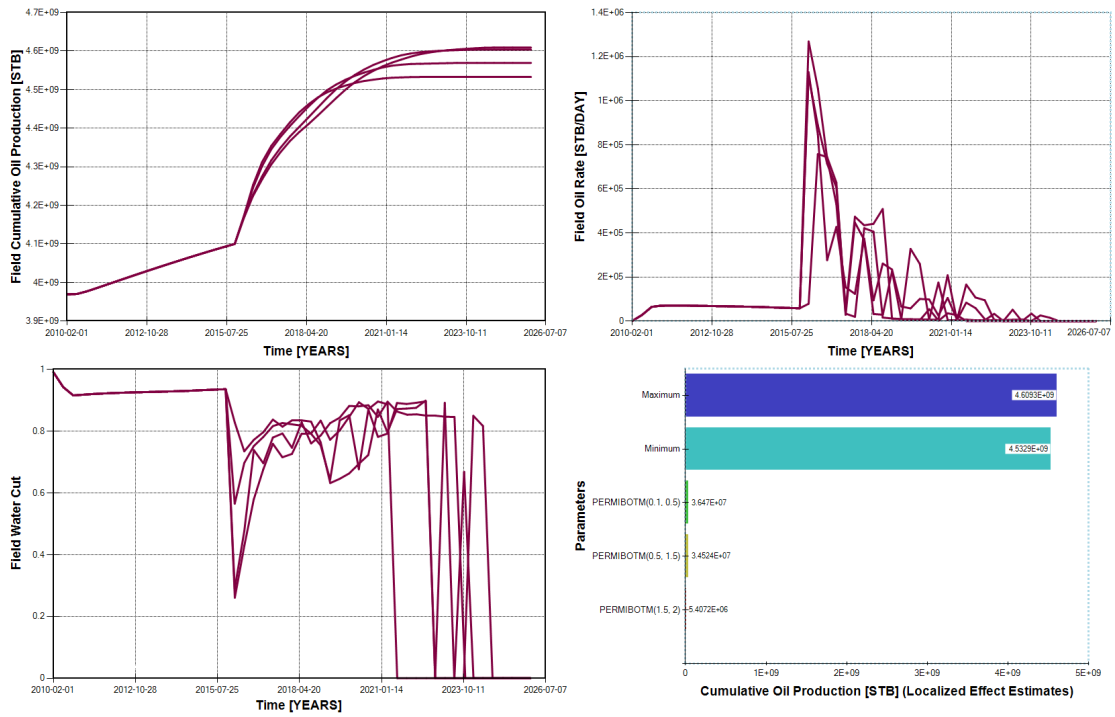


Figure 6.34: Validation Effect of Below-Production Layers' Permeability on the Flow Response



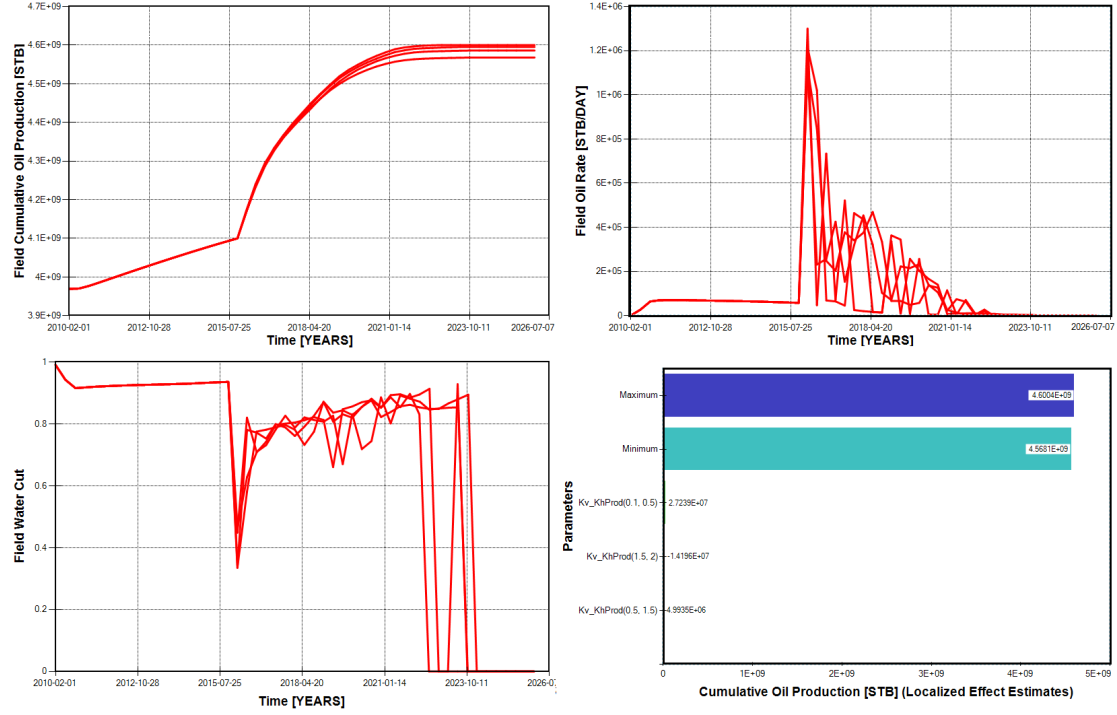


Figure 6.35: Validation Effect of Production Layers' Kv/Kh on the Flow Response

There is matching between the statistical models output about the degree of effects of each factor on the flow response. For instance, permeability of the injection layers are more sensitive than the bottom layers.

Although the porosity in the injection and transition layers seem to have some impact on the flow response, there is no real effect for the mentioned layers and all other layers, as shown in Figures 6.36 and 6.37 that disclose the non-influential role of porosity in production and other layers, correspondingly.

Consequently, Bayesian Model Averaging has reduced the uncertainty in sensitivity analysis because it showed more precise estimation of which factor has the most influential role on the GAGD process than the stepwise elimination. For instance, stepwise selection has indicated that Permeability in the first injection layer is more influence than all other layers, as clarified in Figure 6.21. However, Figures 6.31 and 6.34 illustrate that permeability in bellow-production layers is more sensitive than in production layers. This fact was demonstrated by Bayesian Model Averaging, as in Figure 6.26.

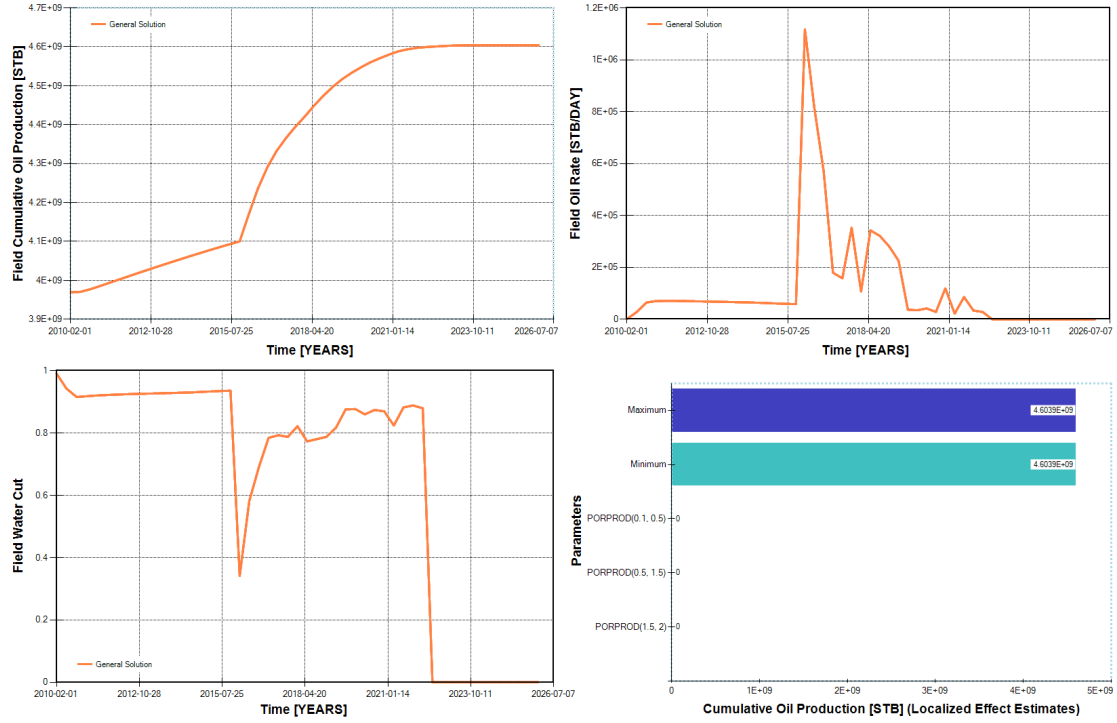


Figure 6.36: Validation Effect of Porosity of Production Layers on the Flow Response

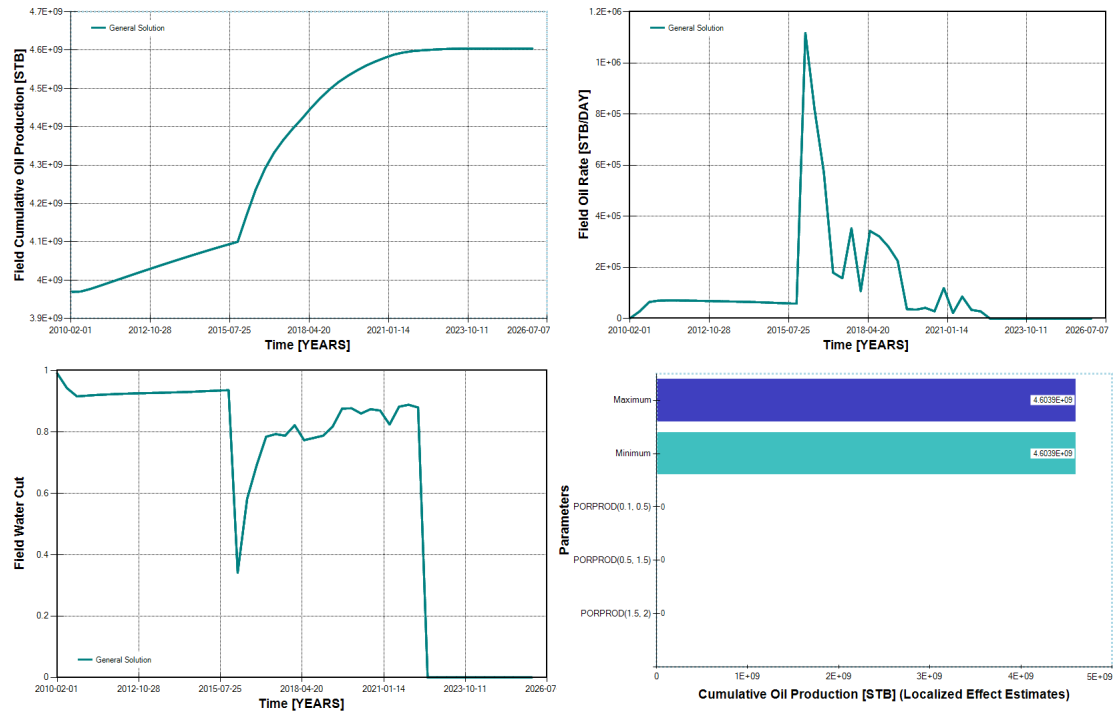


Figure 6.37: Validation Effect of Porosity of Other Layers on the Flow Response

### 6.4.3 Comments on A Previous Sensitivity Analysis Study

A previous study has been conducted on South Rumaila oil field, was presented at SPE conference (SPE-174579), to determine the most influential factors affecting the GAGD process (Al-Mudhafar and Rao, 2015). In that study, the reservoir porosity has the highest influential role than all other factors including reservoir permeability. On the other hand, the current study has shown that reservoir porosity has no effect on the GAGD process performance.

The contradiction between the two studies are explained by the following items:

1. The SPE study has complete different reservoir characterization model from the current study. The SPE study does not have lithology modeling by neglecting the shale and shaly sand facies distribution. The range of permeability and porosity for the SPE study was for sandstone only. However, the current study has distinct ranges for sand, shaly sand, and shale.
2. The reservoir simulation setting in the two studies is completely different. The grid dimension in SPE study is 30 in I-direction, 22 in J-direction, and 5 layers only. On the other hand, the current study has 69 grids in I-direction, 66 grids in J-direction, and 12 vertical layers.
3. The GAGD application setting in the two studies are not similar. The prediction period for the SPE study starts from 2010 to 2022. However, it starts from 2016 to 2026 in the current study. Additionally, the model was run from 2010 to 2016 to start the GAGD process, but the GAGD application started in the SPE study in 2010.
4. The factors setting in the two studies are completely different. The current study has only reservoir parameters (permeability, porosity, and anisotropy ratio). The SPE study has combined static and dynamic parameters: skin factor, porosity, permeability, compressibility, anisotropy ratio along with minimum bottom hole pressure and

maximum oil production rates in the production wells. Different setting of factors leads to distinct strength of influential factors on the GAGD process.

5. The number of oil horizontal producers in the two cases are completely dissimilar. Five horizontal wells with 1500 m length were set in the SPE study, but there is 11 horizontal wells were set in the current study with 3000 m length.
6. Since the number of studied factors in the two studies are not the same, the total number of designed simulation jobs are different. That leads to different statistical analysis between the two studies.
7. In addition, the running of the reservoir model in the SPE study is not correct as changing the parameters started at the 1st time step. For instance, changing the porosity at the 1st time-step led to change the value of initial oil in place. When multiplying by 2 or divide on 2, the initial oil in place with become twice or half, respectively. That setting impedes the effect of other factors and make the porosity is a dominant factor on the GAGD process.

Consequently, the SPE study has inaccurate way of running the simulation jobs. The current study has accurate setting of the GAGD process by considering changing the parameters in the prediction period only (after setting the RESTART option).

## **6.5 Summary and Significance**

To determine the most influential geological parameters that impact the GAGD process, two different approaches of sensitivity analyses were implemented. The first approach implemented sensitivity analysis with respect to reservoir porosity, permeability and anisotropy ratio ( $Kv/Kh$ ) for all the reservoir zones. In the second approach, the reservoir layers were sub-classified into four groups: injection, transition (between injection and production), production, and bottom (below production) layers.

In the first sensitivity analysis approach, the stepwise elimination was conducted to eliminate the non-influential factors on the reservoir flow response. It was investigated

that permeability and anisotropy ratio have the most impact on the GAGD process. The second approach represents a detailed sensitivity analysis with respect to the layers' classification. The Bayesian Model Averaging was then adopted as a stochastic statistical approach in order to select the best model fit through Bayes' theorem and posterior probability. In the detailed sensitivity analysis, the permeability of injection, below-production, and transition layers were investigated to be more influential in the GAGD process than the anisotropy ratio, which is influential only in the production and below-production layers. The horizontal permeability is more influential at the top injection layers than any other layer because of the shale modeling impedes the vertical gas and oil flow, based on the concept of GAGD process.

To validate the sensitivity analysis approaches, the identified most influential factors were changed within different levels and simulated in the reservoir model to check the effect on the GAGD process performance. The influential strength of various parameters were recognized by visualizing the effect of different factors on the field cumulative oil production. The validity and accuracy of the sensitivity analysis models were further checked through a nonparametric bootstrapping that computes the estimated standard errors and confidence intervals of the parameters' coefficients.

## 7. GAGD Process Optimization

### 7.1 Introduction

In order to optimize oil recovery, it is essential to determine the optimal levels of the operational design factors that impact the Enhanced Oil Recovery processes' performance. These factors are mainly comprised of operational constraints in production and injection wells. More specifically, the manner in which the injection and production constraints are set controls the amount of fluid production and injection in the reservoir, Thus, providing a direct impact on the reservoir flow response. Therefore, the optimization of these factors will result in the optimal reservoir performance over time with respect to field cumulative oil production and Net Present Value (NPV) (White and Royer, 2003).

The operational design factors that impact the Gas Assisted Gravity Drainage (GAGD) process represent the production/injection constraints adhering to the limits of fluid rates and pressures. The GAGD process optimization includes determining the optimal oil recovery through production controls, injection pressure, and huff-n-puff optimization settings for 10 years of future reservoir prediction.

### 7.2 Optimization Approaches

The main optimization methods that were adopted in this study are Design of Experiments (DoE) such as Latin Hypercube plus Proxy Model and CMG Designed Exploration and Controlled Evolution (CMG DECE).

### 7.2.1 Latin Hypercube Plus Proxy Model

Latin Hypercube Sampling (LHS) concepts were discussed in Chapter 6. LHS creates minimum simulation jobs in order to obtain maximum information by various combinations of the input parameter values. By adding the proxy model to the LHS method, this enables design quality optimization by incorporating the training data obtained from the Latin hypercube design simulations. More specifically, the training data are the simulation jobs created from Latin hypercube designed experiments, which is then used to build the proxy model. In order to increase the chance of reaching global optima and to improve the prediction accuracy of the proxy model, the training data needs to be validated by verification simulation jobs. These verification simulation jobs are iterative procedures that ensure the achievement of a 95% confidence interval between the predicted and actual objective functions.

Therefore, the proxy model is updated frequently after adding the new set of simulation jobs in order to obtain the true optimal solution (Yang et al., 2011). Latin Hypercube Plus Proxy Model was adopted in the optimization of production controls and injection pressure through the GAGD process. The complete flowchart for Design of Experiments-Proxy model optimization is shown in Figure 7.1.

### 7.2.2 Response Surface Methodology

Response Surface Methodology (RSM) is an optimization tool that is used for modeling response factor as function of input variables. Unlike the linear models, the RSM considers either the polynomial or ordinary kriging model to obtain the expected value of response factor (Kalla and White, 2007). The assumed polynomial function is:

$$y = f(X_1, X_2, \dots, X_k) + \epsilon_i \quad (7.1)$$

where  $X_1, X_2, \dots, X_k$  are the input variables and  $y$  is the expected response factor.

The common model for the polynomial RSM is the second-order quadratic model. For

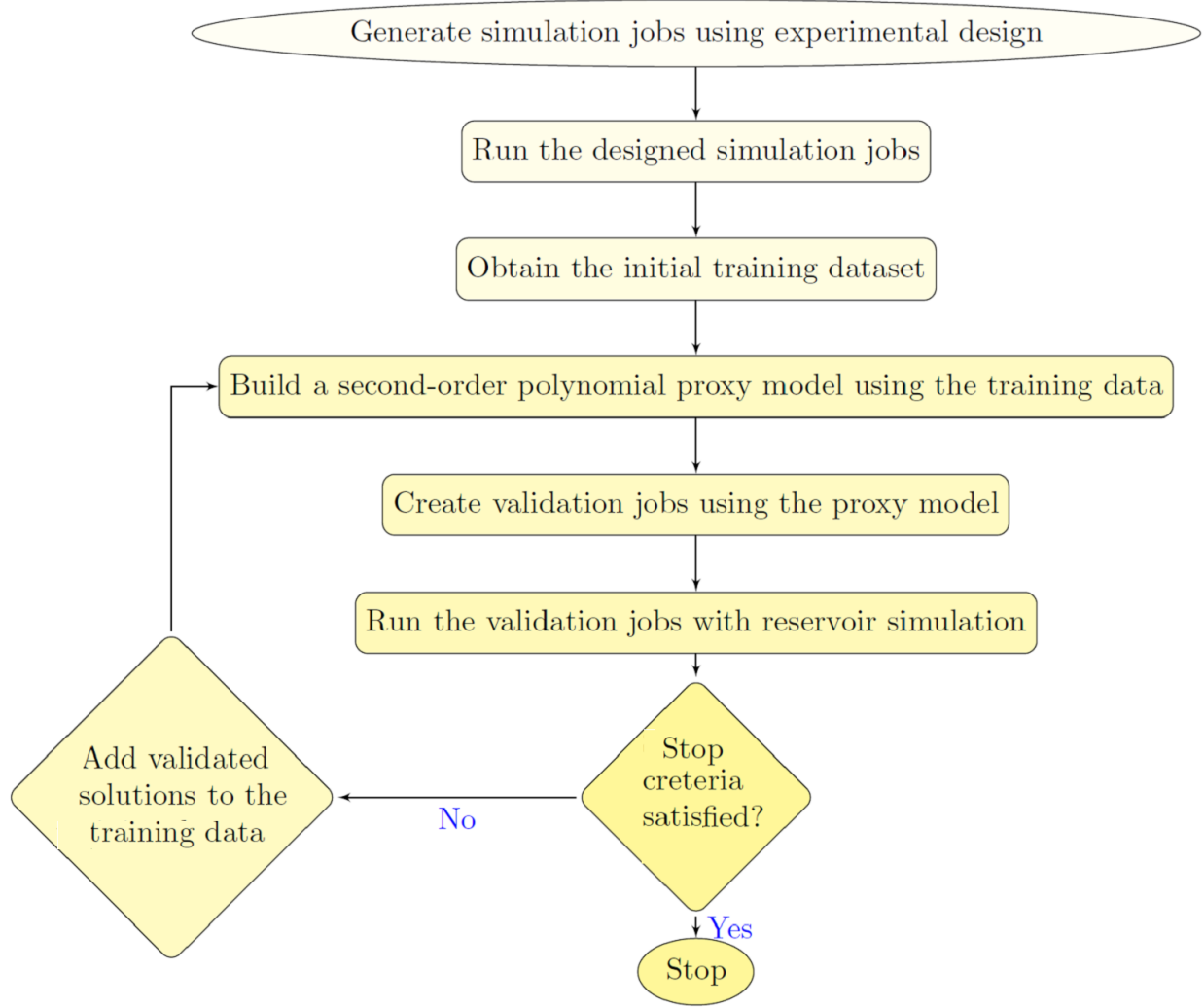


Figure 7.1: Flowchart of Design of Experiments-Proxy Optimization Algorithm

more than two parameters, it is represented by the following formula:

$$y = \alpha_0 + \sum_{j=1}^k \alpha_j X_j + \sum_{j=1}^k \alpha_{jj} X_j^2 + \sum_{i < j} \sum_{j=2}^k \alpha_{ij} X_i X_j + \epsilon_i \quad (7.2)$$

where  $\alpha_j$  is the coefficient of linear term,  $\alpha_{jj}$  is the coefficient of the quadratic terms, and  $\alpha_{ij}$  is the coefficient of the interaction terms between every two variables.

The ordinary kriging algorithm is also used to build the proxy model to be considered for optimization processes (Kalla and White, 2007). It builds the proxy model after evaluating the created observations through the Design of Experiments (Denney, 2010). In ordinary



kriging, the known points of the response factor are combined into a linear combination in which their weights (coefficients) are estimated based on the variogram. Variogram defines the autocorrelation function between known points with a separation distance (lag). After constructing and fitting the variogram, its parameters (range, sill, and nugget) are then employed to estimate the linear weights of the known points at different locations  $x_p = x_1, x_2, \dots, x_M$  based on the following formula (Fedutenko et al., 2013):

$$\hat{y}(x_p) = \sum_{i=1}^n w_i(x_p) y(x_i) \quad (7.3)$$

where the estimator function  $\hat{y}(x_p)$  is the estimated response factor,  $x_1, x_2, \dots, x_M$  are the separation distance between the known points, and  $y(x_i)$  is the known points of response factors all locations (Fedutenko et al., 2013).

The RSM models can be visualized either by contour maps or 3D view to show the optimum regions (Montgomery and Runger, 2003). The response surface methodology was entirely implemented via R-statistical language by *rsm* package (Lenth, 2009).

The RSM-proxy model of second order polynomial equation was adopted in the production control and injection pressure optimization. In addition, it was again employed in the geological uncertainty quantification of the Huff-and-Puff GAGD process, as will be discussed in Chapter 8. Instead, the ordinary kriging proxy model was adopted in the production uncertainty assessment of Huff-and-Puff GAGD process, as will be addressed in Chapter 8.

### 7.2.3 DECE Algorithm

The CMG DECE (Designed Exploration and Controlled Evolution) algorithm is used for generating computer experiments for sensitivity analysis and production optimization. CMG DECE method is conducted through two successive stages: Designed Exploration and Controlled Evolution. In the Designed Exploration stage, simulation runs are designed by random selection of levels for each parameter through Tabu search and experimental

design. In the Controlled Evolution, the objective function is then optimized based on the natural selection. A gene and an allele represent each parameter and levels, respectively (Yang et al., 2007). Identifying which gene has the most effect on the objective function is determined by observing which allele leads to bad results to be deactivated. The purpose of this procedure is to obtain maximum amount of information on a minimum number of simulation experiments.

The main advantage of the CMG DECE algorithm is that it reaches the optimal solution faster than other experimental design tools as well as it does not get trapped in local optima while checking for deactivated alleles. The repeated verification of the deactivated alleles ensures that these alleles will not be reentered into the algorithm. (Yang et al., 2007; CMG, 2011).

DECE was employed in this chapter for the optimization of GAGD Huff-and-Puff process under geological uncertainties in addition to the nominal optimization.

### **7.3 Optimization under Geological Uncertainty**

Capturing more realistic geological environment for the reservoir simulation is a crucial step for field development optimization. Due to the expected impact of reservoir properties on the Gas Assisted Gravity Drainage (GAGD) process performance, the optimization procedure should take into account geological uncertainties that are not based on a single realization. Therefore, multiple realizations were created through the sequential gaussian simulation and evaluated in the compositional reservoir model in order to rank them and determine the nine quantiles (p10, P20,..., P90), which represent the overall geological uncertainty. The entire ranking procedure was addressed in Chapter 5.

The ranked geological realizations were then incorporated in the optimization processed along with Design of Experiments (DoE). This type of optimization under geological uncertainty is called Robust Optimization. The nominal and robust optimization approaches were adopted in the Huff-and-Puff GAGD process.

### 7.3.1 Robust Optimization

The nominal optimization adopts only one single realization of each reservoir property. The algorithm that was employed for this approach is the DECE algorithm that was described in the previous section.

Due to the inherent uncertainty from geological parameters such as permeability and porosity, it is essential to take into account these uncertainties in the optimization process (Fedutenko et al., 2013). In order to quantify the geological uncertainties, multiple realizations of the reservoir stochastic images that describe the unknown uncertainty space need to be simulated in the reservoir model (Yang et al., 2011).

The following equation illustrates the unknown uncertainty space ( $\Theta$ ) encompasses the ranked nine geostatistical realizations for each property ( $\theta_d$ ).

$$\theta_d = \{\theta_1, \theta_2, \dots, \theta_{N_r}\} \subset \Theta \quad (7.4)$$

where  $N_r$  is the total number of ranked realizations.

When the reservoir flow response (Objective Function) in the GAGD process optimization is  $J$  and the periods parameters in the Huff-and-Puff optimization are  $T$ , the objective function under geological uncertainty is (Fedutenko et al., 2013):

$$J = J(T, \theta_d). \quad (7.5)$$

The objective function in robust optimization, which takes into account the geological uncertainties, is represented as  $J_{RO}$ , where it is calculated at each realization in the finite space:

$$J_{RO} = E_{\theta_d}[J(T, \theta_d)] - r \cdot \sum_{\theta_d} [J(T, \theta_d)] \quad (7.6)$$

where  $E_{\theta_d}$  is the expected value over each realization  $\theta_d$ ,  $\sum_{\theta_d}$  is the standard deviation, and  $r$  is the risk aversion factor (Van Essen et al., 2006).

Since all the realization are equiprobable stochastic images, the expected value equals the average of all realizations:

$$E_{\theta_d}[J(T, \theta_d)] = \frac{1}{N_r} \sum_{i=1}^{N_r} J(T, \theta_d) = \bar{J}. \quad (7.7)$$

The final equation of robust optimization objective function can be rearranged as (Yang et al., 2011):

$$J_{RO} = \bar{J} - r \cdot \sqrt{\frac{1}{N_R - 1} \sum_{i=1}^{N_r} (J(T, \theta_d) - \bar{J})^2}. \quad (7.8)$$

The entire workflow of robust optimization procedure is outlined below (Fedutenko et al., 2013; Yang et al., 2011):

1. Creation of large number of Geostatistical realizations of reservoir properties.
2. Ranking the realizations to select the quantiles from P10, P50, and P90 that represent the overall geological uncertainty.
3. Incorporating three of the ranked realizations (less-likely, median, and most-likely) in the robust optimization algorithm.
4. Review the optimization outcome that has the optimal flow response with optimal set of design parameters.
5. Implement the uncertainty assessment via integrate all the nine ranked realizations.
6. Analyze the decision risk of the future prediction after obtaining the less-likely, median, and most-likely flow forecast.

The robust optimization flowchart is illustrated in Figure 7.2.

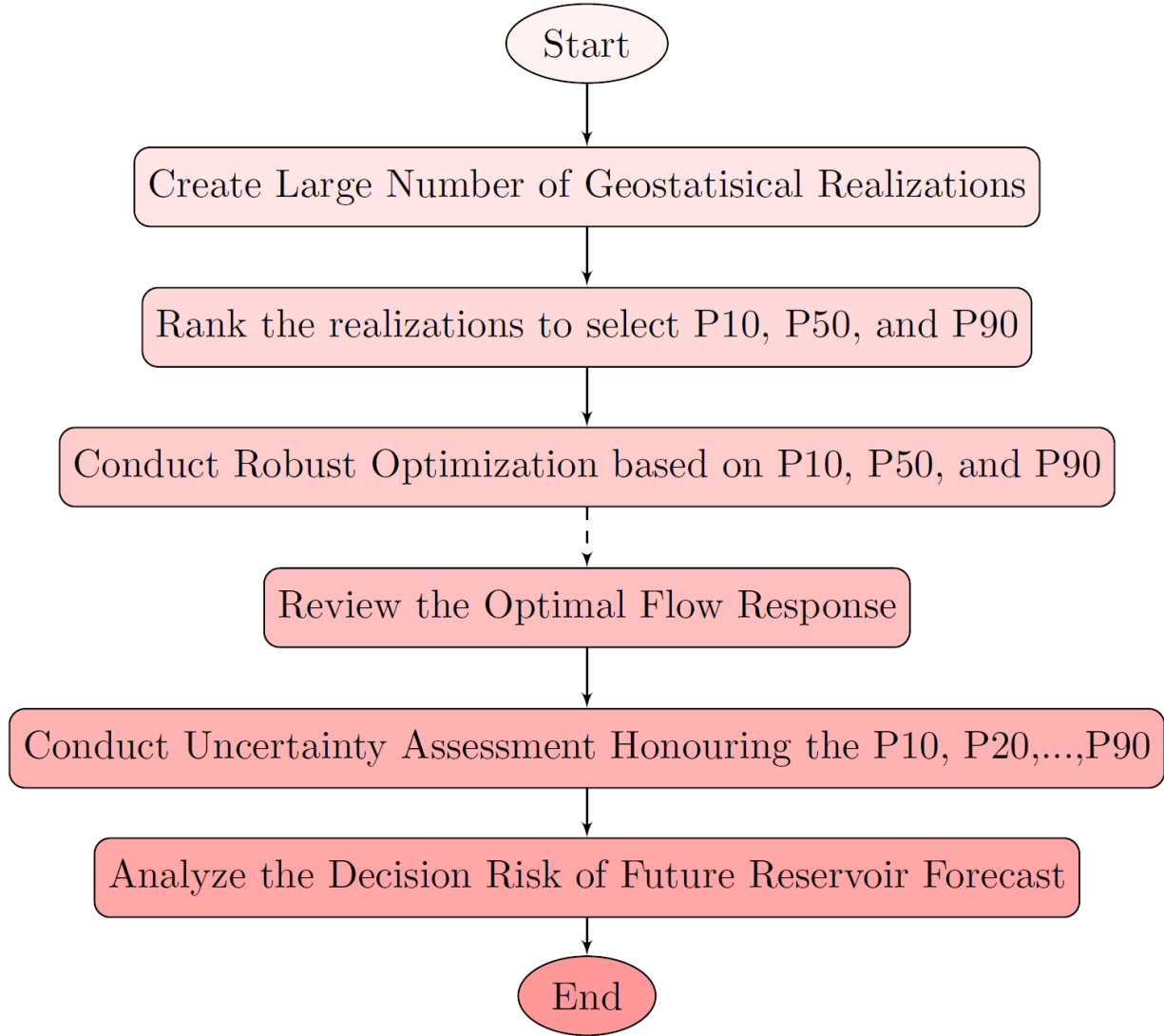


Figure 7.2: Flowchart of Robust Optimization Procedure of Huff-and-Puff GAGD Process

Three different optimization schemes were adopted to determine the optimal future reservoir performance through the GAGD process. More specifically, two consecutive optimization schemes were performed while considering the immiscible production controls optimization and the injection pressure optimization in miscible mode. After that, the optimal solution from the immiscible production controls optimization was then considered for

the immiscible Huff-and-Puff optimization to test the effectiveness of periodic gas injection through the GAGD process. The immiscible Huff-and-Puff optimization was implemented through a comparison of nominal procedure as well as robust under geological uncertainty.

## 7.4 Results and Discussion

### 7.4.1 Production Control Optimization

The six factors that were investigated for production controls optimization in the immiscible GAGD process are: maximum oil production rates (MAX\_STO), minimum bottom hole pressure (MIN\_BHP), water-cut (MAX\_WCUT), and skin factor in production wells, along with maximum gas rate (MAX\_BHG) and minimum bottom hole injection pressure (MAX\_BHP) in injection wells. Table 7.1 illustrates the values of the base case simulation of the GAGD process, which was addressed in Chapter 5, in addition to the ranges of each factor (minimum and maximum levels) in the optimization process. The levels of each fac-

Table 7.1: Parameters of the Production Controls Optimization

Response CumOilProd, STB	Min	Base Case	Max	Optimal
	4.2208E09	4.3887E09	4.6039E09	4.6039E09
Factor	Min	Base Case	Max	Optimal
MAX_STO, STB/DAY	350000	750000	625000	500000
MIN_BHP, Psia	2000	2660	3500	2000
MAX_WCUT	0.7	0.95	0.9	0.9
Skin Factor	5	0	15	5
MAX_BHG, $ft^3/DAY$	7.5E06	10E06	15E06	15E06
MAX_BHP, Psia	2000	3000	3500	3500

tors in Table 7.1 were combined by the Design of Experiments approach, which is referred to as the Latin Hypercube Plus Proxy Model, in order to produce hundreds of computer experiments (simulation jobs). Then, the designed experiments were evaluated through the compositional reservoir simulator to calculate the cumulative oil production by the end of the 10 years prediction period (January 1, 2026). The optimal solution referred to as the the simulation job that results in the maximum cumulative oil production by the end of prediction period, as shown in Figure 7.3.

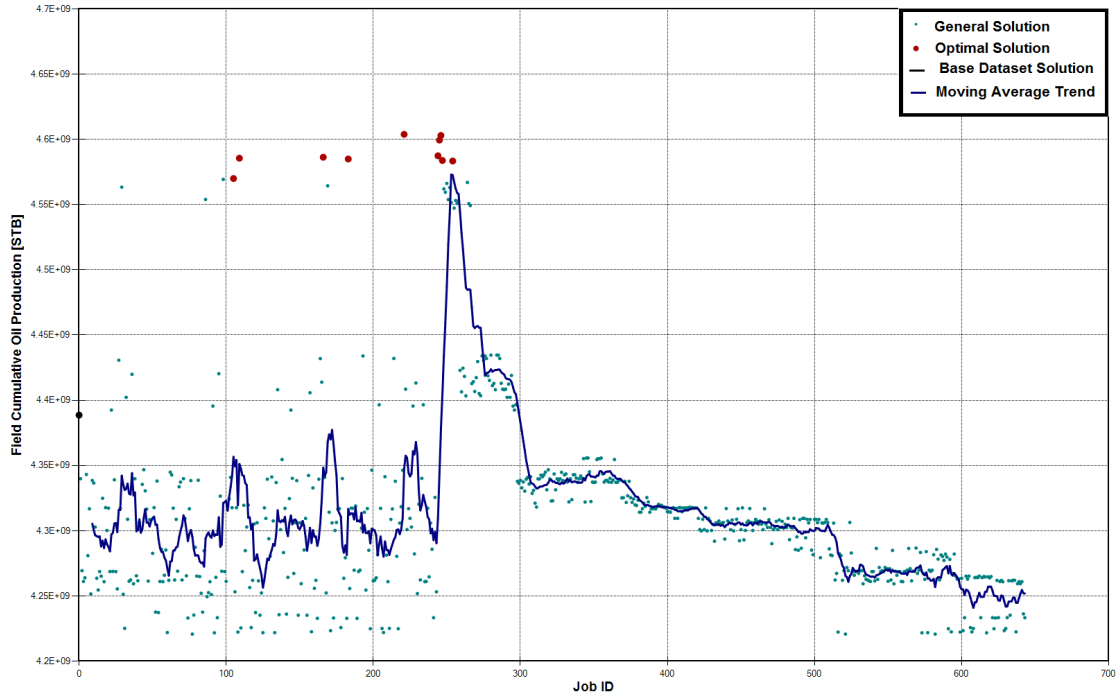


Figure 7.3: General and Optimal Solutions in GAGD Production Control Optimization

This figure also illustrates the field cumulative oil production through the base case of the GAGD process in addition to the general solutions that represent the non-optimal cases. The total number of the generated simulation jobs including the optimal solution was initially 256 runs that include the optimal solution. More than 400 extra verification runs were iteratively generated through the Latin Hypercube Sampling approach to validate the built proxy model. The proxy model is an alternative equation for the reservoir simulator that encompasses the six factors, which were included in the optimization process. The reservoir flow response, cumulative oil production, can be easily calculated by the validated proxy model for uncertainty evaluation purposes.

In this optimization case, the reservoir flow response from the calculated field cumulative oil production was statistically fitted as function of the six production controls factors given by the 256 simulation jobs (training jobs). Then, the proxy model was built in linear and non-linear second-order polynomial forms based on the training jobs. After that, the

verification jobs were generated and evaluated by the built proxy model to ensure that their values lie within the 95% confidence interval, as shown in Figure 7.4. Below illustrates the four different iterations of the polynomial proxy model verification. The linear

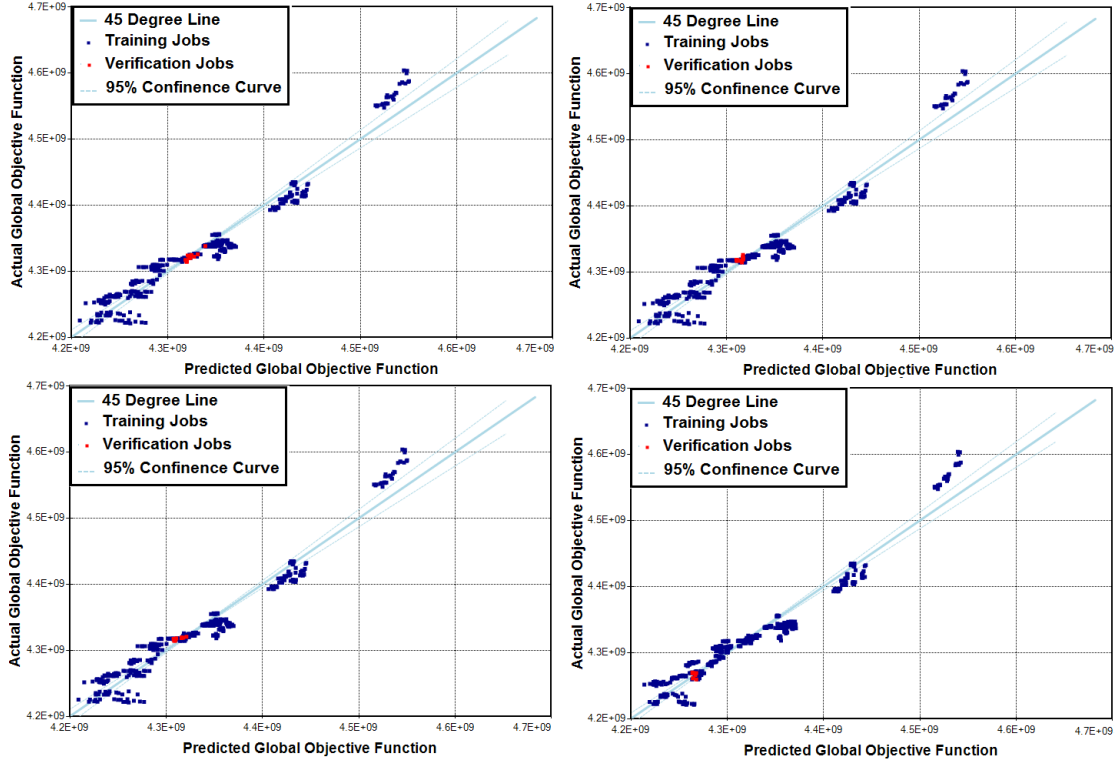


Figure 7.4: Iterative Validation of Proxy Model in the Production Control Optimization

and polynomial proxy model were again validated through Figure 7.5 that illuminates the reservoir flow response comparison that is calculated by the reservoir simulator and predicted by the built proxy model. In this figure, the red line refers to the linear proxy model with  $R^2 = 0.94$ . However, the polynomial proxy model, represented by the green line, has a higher  $R^2$  than the linear model ( $R^2 = 0.9555$ ). Since both of the proxy models have an acceptable coefficient of determination values, the current optimization process can be addressed by the proxy model to calculate the field cumulative oil production given the aforementioned six production control factors.

Additionally, the partial t-test was adopted to validate the linear and polynomial proxy models in order to determine the most influential production control factors that lead to



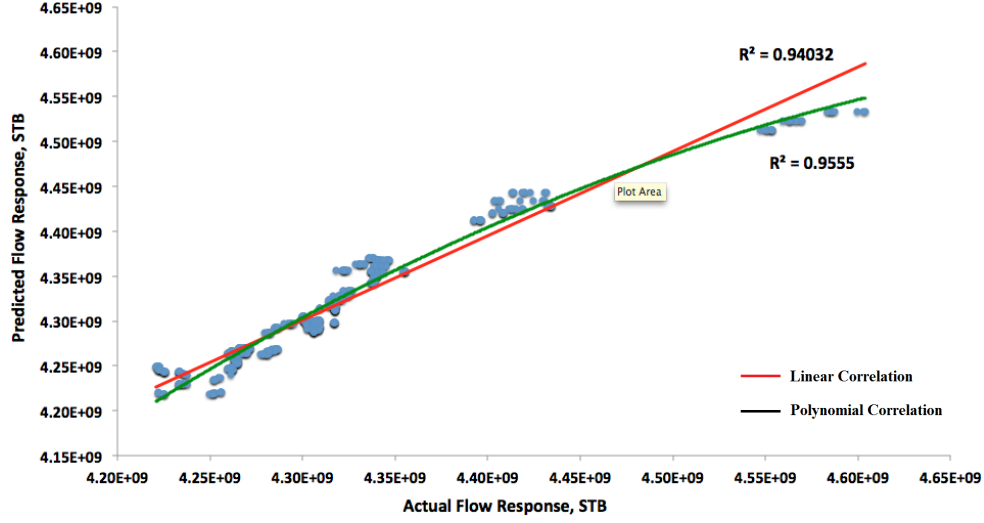


Figure 7.5: Validation of the Linear and Polynomial Proxy Models For the Production Controls Optimization Case

the optimal solution. The non-influential factors are removed from the model when they fail to reject the null hypothesis of on-effect factor, as was extensively discussed in Chapter 6. Figure 7.6 and Figure 7.7 show the tornado plot of most influential production factors inferred from full and reduced linear proxy model. It is noticeable in the last

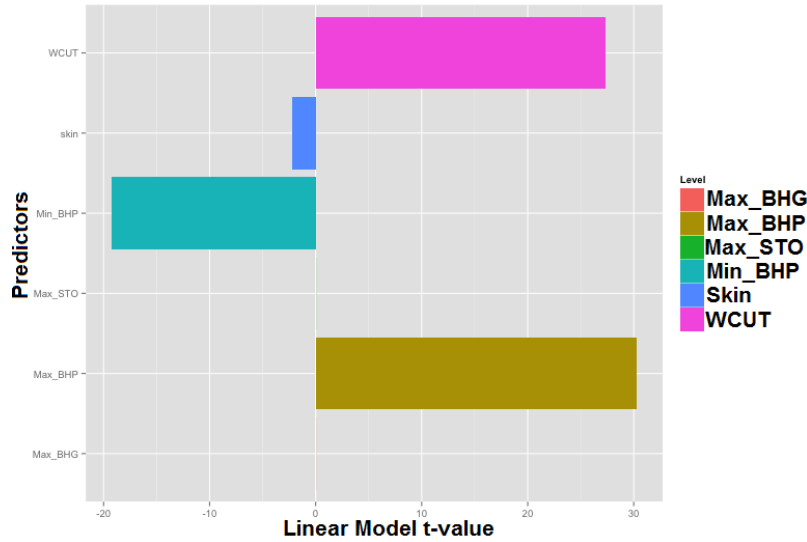


Figure 7.6: Most Influential Production Factors from the Full Linear Proxy Model

three figures that MAX.BHP in injection wells and WCUT and MIN.BHP in production wells are the most influential factor that affect the GAGD production control optimiza-

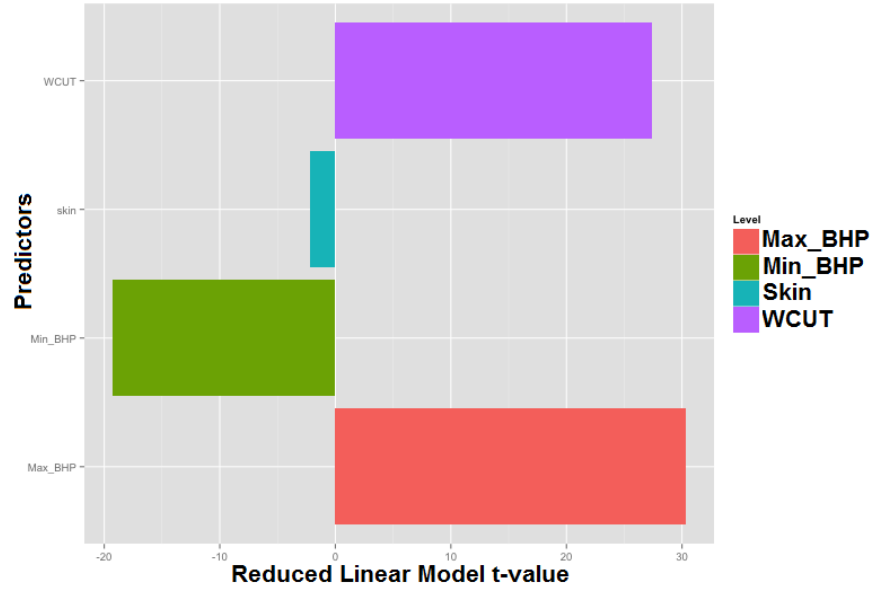


Figure 7.7: Most Influential Production Factors from the Reduced Linear Proxy Model

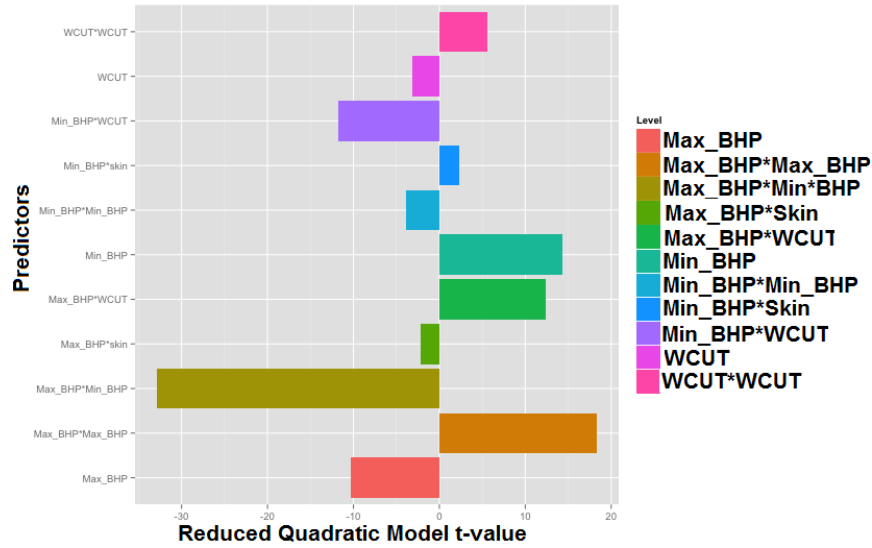


Figure 7.8: Most Influential Production Factors from Reduced Quadratic Proxy Model

tion process. Therefore, these three factors were considered for Response Surface Modeling (Proxy Model) to fully understand the effect of each of the two factors on the cumulative oil production conditioning to the third one.

Figures 7.9, 7.10, and 7.11 decorate the response surface models for MAX\_BHP and MIN\_BHP, MAX\_BHP and WCUT, and MIN\_BHP and WCUT, receptively. The red

bold color regions in the RSM figures show the optimum regions of obtaining the highest cumulative oil production as function of every two predictors. As a result, the cumulative oil production is getting higher at maximum of MAX\_BHP and WCUT and minimum MIN\_BHP.

The final form of the reduced polynomial proxy model for the production control opti-

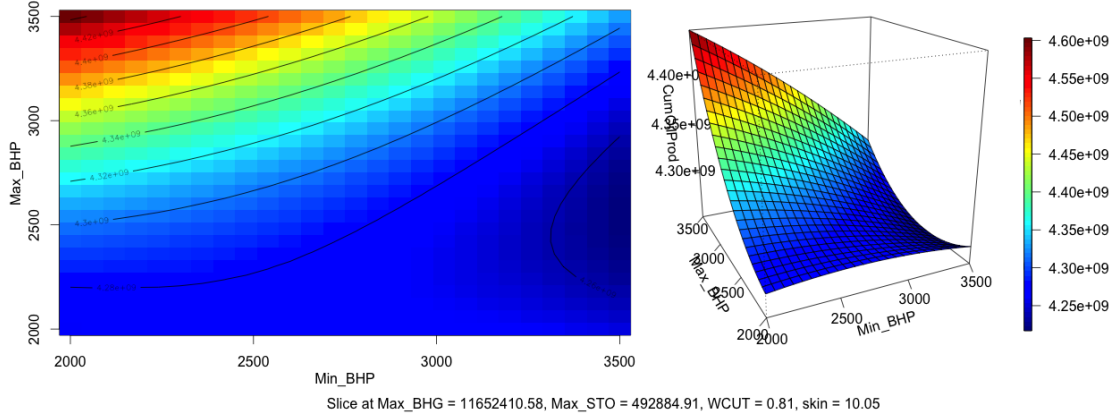


Figure 7.9: Response Surface Model between MAX\_BHP and MIN\_BHP

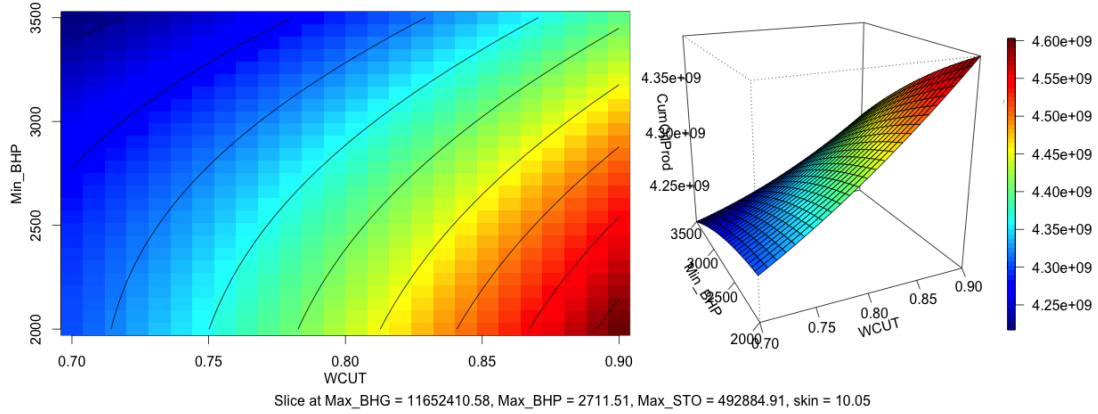


Figure 7.10: Response Surface Model between MAX\_BHP and WCUT

mization case is represented by the following equation:

$$\begin{aligned}
 GlobalObj = & 4.19345E09 - 230296 \times Max\_BHP + 308460 \times Min\_BHP - 6.80225E08 \\
 & \times WCUT + 3.18043E06 \times skin + 59.3972 \times Max\_BHP \times Max\_BHP \\
 & - 59.1208 \times Max\_BHP \times Min\_BHP + 176752 \times Max\_BHP \times WCUT \\
 & - 749.155 \times Max\_BHP \times skin - 8.92493 \times Min\_BHP \times Min\_BHP \\
 & - 183727 \times Min\_BHP \times WCUT + 533.694 \times Min\_BHP \times skin \\
 & + 8.00025E08 \times WCUT \times WCUT - 4.18306E06 \times WCUT \times skin
 \end{aligned} \tag{7.9}$$

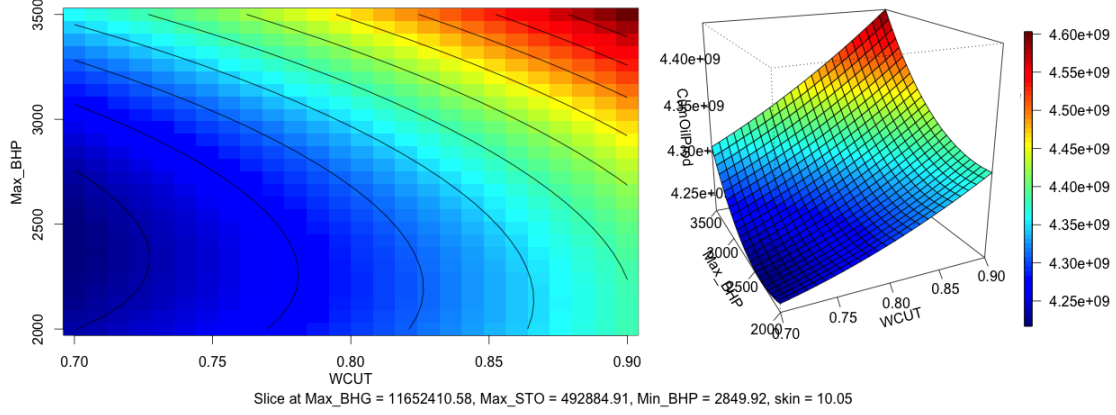


Figure 7.11: Response Surface Model between MIN\_BHP and WCUT

where:

*GlobalObj*: the cumulative oil production by the end of the prediction period.

*Max\_BHP*: the maximum bottom hole pressure in the injection wells.

*Min\_BHP*: the minimum bottom hole pressure in the production wells.

*WCUT*: the field water cut.

*skin*: the skin factor.

After validating the optimization process, verifying the proxy models, and determining the optimal production factors that affect obtaining the maximum flow response, the optimal solution needs to be identified. The comparison between the base case of the GAGD process and all other simulation jobs was performed with regards to calculating and visualizing the field cumulative oil production and oil production rate, as were outlined in Figure 7.12 and Figure 7.13, respectively.

The optimal solution represents the maximum field cumulative oil production by the end of the 10 years prediction period. This can also be identified based on the field oil rate during to the prediction period. The general solutions, represented by the green curves, in Figures 7.12 and 7.13 refer to the least flow response that combine low and/or poor combinations of the factors' levels. Hence, why they led to low field cumulative oil production. The cumulative oil production by the end of the prediction period through base case-GAGD process was 4.3887 MMMSTB. However, the optimal solution that is obtained

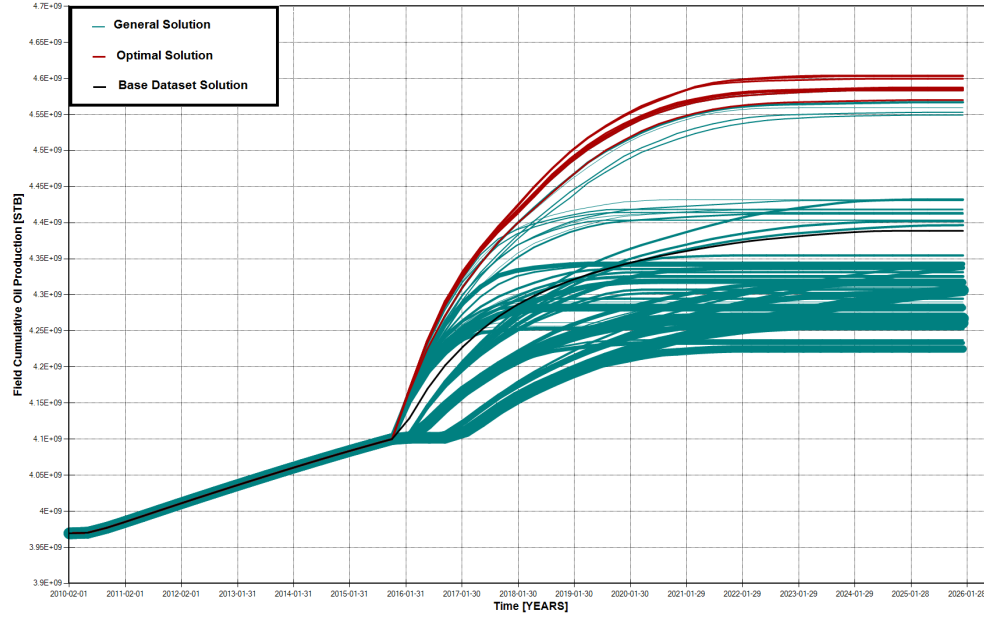


Figure 7.12: Optimal Field Cumulative Oil Production by the GAGD Production Control Optimization Case

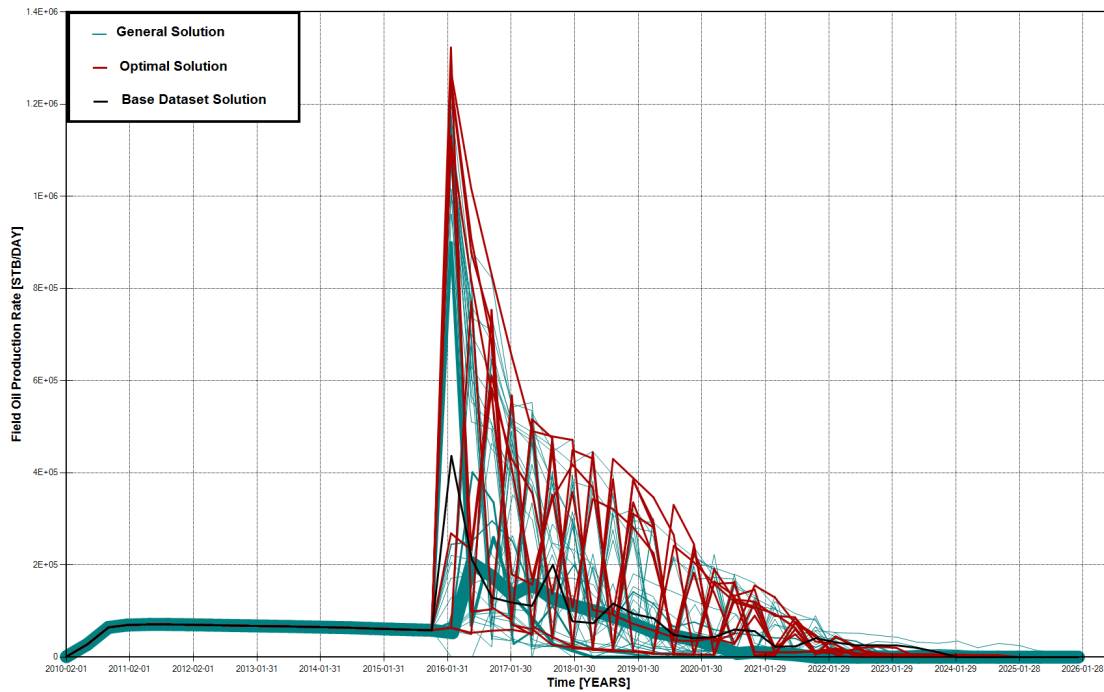


Figure 7.13: Optimal Field Oil Production Rate by the GAGD Production Control Optimization Case

from the production control optimization case (OptimalCase1) has led to increase the cumulative oil production to 4.6039 MMMSTB. The incremental oil recovery is 215.2 millions

STB, as illustrated in Figure 7.14, compares the base case and optimal GAGD process performance along with the primary production case of no injection. The optimal case has significant high levels of daily field oil production rates.

The optimal cumulative oil production was acquired by obtaining the optimal levels of all the production control factors, which are illustrated in Table 7.2.

Furthermore, Figure 7.15 shows the frequency usage of each level at each factor in

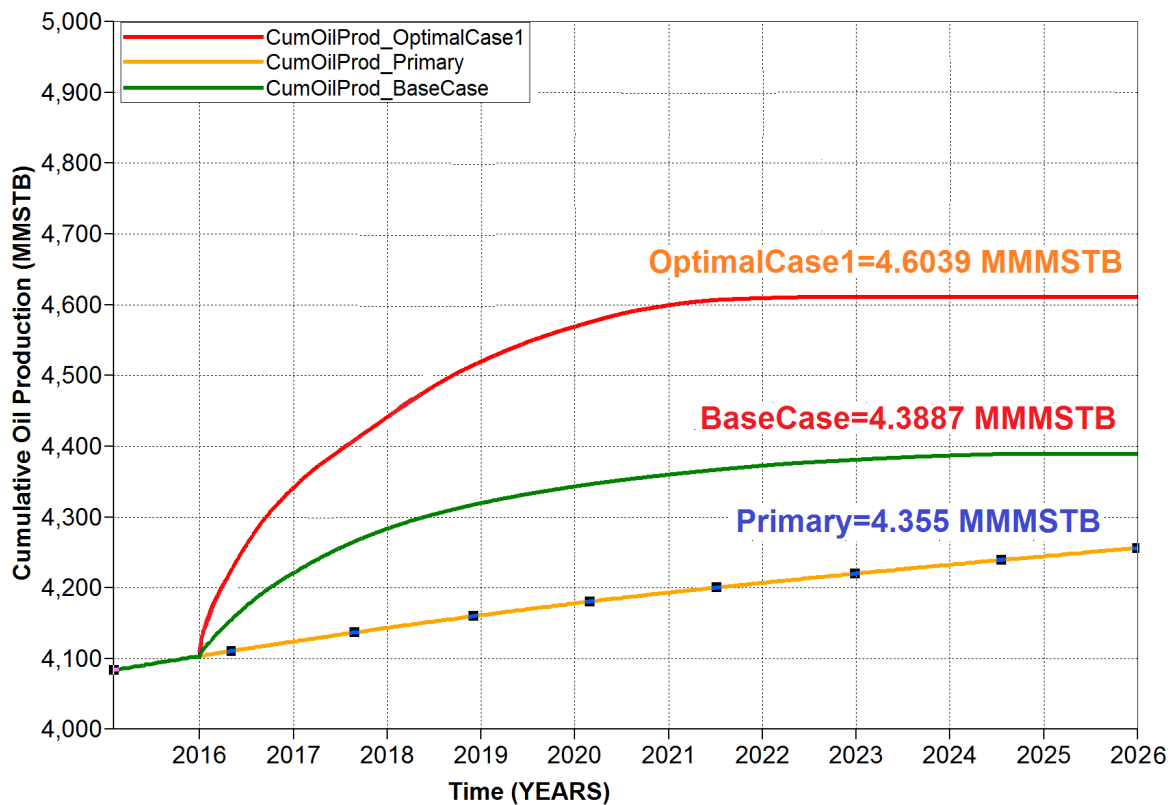


Figure 7.14: Comparison of Oil Production between Primary, Base, and Optimal Cases

the control production optimization process.

All the levels of MAX\_BHG and MAX\_STO were participated in obtaining the optimal solution. On the other hand, only the minimum level of MIN\_BHP in production wells was high frequently selected to obtain the optimal solution. However, the optimal solution frequently used the maximum levels of MAX\_BHP and maximum WCUT. It can be concluded that MAX\_BHP and WCUT are the most influential production control factors that impact the optimization process as was stated before in Figure 7.7.

Table 7.2: Optimal Levels of the Embedded Factors in the Production Controls Optimization in Comparison with the Base Case

Response	Base Case	Optimal
CumOilProd, STB	4.3887E09	4.6039E09
Factor	Base Case	Optimal
MAX_STO, STB/DAY	750000	500000
MIN_BHP, Psia	2660	2000
MAX_WCUT	0.95	0.9
Skin Factor	0	5
MAX_BHG, $ft^3/DAY$	10E06	15E06
MAX_BHP, Psia	3000	3500

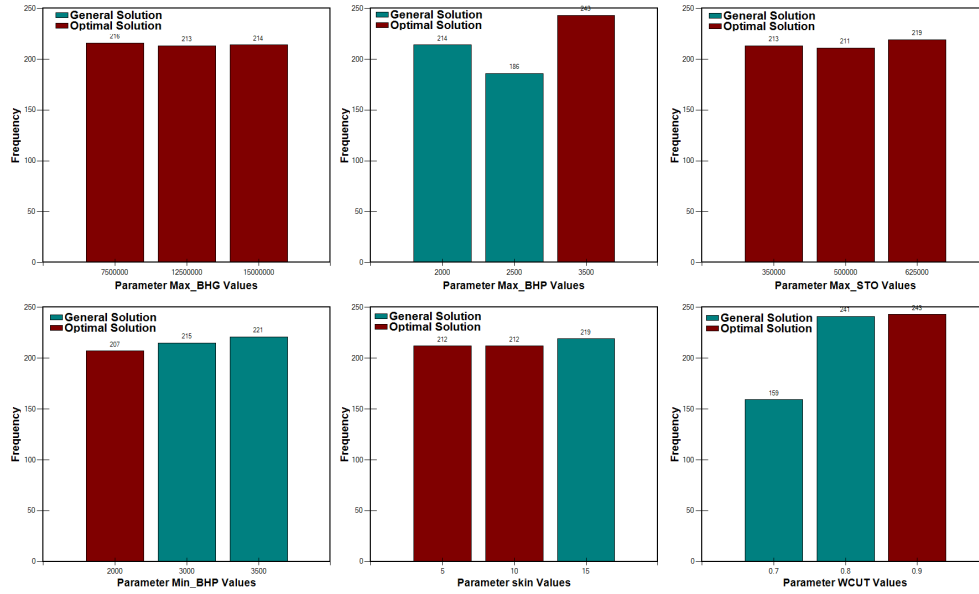


Figure 7.15: Frequency Charts of Factor Levels in the Production Control Optimization

Since (MAX\_BHP) factor has the highest impact on the cumulative oil production through the production control optimization case, it was annually changed through multiple levels to include both immiscible and miscible cycling optimization of the GAGD Process performance, as demonstrated in the next section.

#### 7.4.2 Injection Pressure Optimization

The previous section includes optimization of the Gas Assisted Gravity Drainage Process (GAGD) with regard to determining the optimal set of production control factors. It was determined that maximum bottom hole pressure (MAX\_BHP) in the injection wells is

the most influential factor that impacts the reservoir performance in the immiscible GAGD process.

In this section, the maximum bottom hole injection pressure (MAX\_BHP) factor was considered for cycling optimization of the GAGD process to encompass the immiscible and miscible modes of CO<sub>2</sub> flooding. The injection pressure was annually changed from minimum levels of zero psi to maximum levels of 5000 psi. The pressure levels less than 3500 psi were adopted to simulate the immiscible flooding. Nevertheless, the 4000 psi and 5000 psi represent the miscible injection mode.

More than 400 simulation jobs were created via Design of Experiments-Latin Hypercube Sampling Plus Proxy Model. Initially, 243 simulation jobs were created to obtain the early set of training simulation jobs. Then, 180 simulation jobs were iteratively constructed to build, validate, and verify the second-order polynomial proxy model. The overall generated experiments were evaluated by the compositional reservoir simulator to calculate the cumulative oil production by the end of the 10 years prediction year. Figure 7.16 illustrates the flow response of field cumulative oil production for the training and verification simulation runs. Unlike the proxy model that was discussed in the production control optimization case, the current proxy model does not get trapped in local optima in the verification stage because there was only one factor that was changed in different levels to obtain the optimal solution.

The validation process of the linear and polynomial proxy optimization models was verified via multiple iterations of comparison between the predicted values calculated by the proxy model, and actual global objective function calculated by the reservoir simulator. Each iterations has approximately 10 simulation jobs that were evaluated by the reservoir simulator to calculate the field cumulative oil production.

The four proxy model validations were decorated in Figure 7.17 that show how the verification jobs were located inside the 95% confidence interval. The linear and polynomial proxy model in the injection pressure optimization case were also validated through Figure



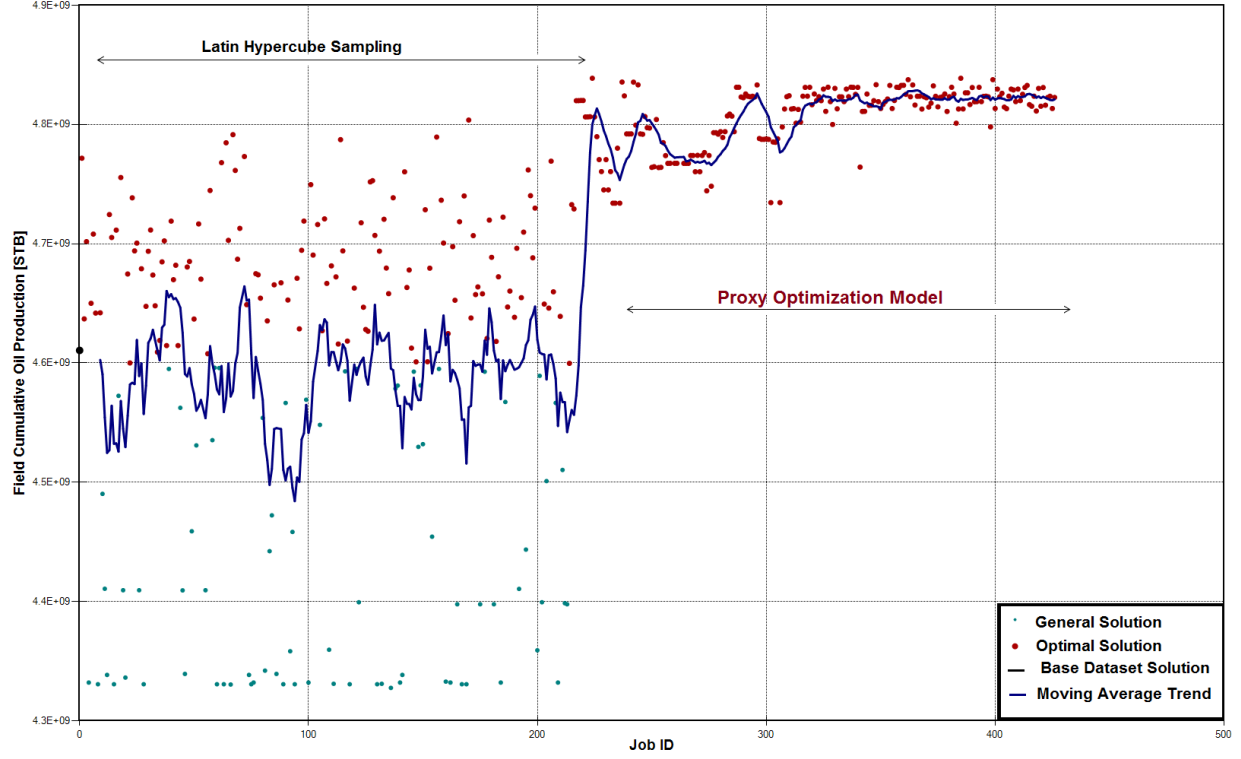


Figure 7.16: General and Optimal Solutions in GAGD Injection Pressure Optimization

7.5 that disclose the comparison between calculated reservoir flow response by the reservoir simulator and its predicted value through the constructed proxy model. In this figure, the red line refers to the linear proxy model with a value of coefficient of determination  $R^2 = 0.913$ . However, the green curve denotes to the polynomial proxy model and it has a value of ( $R^2 = 0.9245$ ) higher than the linear model. Since both of the proxy models have an acceptable ( $R^2$ ), the current optimization process can be addressed by the proxy model to calculate the field cumulative oil production given the aforementioned periodic bottom hole injection pressure.

Before determining the optimal solution that has the maximum cumulative oil production, it is essential to determine the most influential factors that impact the GAGD process through the current optimization case. The reservoir flow response was fitted to the input variables, which are the annual bottom hole injection pressure, via linear and polynomial regression models. The partial t-test was considered as an indicator for the

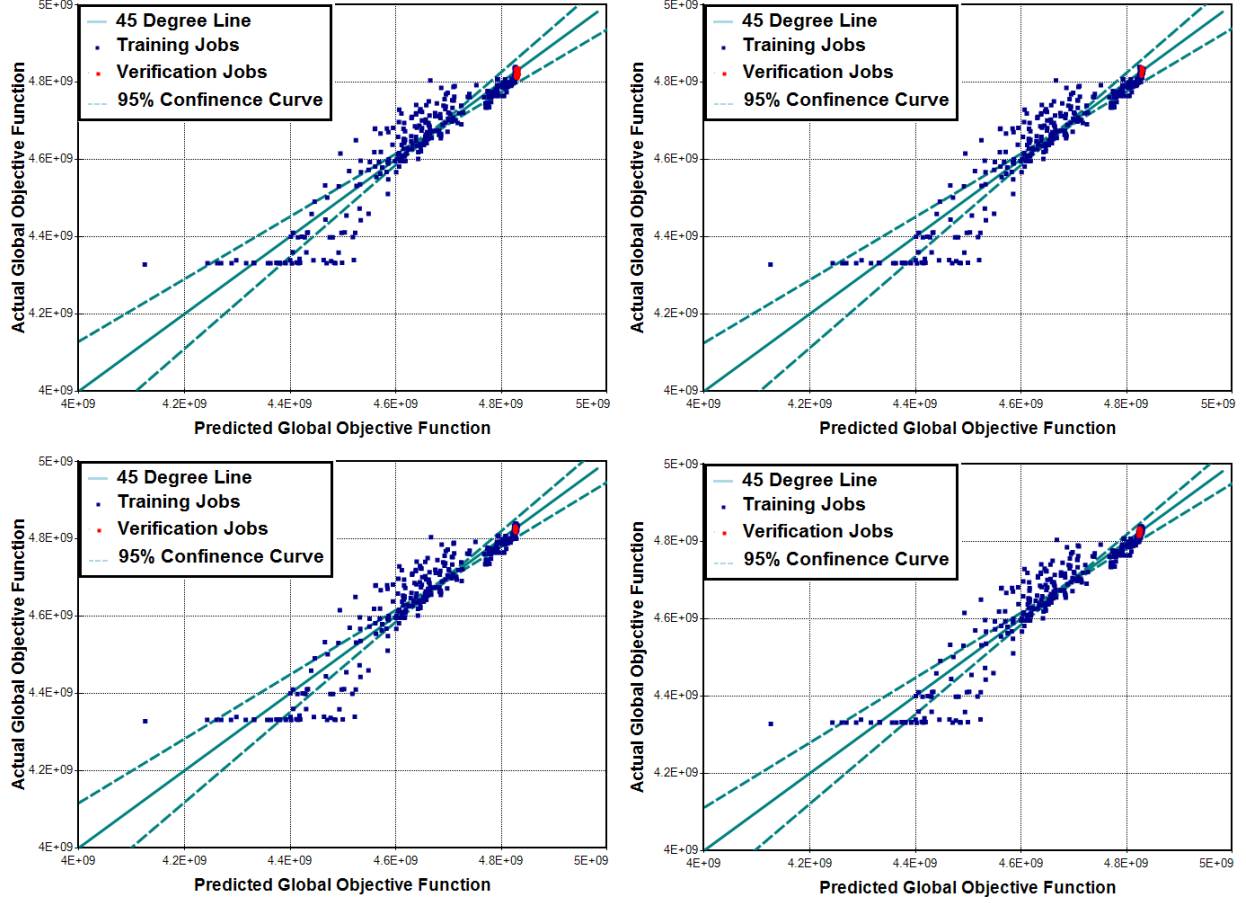


Figure 7.17: Validation of Proxy Model in the Injection Pressure Optimization

most sensitive factors. When a factor has high t-value, it highly rejects the null hypothesis of its coefficient being zero. Therefore, the highest t-values refer to the most sensitive factors. Figure 7.19 and Figure 7.20 represent the tornado plots of the influential factors in full and reduced linear proxy models, respectively.

However, Figure 7.21 depicts the most influential factors obtained by the reduced quadratic regression model. All of the linear and quadratic models indicated that MAX\_BHP in years 2018, 2017, and 2016 are the most influential factors that impact obtaining the optimal solution. The response surface models were constructed to delineate the effect of each two factors on the cumulative oil production as depicted in Figures 7.22, 7.23, and 7.24. The red bold color regions in the RSM figures show the optimum regions of obtaining the highest cumulative oil production as function of every two predictors. Therefore, the

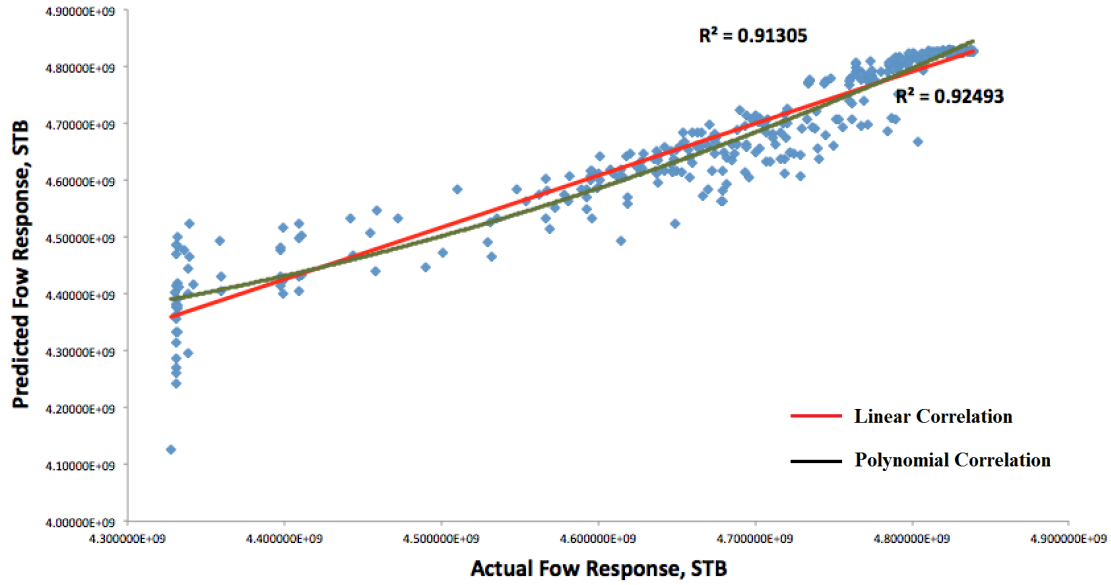


Figure 7.18: Validation of the Linear and Polynomial Proxy Models For the Injection Pressure Optimization Case

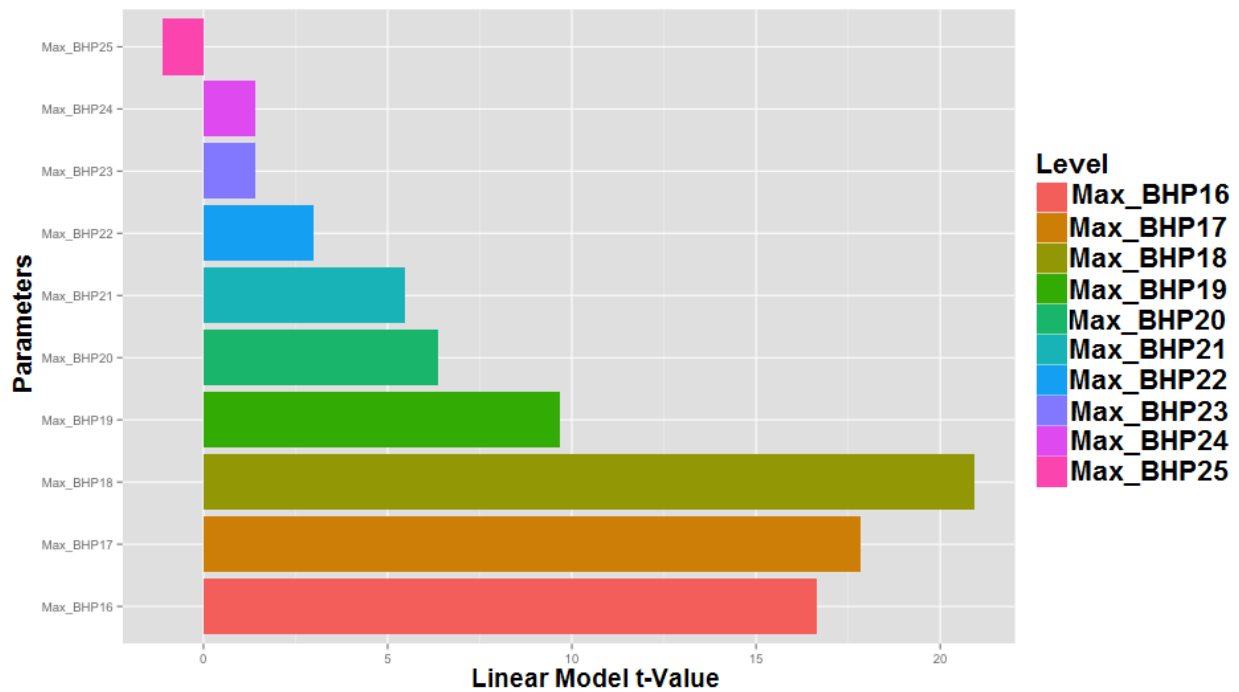


Figure 7.19: Most Influential Pressure Factors from Linear Proxy Model

cumulative oil production is getting higher at maximum of *MAX\_BHP* in periods 2018, 2017, and 2016. This fact also approves the efficiency of miscible CO<sub>2</sub> flooding to improve oil recovery through the GAGD process.

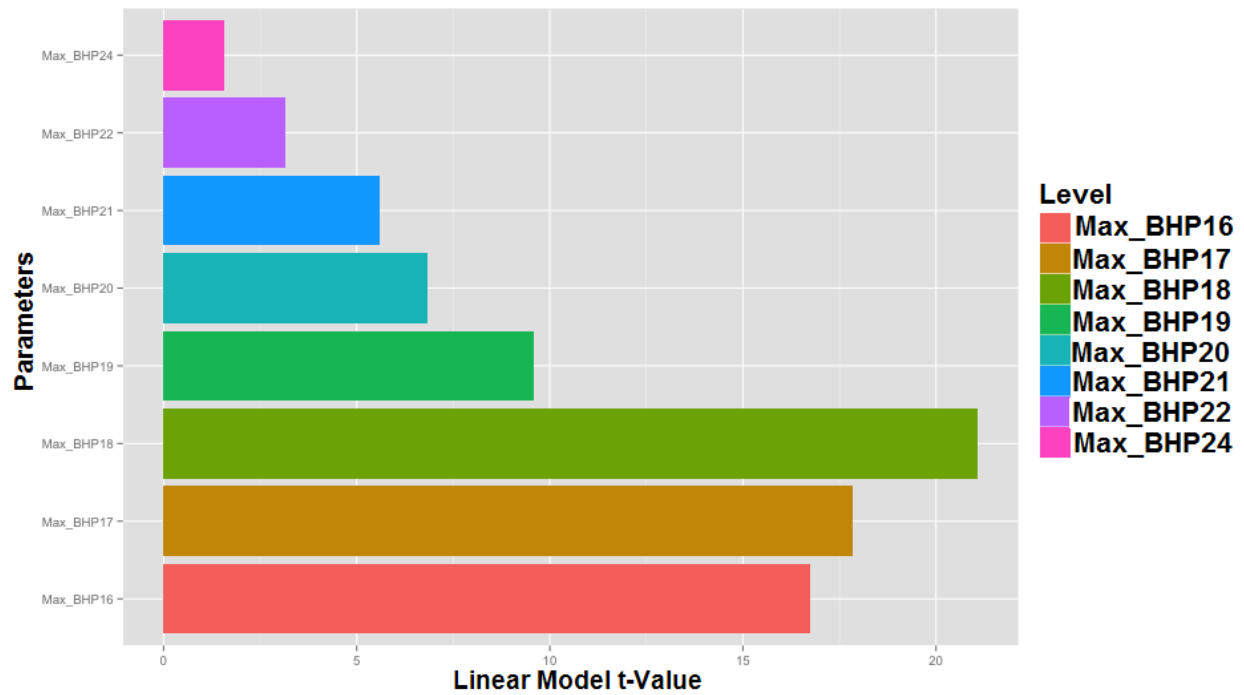


Figure 7.20: Most Influential Pressure Factors from Reduced Linear Proxy Model

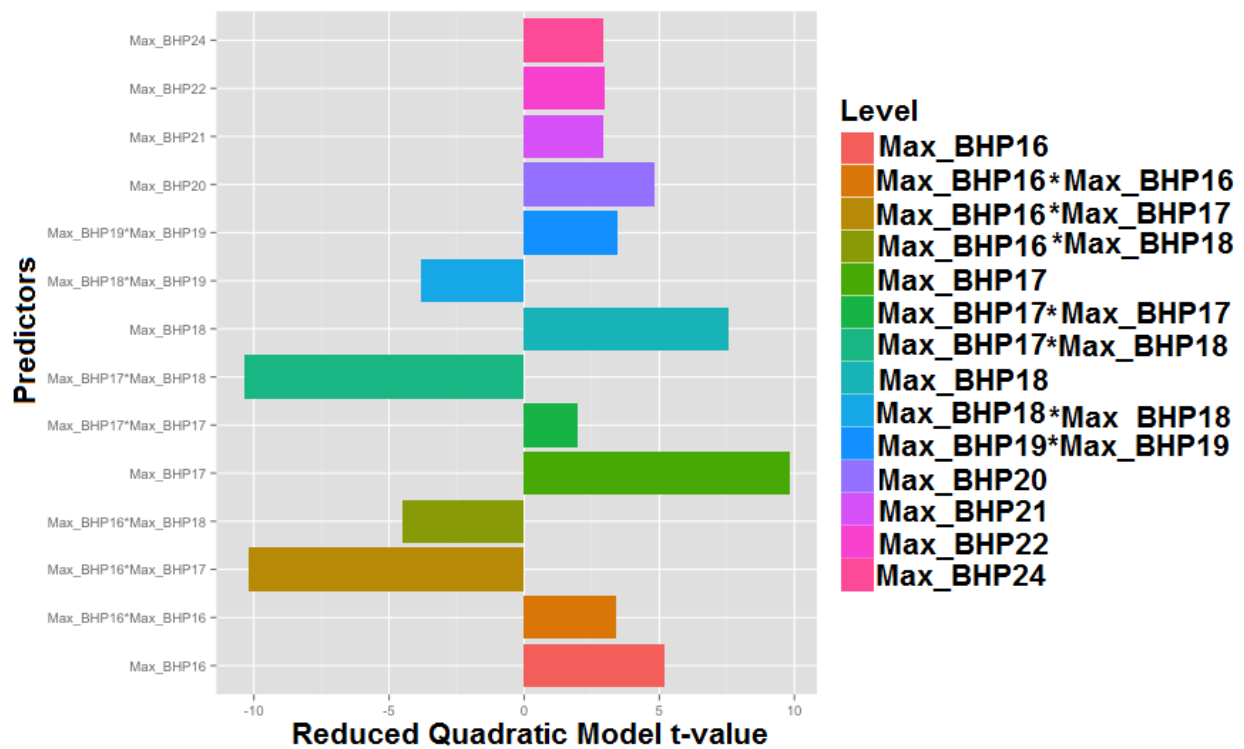


Figure 7.21: Most Influential Pressure Factors from Reduced Polynomial Proxy Model

In the current optimization case, the field cumulative oil production and oil rate were calculated to provide an indicator for the optimal solution given all the generated training and verification jobs, as shown in Figures 7.25 and 7.26.

One can notice that the optimal solution has the highest field cumulative oil pro-

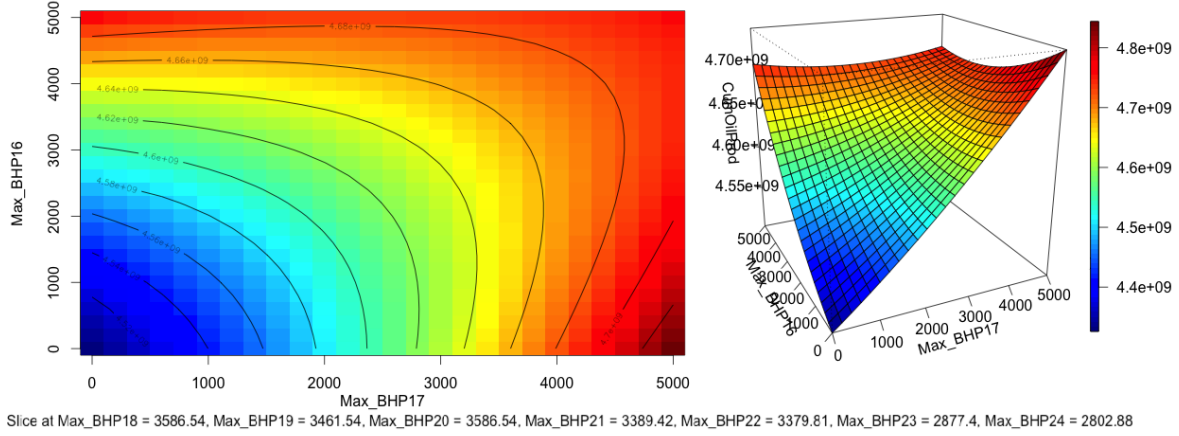


Figure 7.22: Response Surface Model MAX\_BHP in 2016 and 2017

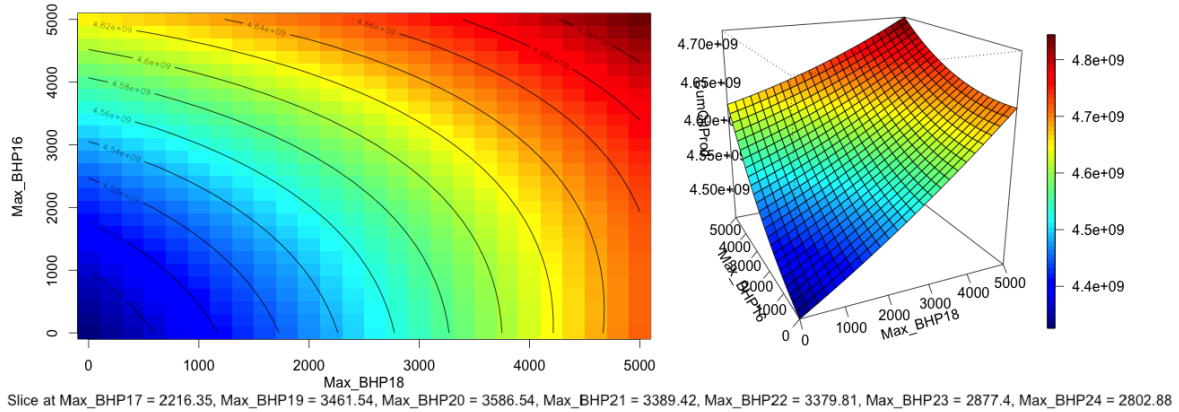


Figure 7.23: Response Surface Model MAX\_BHP in 2016 and 2018

duction and high oil rate during the prediction period, as illuminated by the red color. The optimal case from this optimization case was obtained within the miscible simulation runs because the maximum bottom hole injection pressure in injection wells was 5000 psi for the years from 2016-2014. However, the injection pressure at last year of the prediction period (2026) in optimal solution was determined through the optimization process to be



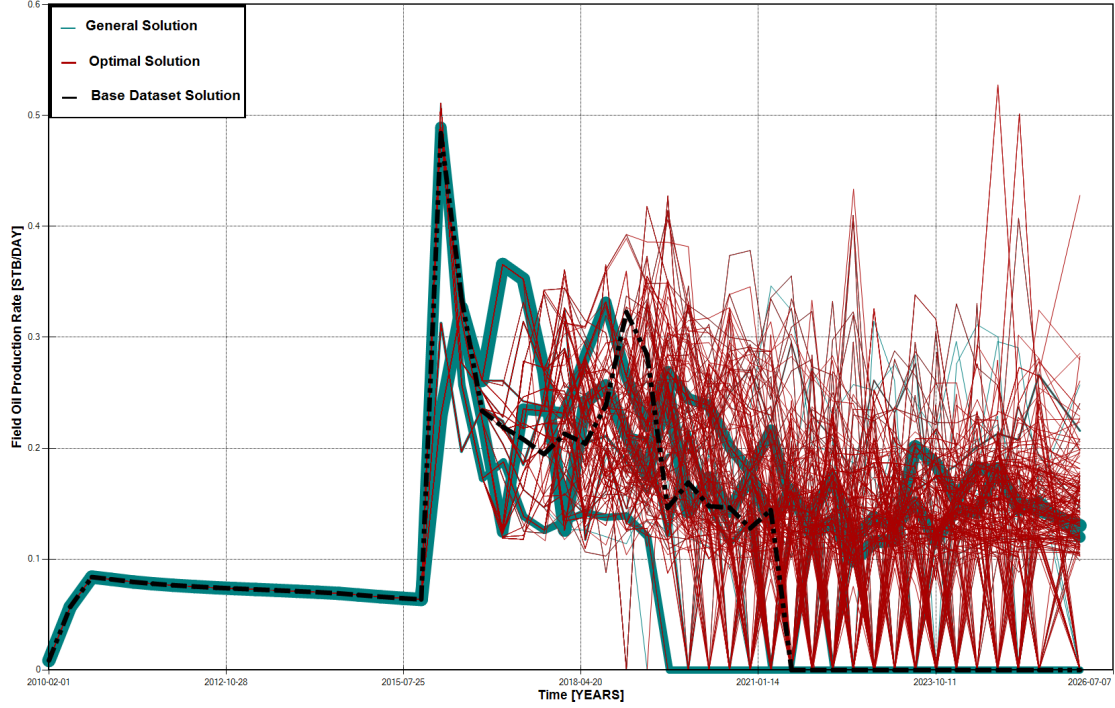


Figure 7.26: Optimal Field Oil Production Rate by the GAGD Injection Pressure Optimization Case

Control Optimization (OptimalCase1) because the current optimal case includes miscible CO<sub>2</sub> flooding during most of the prediction period. Figure 7.27 shows that there is 226.4 MMSTB as an extra increment in field cumulative oil production from this case over than the production control optimization case (OptimalCase1). Additionally, it can be seen how the OptimalCase2 has higher daily production rate than OptimalCase1 at all the time intervals of the prediction period.

The previously stated fact about the importance of the periodic gas flooding in the GAGD process can be also explained by Figure 7.28 that disclose the frequency usage of each level for the  $MAX\_BHP$  factor over all the prediction period. It is noticeable that  $MAX\_BHP = 5000$  psi is the most frequent optimal level that is used in the optimization process. That level represents the miscible CO<sub>2</sub> flooding that led to obtain the optimal cumulative oil production.

Due to economic constraints, the CO<sub>2</sub> miscible flooding may be impractical to be considered to enhance oil recovery in the oil field understudy. Therefore, new trends of op-



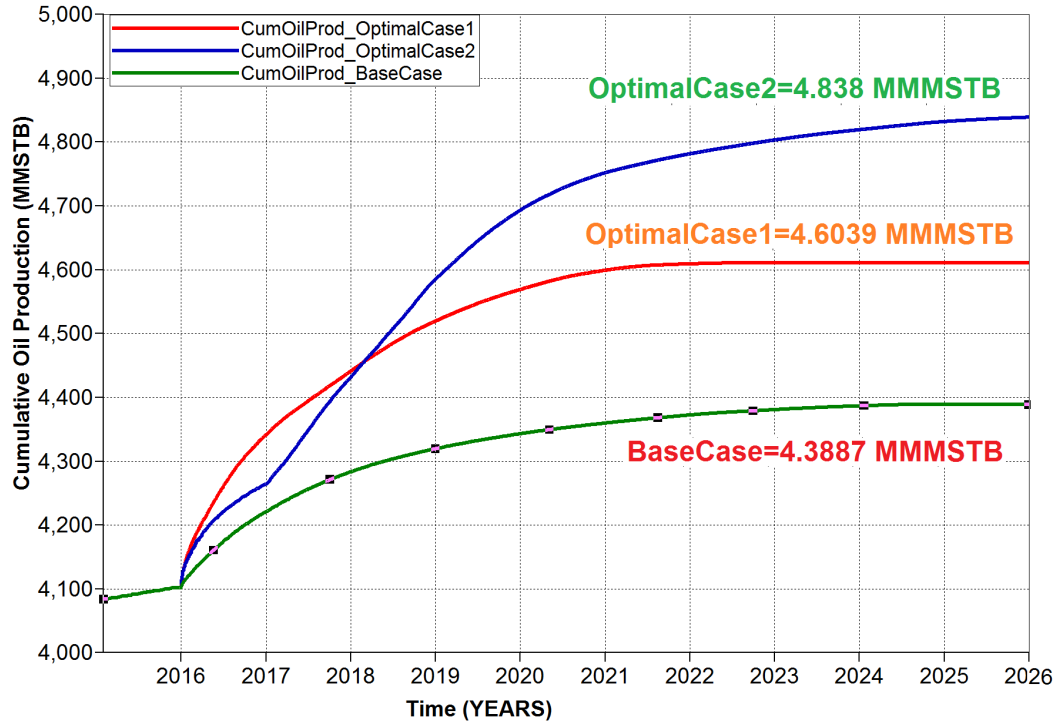


Figure 7.27: Comparison of Cumulative Oil Production between Optimization of Injection Pressure and Production Control, with the Base Case

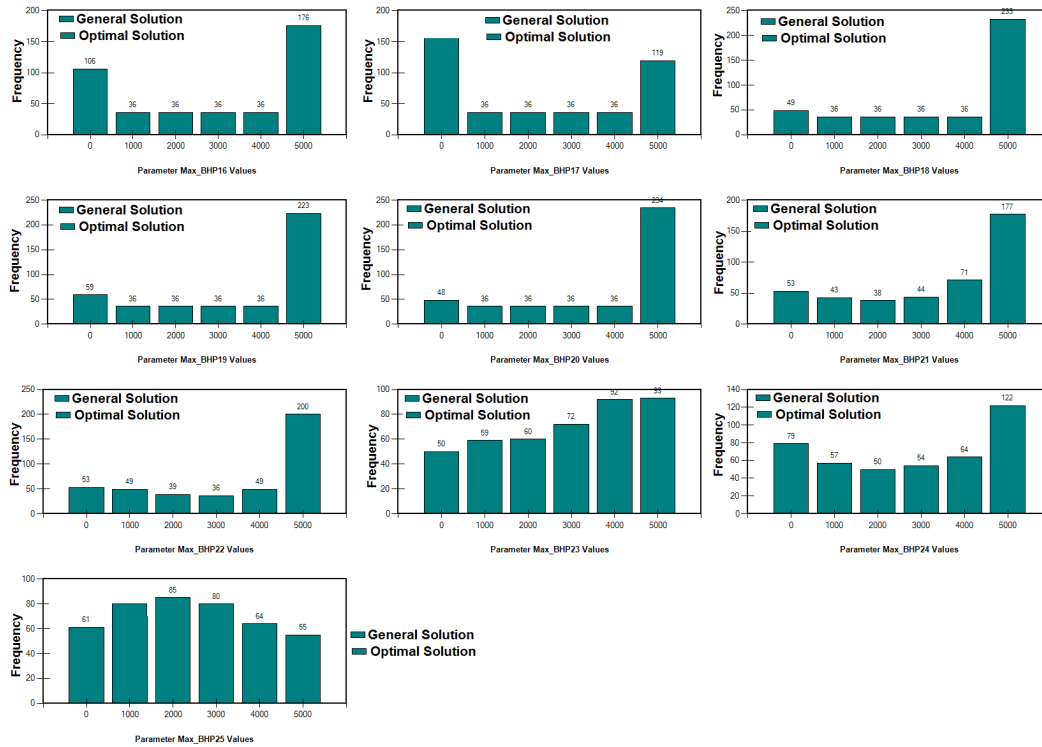


Figure 7.28: Histograms of Factor Levels in the Injection Pressure Optimization



timization are discussed in the next section to determine the optimal oil recovery through the Huff- and-Puff GAGD process with regard to nominal and robust optimization (optimization under geological uncertainty).

### 7.4.3 Optimization Results of Huff-n-Puff GAGD Process

Huff-and-Puff process is an efficient way to trigger long-term oil production through the cycling injection. It is difficult to randomly assign the injection and soaking durations. In this section, we propose an integrated approach to optimize the periods of gas injection, soaking, and oil production, along with the number of cycles through nominal and robust optimization approaches. In addition, the approach includes determining the optimal minimum BHP in production wells that changes periodically during the future prediction period.

Firstly, the nominal optimization procedure did not include the geological uncertainties. However, the robust optimization integrated the three quantiles of the geostatistical realizations of porosity and permeability. More specifically, the three realizations represent the minimum, median, and maximum levels of the 3D geostatistical property distribution.

The base case of the Huff-and-Puff GAGD process simulation led to obtain higher cumulative oil production than the original continuous base case that was implemented in Chapter 5. In particular, the cumulative oil production by the end of the prediction period in base case of continuous GAGD process produced 4.3887 billion bbls. However, the Huff-and-Puff base case with default periods setting resulted in obtaining 4.535 billion bbls. Figure 7.29 illustrates the field cumulative oil production and oil rate in the case of Huff-and-Puff GAGD process simulation (HnP) in comparison with the continuous base case of GAGD process.

In this figure, it is noticeable the advantages of Huff-and-Puff simulation of the GAGD process to obtain a significant incremental oil recovery about 142.71 millions bbls over than the continuous GAGD process base case. As a result, the nominal and robust optimization approaches, in addition to the uncertainty assessment were implemented based on the

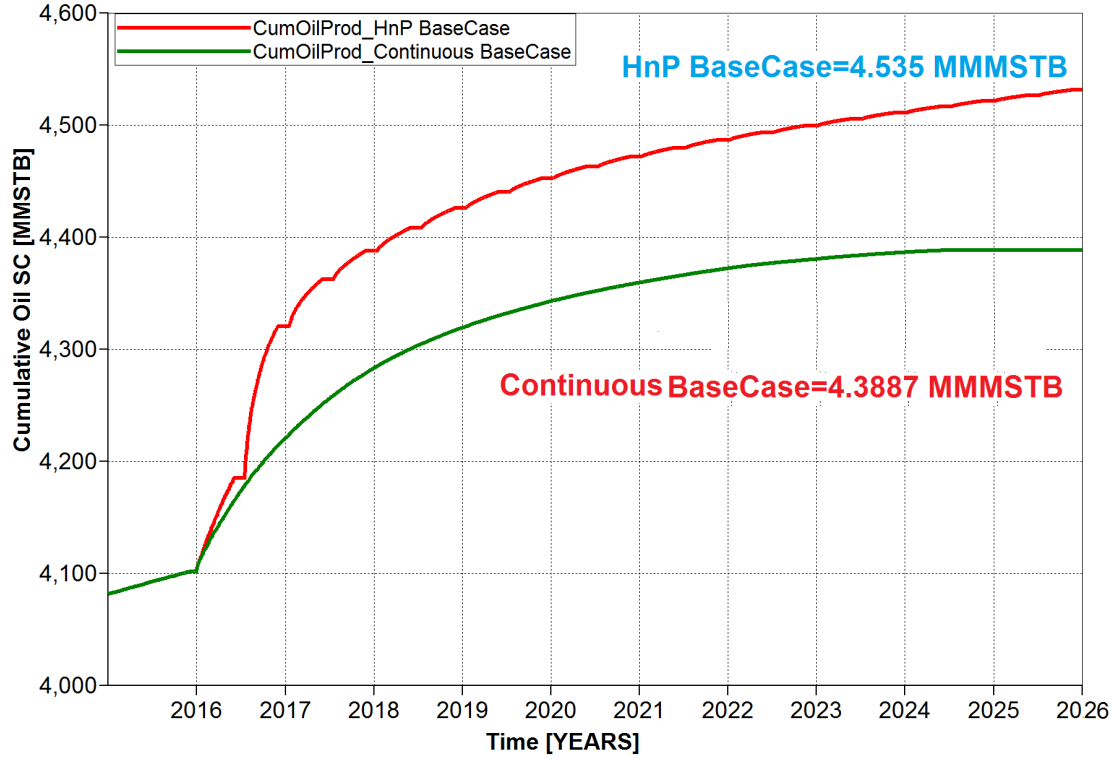


Figure 7.29: Comparison of Cumulative Oil Production between Huff and Puff Case and Production Base Case

Huff-and-Puff simulation of the GAGD process. The DECE algorithm was conducted for nominal optimization of the Huff-and-Puff GAGD process that approaches the optimal solution faster than other described Design of Experiments algorithms. To be more specific, the DECE algorithm calculates the optimal solution with smaller number of simulation runs than other DoE methods, which were adopted in the previous two optimization cases.

The base case parameters and nominal design optimization of the Huff-and-Puff simulation of the GAGD process were outlined in Table 7.3 that uncovers the minimum and maximum levels of the optimization constraints. Figure 7.30 illustrates all the simulation jobs of the Huff-and-Puff nominal optimization process with regards to calculating the cumulative oil production by the end of the prediction period. In this figure, the optimal solution was obtained in small number of simulation runs without getting trapped in local optima. The general solutions, represented by green filled circles, have low values of oil production because they were obtained from a combination of low levels of the optimization

Table 7.3: Decision Parameters of Nominal Huff and Puff GAGD Optimization

Parameter	BaseCase	Min	Max	No. of Levels
Injection period, days	30	15	60	4
Soaking Period, days	15	15	60	4
Production Period, days	137	62	152	7
Number of Cycles	1	2	3	2
MIN_BHP, psi	1000	100	2500	7

parameters. On the contrary, the red filled circles represent the frequent optimal solution with high oil production.

Exclusively, the Huff-and-Puff GAGD base case, general, and optimal solutions were

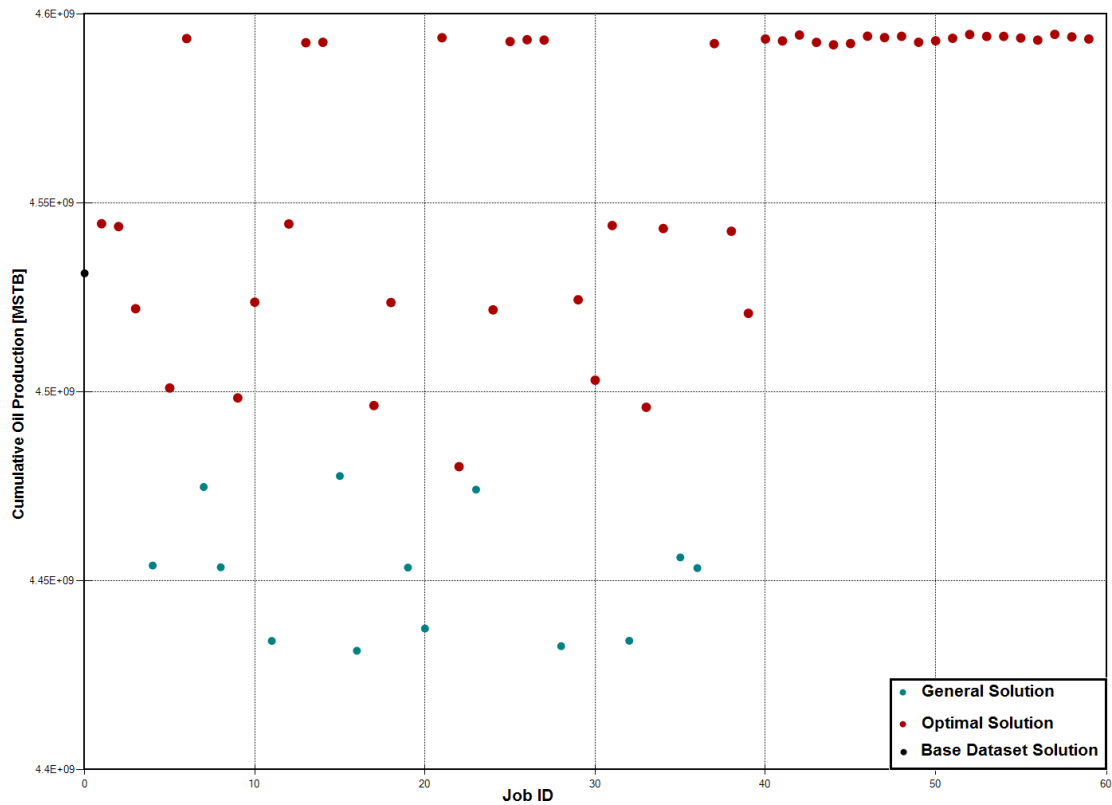


Figure 7.30: General and Optimal Objective Function of Nominal Huff-and-Puff GAGD Optimization

depicted in Figure 7.31 and Figure 7.32 with respect to the field cumulative oil production and oil rate, respectively. The optimal solution has the highest values of cumulative oil production and oil rates.

The frequencies of the selected parameters' levels in the optimization process are decorated in Figure 7.33. The high frequent parameters' levels were mostly employed in the optimization solutions. Finally, the comparison between the base case of Huff-and-Puff

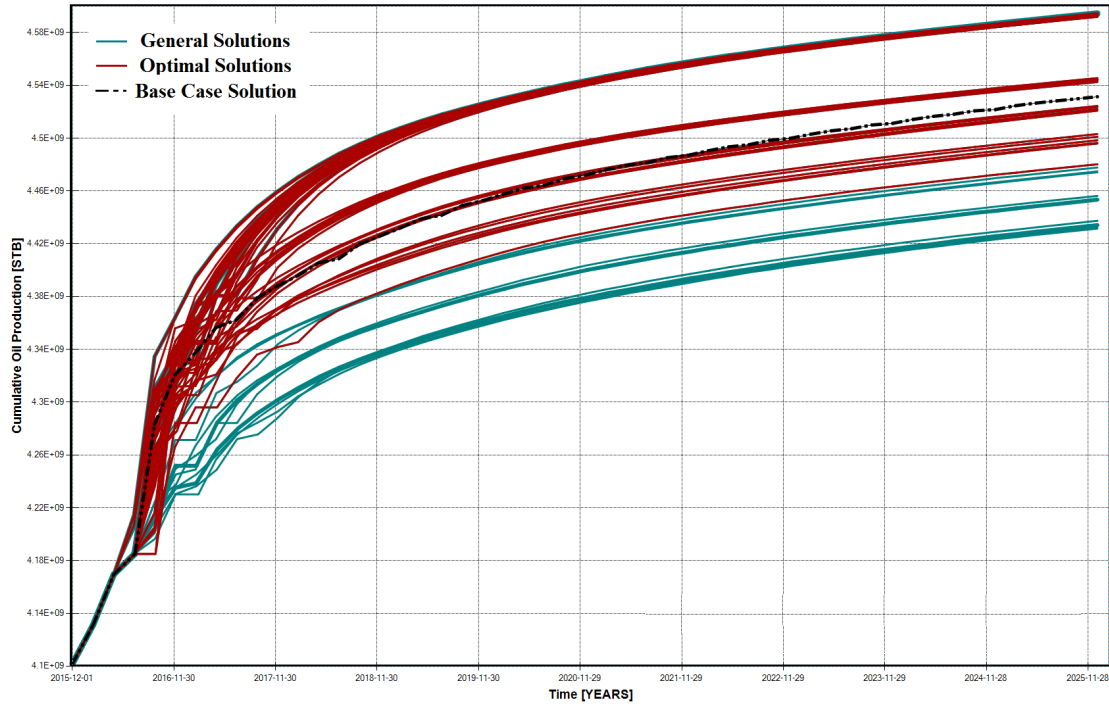


Figure 7.31: Cumulative Oil Production from Nominal Huff-and-Puff GAGD Optimization

simulation and the optimal solution of the nominal Huff-and-Puff GAGD process resulted in a significant increase in oil recovery, as illustrated in Figure 7.34. This is concluded by the daily field oil rate in the optimal solution is higher than the base case. More specifically, the optimal cumulative oil production is 4.59726 million bbls and it is 62.26 million bbls larger than the base case.

The above optimization procedure adopted only one single realization of the reservoir properties (nominal optimization). However, the less-likely, median, and most-likely geostatistical realizations of porosity and permeability were incorporated in the robust Huff-and-Puff optimization of the GAGD process in order to take into account the geological uncertainties. The same levels of injection, soaking, and production periods, which were incorporated in nominal optimization case, were again re-incorporated in the robust opti-

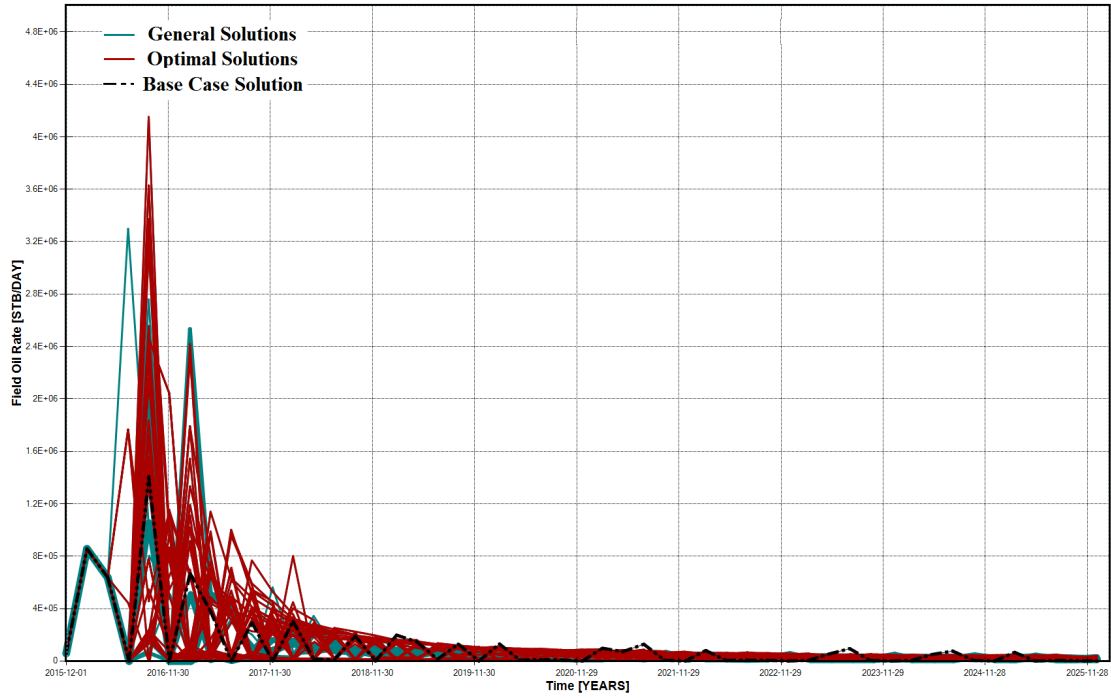


Figure 7.32: Field Oil Rate from the Nominal Huff and Puff GAGD Optimization

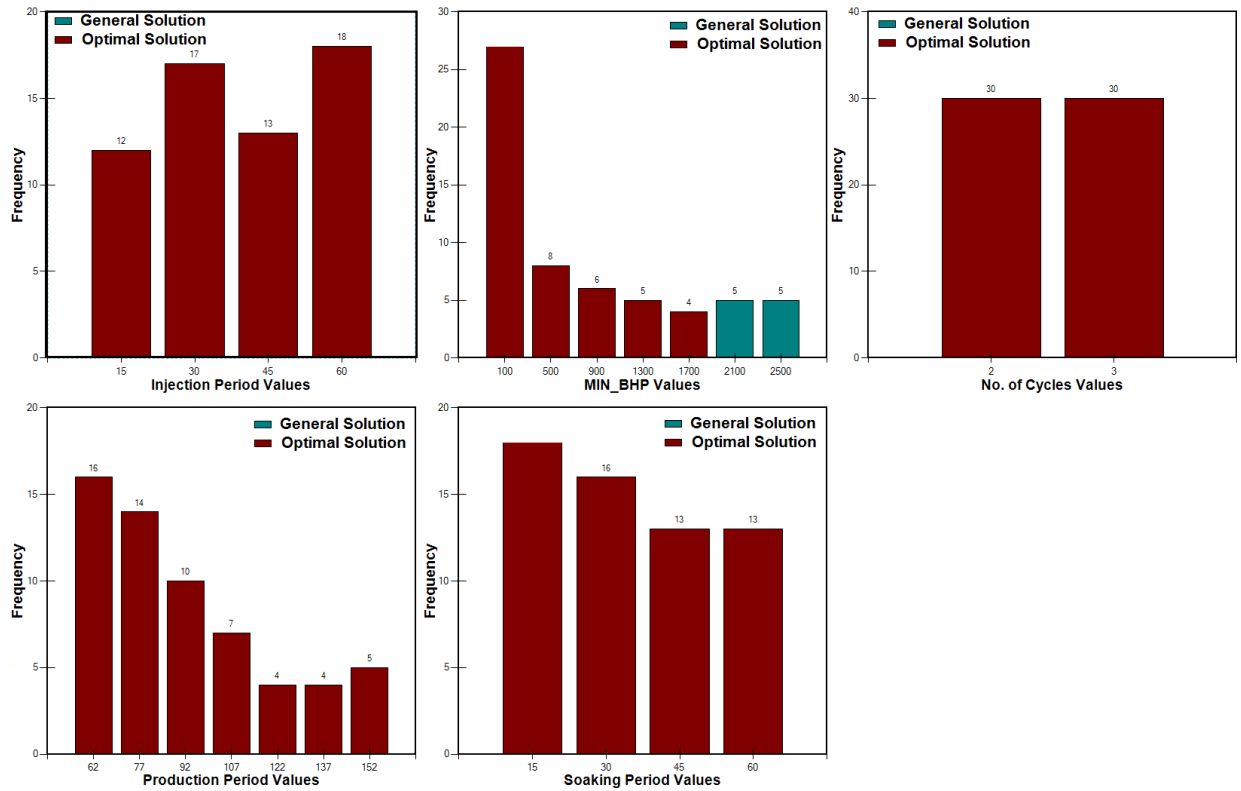


Figure 7.33: Frequency of Optimization Parameters in Huff-and-Puff GAGD optimization

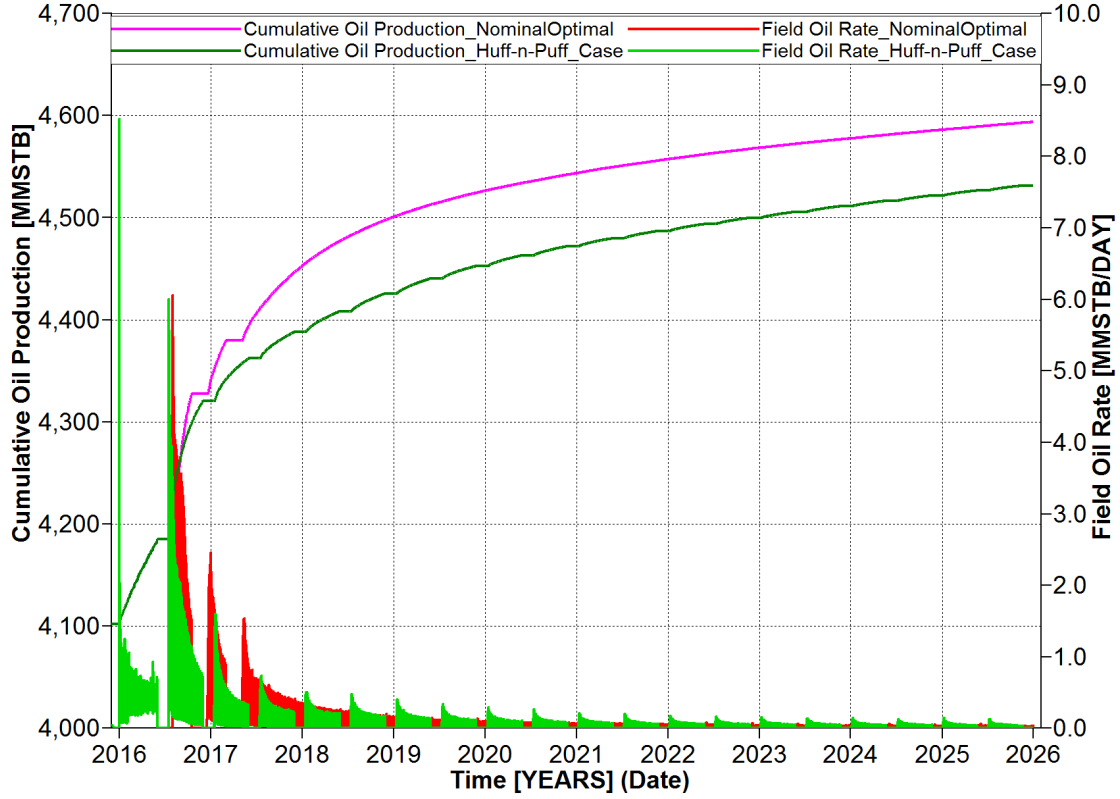


Figure 7.34: Comparison between Base Case and Nominal Optimal Flow Responses of Huff and Puff GAGD Optimization

mization. Additionally, the Latin Hypercube Sampling Plus Proxy Optimization approach was again used to find the optimal cumulative oil recovery in the robust optimization.

Figure 7.35 shows the field cumulative oil production given all the simulation jobs. The optimal solution of the highest field cumulative oil production, obtained by the end of the prediction period, is represented by the simulation jobs with ID 36 and 52. Figure 7.36 and Figure 7.37 illustrate the general and optimal solutions of the robust Huff-and-Puff GAGD process optimization in terms of cumulative oil production and field oil rate, respectively. The red curves refer to the optimal solution of the highest field cumulative oil production and the continuous high oil rate along during the prediction period.

The frequency of parameters values that were selected during the robust optimization procedure are depicted in Figure 7.38. The most frequent used levels were employed in the optimal solutions. More specifically, the most-likely levels of permeability and anisotropy

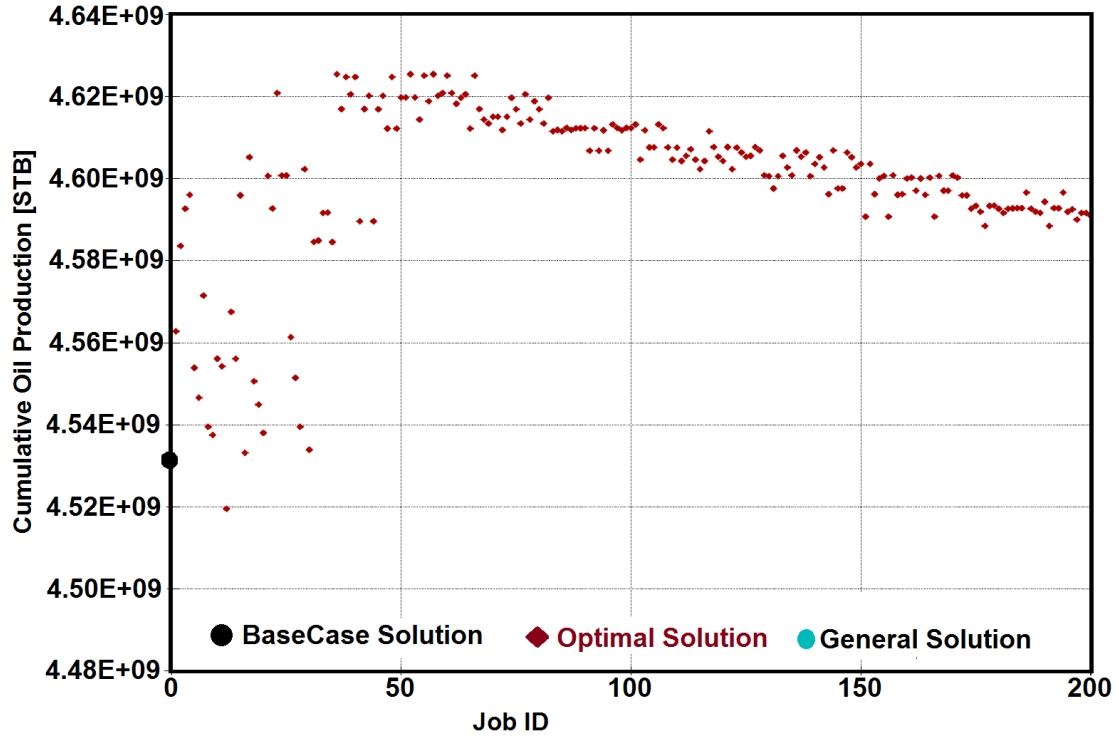


Figure 7.35: Robust Optimization of Huff and Puff GAGD Process Simulation

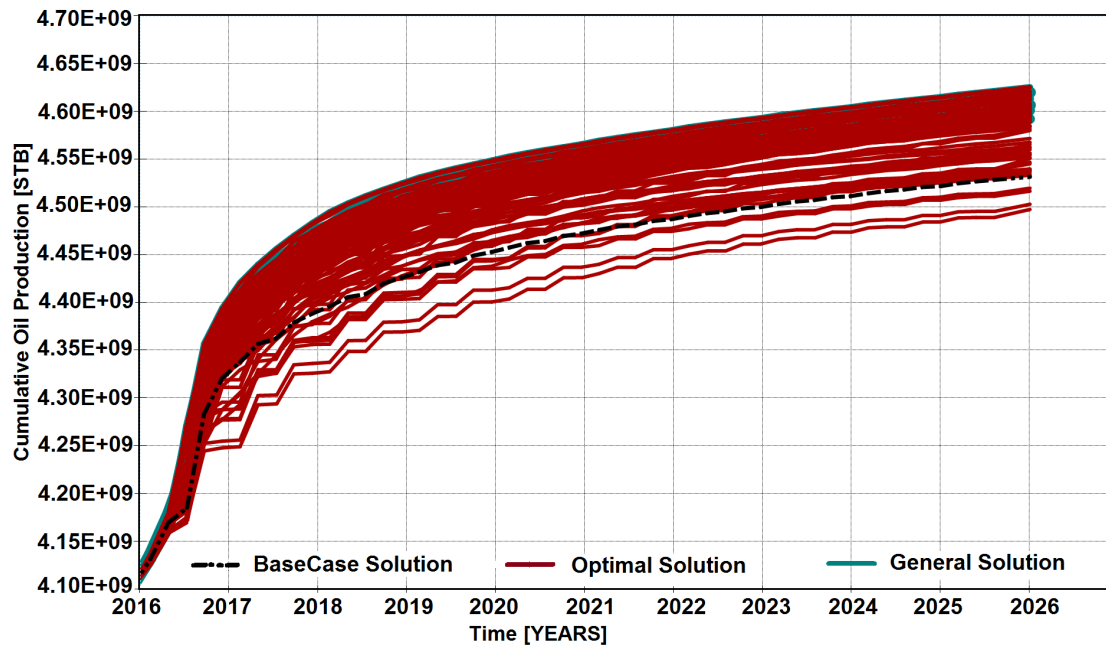


Figure 7.36: Optimal and General Cumulative Oil Production of Robust Huff and Puff GAGD Process Simulation

ratio along with the less-likely level of porosity have higher frequent usage in the selection of optimal solutions. The other levels of porosity have also contributed to obtain the optimal

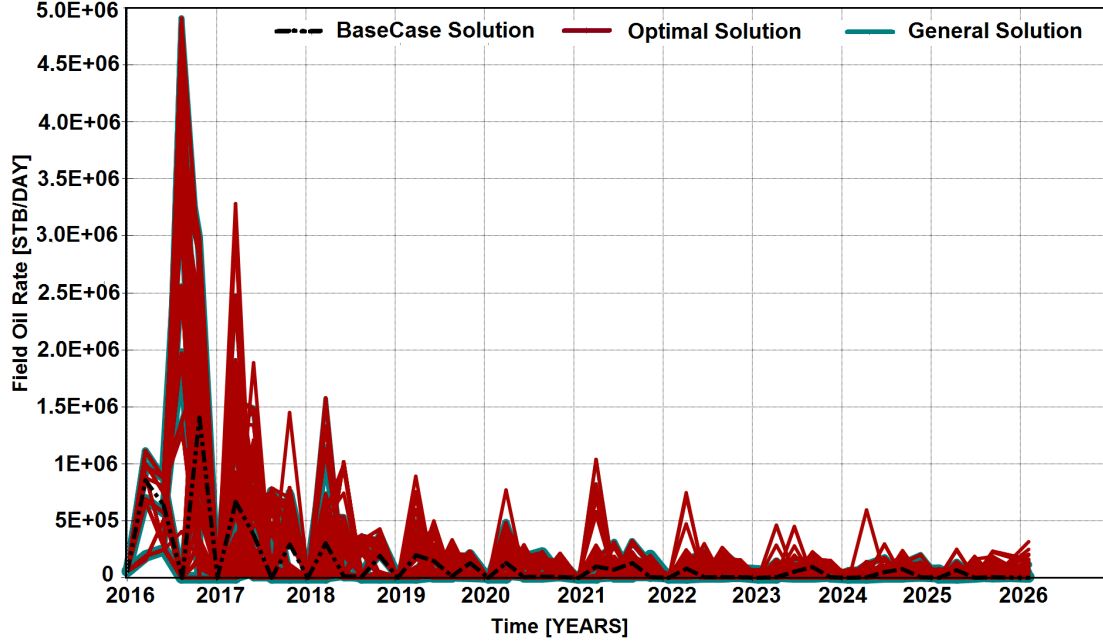


Figure 7.37: Optimal and General Field Oil Rate of Robust Huff and Puff GAGD Process Simulation

solutions because the reservoir porosity has limited effect on the reservoir flow response, as it was stated in Chapter 6. On the other hand, the low levels of the production and soaking periods in addition to the producers' minimum bottom hole pressure have the highest frequent embedding in obtaining the optimal solutions. The comparison between the nominal and robust Huff-and-Puff simulation showed that there a significant impact when considering the geological uncertainties into the optimization of the GAGD process. Figure 7.39 depicts the comparison of field cumulative oil production and daily oil rate between the nominal and robust optimization resulted from the Huff-and-Puff GAGD optimization. The robust optimization process led to the result of 4.62547 million bbls resulting in about 28.21 millions barrels of oil higher than the nominal optimization. This fact reflects the feasibility of embedding the geological uncertainties in the optimization process. The optimal set of the decision parameters in robust optimization case, including the geostatistical realizations, are outlined in Table 7.4.



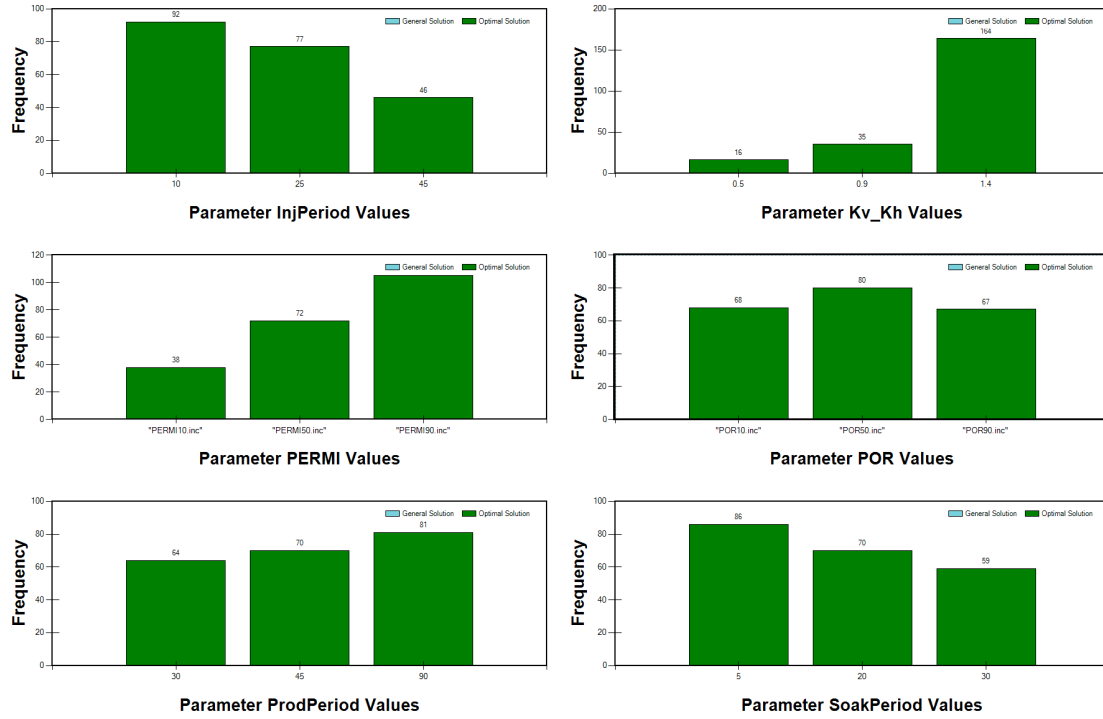


Figure 7.38: Frequency of Factor Levels in the Robust GAGD Optimization

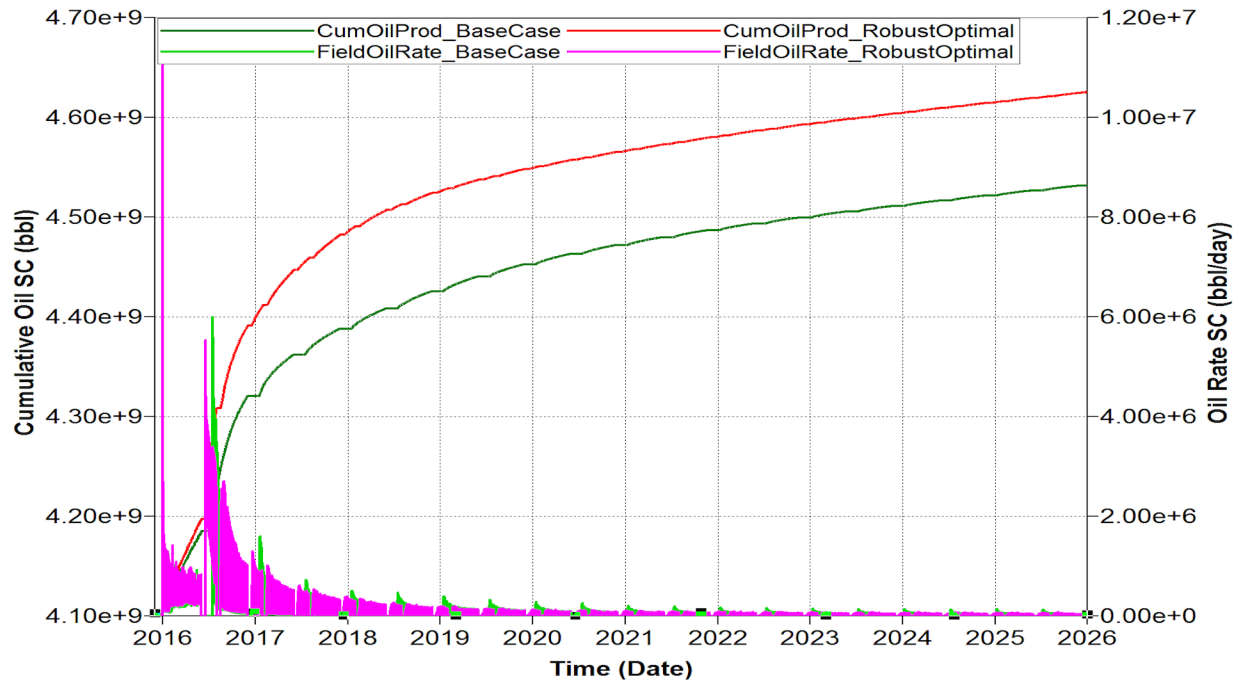


Figure 7.39: Comparison of Reservoir Flow Response between BaseCase and Robust GAGD Optimization

Table 7.4: Optimal Decision Parameters of Robust Huff and Puff GAGD Optimization

Parameter	BaseCase	Optimal
Injection period, days	30	10
Soaking Period, days	15	5
Production Period, days	137	45
Kv/Kh	0.7	0.9
PERMI, md	P50	P90
POR	P50	P50

## 7.5 Summary and Significance

The setting of production controllable parameters in the GAGD reservoir simulation can be changed to obtain different values of the reservoir flow response, especially the cumulative oil production. Three optimization approaches were conducted to determine the optimal future reservoir scenario that reaches the highest oil recovery. The optimization approaches adopted the Design of Experiments plus proxy model and DECE algorithm to generate different simulation jobs in order to determine the optimal solution.

The 1st approach includes determining the optimal set of operational design parameters through immiscible GAGD process. The parameters are CO<sub>2</sub> injection rate and maximum BHP in the injection wells, along with the maximum oil production rate, minimum BHP, skin factor, and maximum water cut in the horizontal producers. This optimal case led to obtain 4.6039 million bbls with increment of 212.5 million bbls of oil over the base GAGD case (360 Million bbls over the primary production case).

The most influential parameter from 1st optimization approach is the maximum BHP in the injection wells, which was considered in the second optimization approach as a dominant factor to control the optimization of oil recovery in both immiscible and miscible GAGD process modes. The second optimization approach involves annually changing the maximum CO<sub>2</sub> injection pressure in different levels from immiscible and miscible ranges. The purpose was to obtain the optimal set of periodic injection pressure that lead to maximum oil recovery. The 2nd approach increased to cumulative oil production to 4.838 million

bbls with 226.4 millions barrels of oil higher than the previous optimal case (438.9 million bbls of oil over the GAGD base case). The two approaches demonstrated the feasibility of the GAGD process to improve oil recovery in different parameter setting and different injection modes.

The Huff-and-Puff GAGD process optimization was conducted in this work in immiscible injection mode to test the effect of cycling injection/production on improve oil recovery. The base case of Huff-and-Puff GAGD simulation has led to trigger the cumulative oil production to be 142.71 million bbls larger than the 1st base case of GAGD process. The optimization of Huff-and-Puff GAGD process was implemented based on nominal and robust optimization approaches. Both approaches include sampling the Huff-and-Puff parameters of injection, soaking, production periods, and number of cycles in addition to the minimum bottom hole pressure. The nominal optimization led to obtain 62.26 Million bbls larger than the Huff-and-Puff base case.

The robust optimization that takes into account the geological uncertainty in the Huff-and-Puff optimization process. The less-likely, median, and most-likely of the ranked realizations were incorporated in robust Huff-and-Puff optimization of GAGD process. The optimal solution has increased the cumulative oil production to be 28.21 million bbls of oil larger than the nominal optimization. The DECE optimization algorithm was conducted in Huff-and-Puff GAGD simulation because it faster than other experimental design methods. DECE algorithm remove the weak levels that badly affected the objective function values. Therefore, it avoids from trapping in local optima.

The Huff-and-Puff base case has a higher cumulative oil production than the base case, which was addressed in the production control optimization. In particular, the periodic injection, soaking, and production activities led to trigger the oil production during the prediction period. Additionally, incorporating the geological uncertainties in the immiscible Huff-and-Puff GAGD process resulted in obtaining higher oil recovery than the optimal case of immiscible production control optimization case (OptimalCase1). More specifically,

the cumulative oil production of 4.62547 million bbls predicted from the Huff-and- Puff optimization process was higher than the 4.6039 million bbls obtained from the production control optimization case.

## 8. Uncertainty Assessments

### 8.1 Introduction

Uncertainty assessment in reservoir simulation studies is a crucial task to reduce the risk regarding the future reservoir performance. The risk of making a decision with regards to future reservoir prediction comes from uncertainty in the reservoir data used to build the reservoir model (Guyaguler and Horne, 2001). The uncertainty is quantified through a large number of equiprobable reservoir stochastic images (realizations) with different sets of dynamic data to obtain a flow response distribution through the Cumulative Probability Function (CPF). Then, nine quantiles, which represent a range of realizations from less-likely to most-likely, are collected in order to generate the same range of the calculated field cumulative oil production and/or Net Present Value (Liu et al., 2001). As result, uncertainty quantification includes three main steps in integrated reservoir simulation studies through experimental design approaches: sensitivity analysis for parameters screening, reservoir modeling, and Monte Carlo simulation (Carpenter, 2013).

In this chapter, the geological uncertainty was quantified for the current reservoir model to reduce the risk regarding the optimal cumulative oil production obtained through the GAGD process. In particular, the uncertainty assessment was quantified through the optimal Huff-and-Puff GAGD process, which was obtained from Chapter 7, in terms of geological and production parameters. The purpose was to determine an actual optimal solution through cycling optimization of the CO<sub>2</sub>-GAGD process in a giant heterogeneous sandstone oil reservoir, South Rumaila field.

### 8.1.1 Box-Behnken Designs

Box-Behnken design is an independent design, which include neither factorial nor fractional factorial designs, that mostly fits to the quadratic response surface design (Montgomery, 1997). Box-Behnken design creates the experiments by selecting combinations at the midpoints and centers of the process space edges. The number of factors and levels should be more than two to be treated in the Box-Behnken design (Kalla and White, 2007).

Figure 8.1 depicts a Box-Behnken design for three factors example.

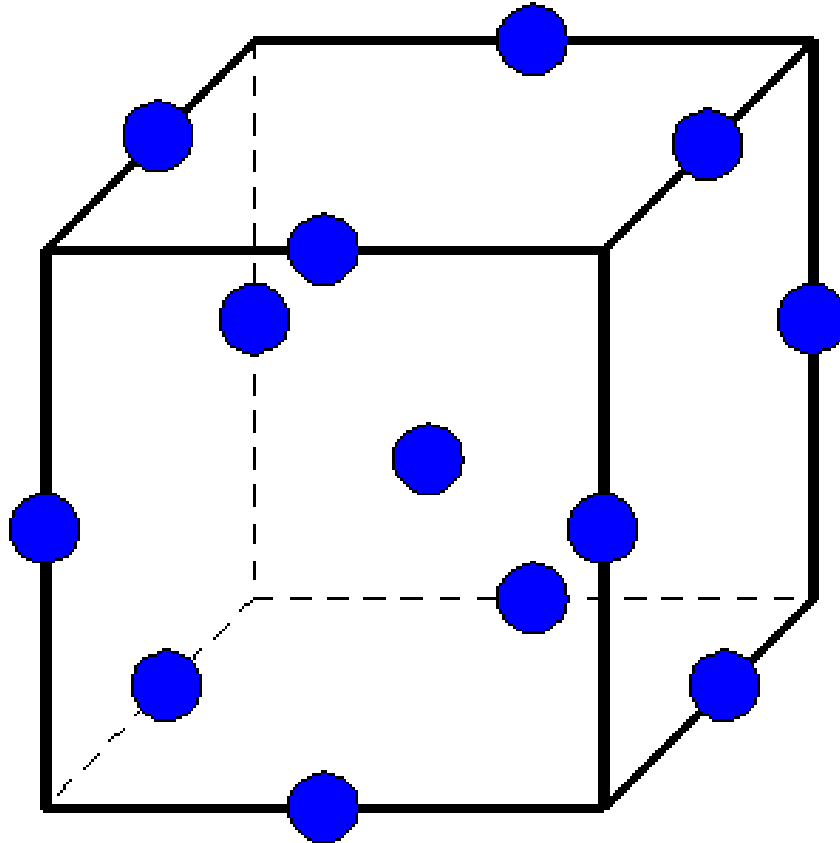


Figure 8.1: Box-Behnken design for Three Factors Example (NIST, 2012)

Box-Behnken design was adopted for the production uncertainty assessment given the most-likely model of the Huff-and-Puff GAGD process, which resulted after quantifying the geological uncertainty assessment.

## 8.2 Geological Uncertainty

The uncertainty assessments were successively quantified to count the further geological and operational design parameters uncertainties through the Huff-and-Puff GAGD process optimization. The geological uncertainty assessment was conducted by embedding the nine quantiles of the geostatistical realizations (p10, P20,..., P90) for permeability and anisotropy ratio in the GAGD reservoir simulation. The nine realizations of permeability were depicted in Figure 4.28, which was shown in Chapter 5. Based on the optimal Huff-and-Puff GAGD process (Chapter 7), the optimal set of injection, soaking, and production periods along with the number of cycles and minimum bottom hole pressure in production wells were set as nominal values in the geological uncertainty assessment. Since there are nine quantiles from permeability and  $K_v/K_h$ , the total number of simulations in the geological uncertainty was 81 ( $9 \times 9 = 81$ ). The workflow of the geological uncertainty quantification is outlined below (CMG, 2011):

1. Create discrete realizations of static reservoir properties that represent model uncertainty.
2. Create a parameter that includes these realizations as candidate values in the assessment model.
3. Generate multiple simulation jobs honoring the uncertain space realizations.
4. Run the reservoir simulation jobs, including all the realizations.
5. Evaluate and analyze the simulation outcome, then select the three quantiles of the model outcome.

Flowchart of geological uncertainty is illustrated in Figure 8.2.

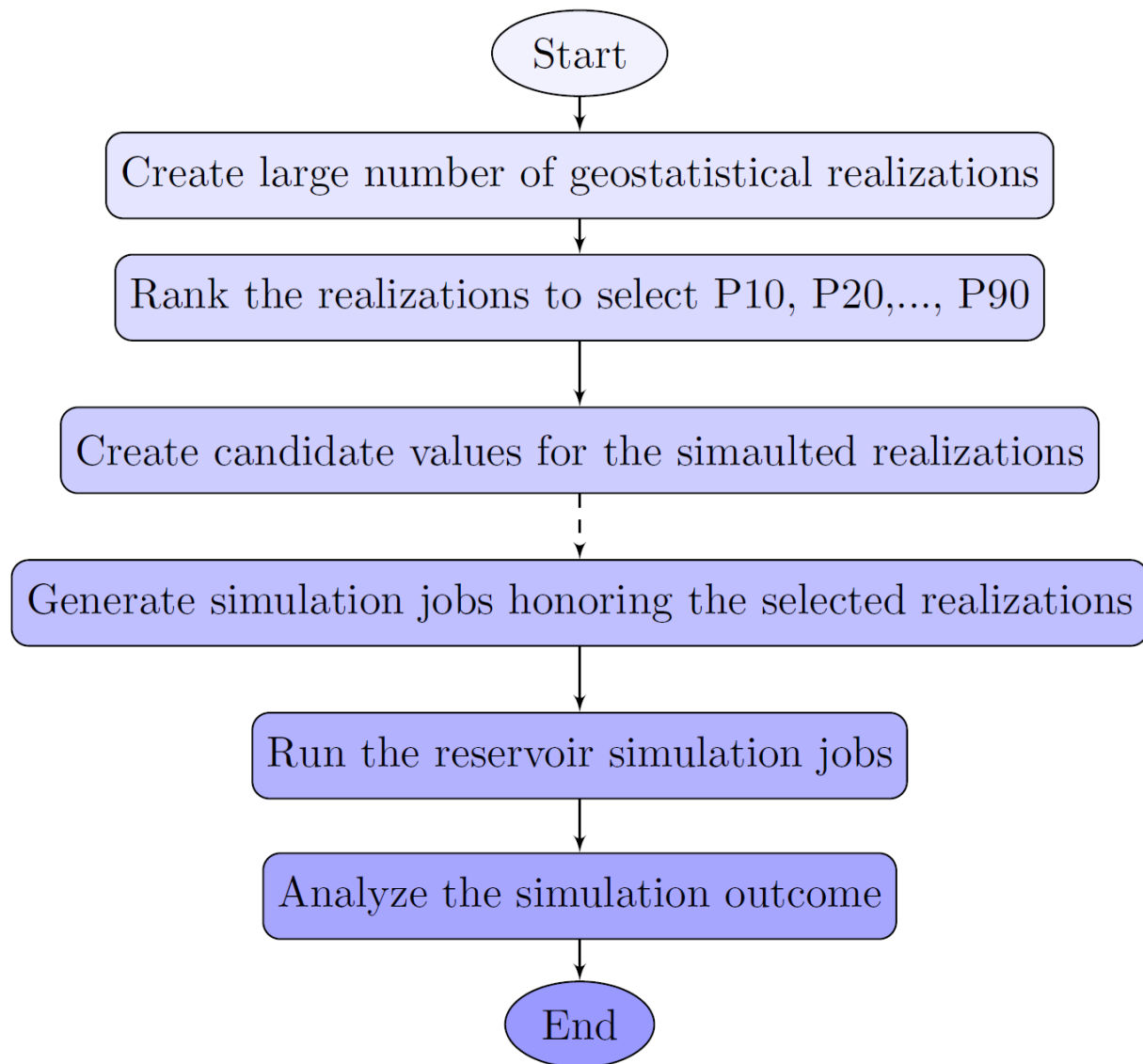


Figure 8.2: Flow Chart of Huff and Puff Geological uncertainty

### 8.3 Production Uncertainty

After conducting the geological uncertainty quantification and obtaining the most-likely reservoir model that has the actual cumulative oil production (objective function) given the optimal set of permeability and anisotropy realizations, production uncertainty was conducted to assess the uncertainty in the independent uncertain operational design parameters of the Huff-and-Puff GAGD process. The independent uncertain operational



design parameters of the Huff-and-Puff GAGD process are injection, soaking, and production periods in addition to the number of cycles and minimum bottom hole pressure in the producers. The optimal set of these parameters was obtained from the optimal Huff-and-Puff GAGD process, which was addressed in the last section of Chapter 7. These optimal parameters might have some uncertain values that may lead to obtaining higher objective function. The uncertain values may result from the significant difference between each two levels in the suggested values for each parameter in the optimization process.

Consequently, assessing the uncertainty in the varied Huff-and-Puff parameter levels were conducted through the Design of Experiments, proxy modeling, and Monte-Carlo simulation. First, the Box-Behnken design was employed to generate the training and verification simulation jobs by combining the levels of the four most influential Huff-and-Puff parameters. These four parameters are periods of injection, soaking, and production in addition to the minimum bottom hole pressure in the production wells. In order to capture the overall uncertainty space with less number of observation, three levels of each parameter (min, median, and max) were sampled based on uniform, triangle, and truncated normal distributions. These distribution types were adopted to generate normal or semi-normal sampled data behavior as the all the original data, used to build the geological model, was transformed to the normal distribution. Therefore, the total number of simulation jobs sampled by Box-Behnken design that capture a wide variety of the uncertain parameters is twenty five. Next, the ordinary kriging algorithm, based on the reservoir outcome from the training simulation jobs, built a proxy model in this case. Then, the proxy model was validated by the verification simulation jobs. Finally, based on the verified proxy model, 65,000 simulation jobs were created by the Monte-Carlo simulation to quantify and analyze the uncertainty in the reservoir cumulative oil production. The complete flow chart of Huff-and-Puff uncertainty assessment procedure is illustrated in Figure 8.3.

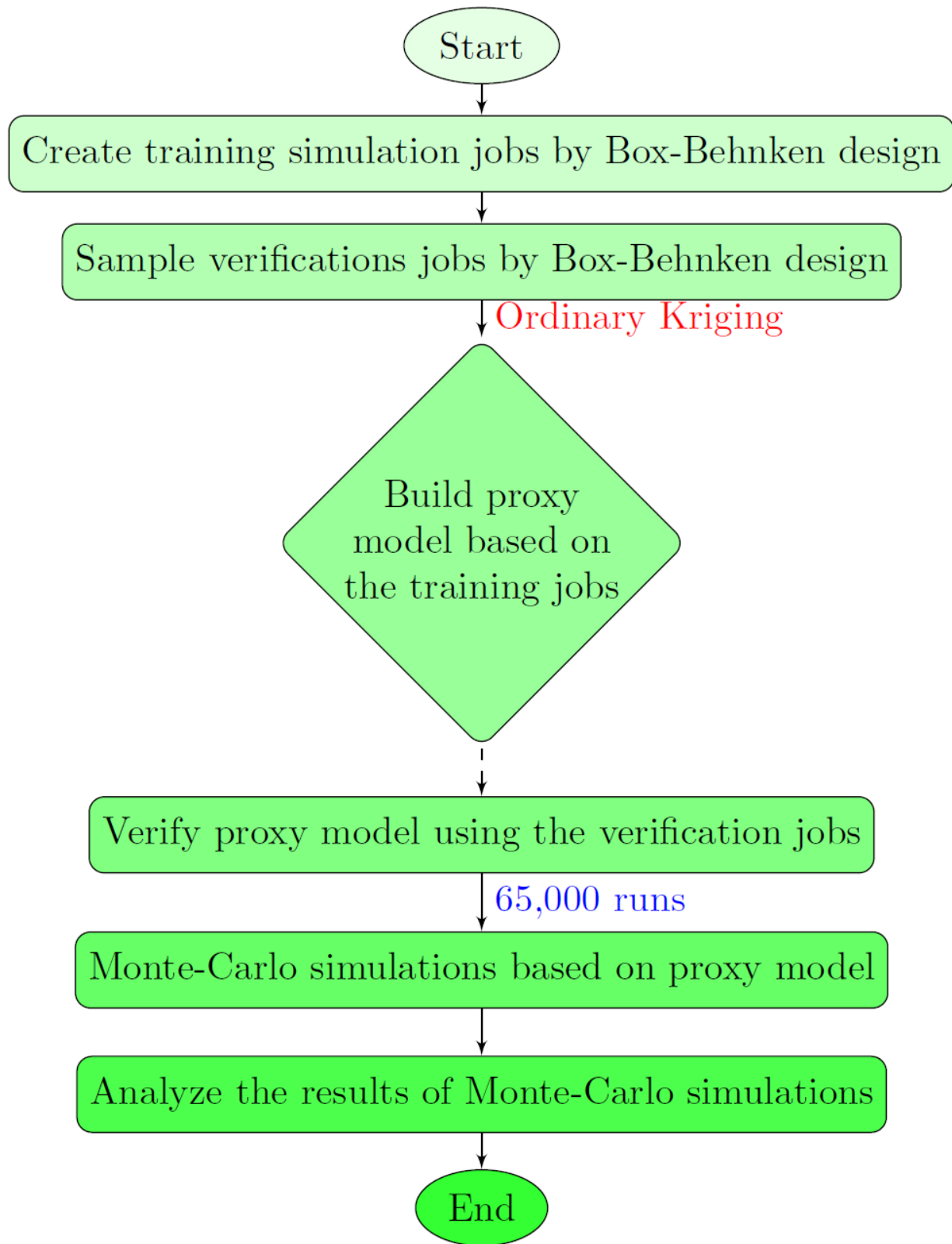


Figure 8.3: Flowchart of Production Huff and Puff Uncertainty Assessment

## 8.4 Heterogeneity and Anisotropy Effects

Reservoir heterogeneity and anisotropy have a direct impact on the fluid movement through the porous media. More specifically, the oil production through the EOR processes is controlled by the reservoir complexity in terms of the spatial variation of the reservoir permeability as well as the ratio of vertical to horizontal permeability. Thus, the reservoir heterogeneity and anisotropy are quantified through the generation of multiple stochastic reservoir images (realizations) to be integrated into the reservoir flow simulation. These realizations capture a wide variety of the property ranges, especially when the geostatistical model includes multiple lithology distributions.

## 8.5 Results and Discussion

### 8.5.1 Geological Uncertainty Quantification

In the geological uncertainty assessment, 81 simulation runs were conducted considering the combined realizations of permeability and anisotropy ratio. Figure 8.4 and Figure 8.5 illustrate the outcome of integrating the geological uncertainties in the GAGD reservoir simulation with regards to cumulative oil production and field oil rate, respectively. The figures show how the geological uncertainty assessment affected the reservoir flow response.

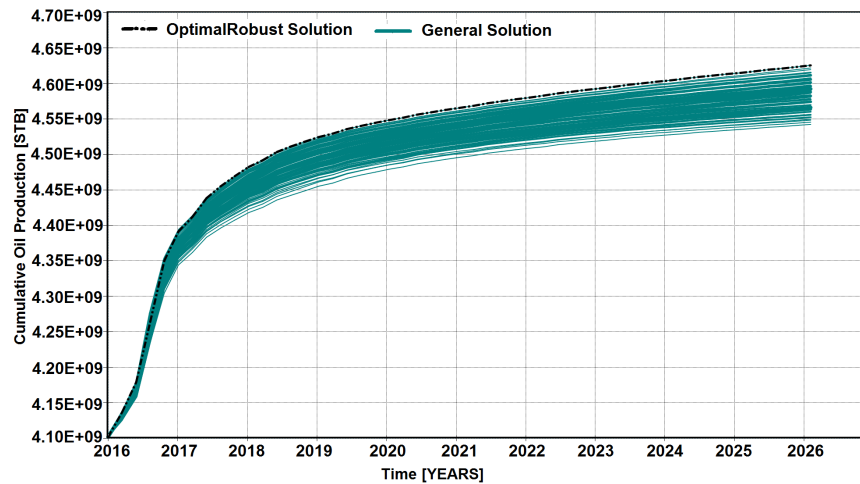


Figure 8.4: Effect of Geological Uncertainty on Field Cumulative Oil Production

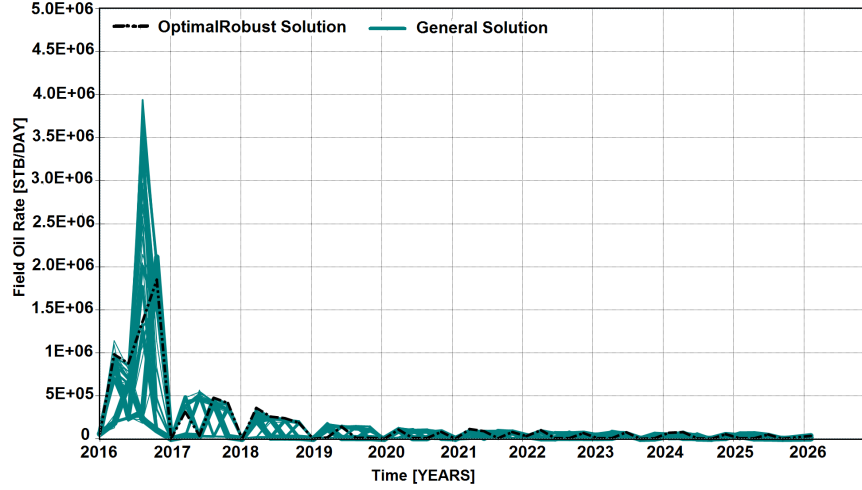


Figure 8.5: Effect of Geological Uncertainty on Field Oil Production Rate

Figure 8.6 shows the histogram of the quantified uncertainty effect on the reservoir flow response. The figure also illustrates the less-likely (P10), median (P50), and most-likely (P90) cumulative oil production. Furthermore, Figure 8.7 depicts the cumulative probability distribution function in addition to the less-likely, median, and most-likely cumulative oil production values. The wide variety range of the calculated reservoir flow response reflects the importance of quantifying the geological uncertainty in the GAGD reservoir simulation.

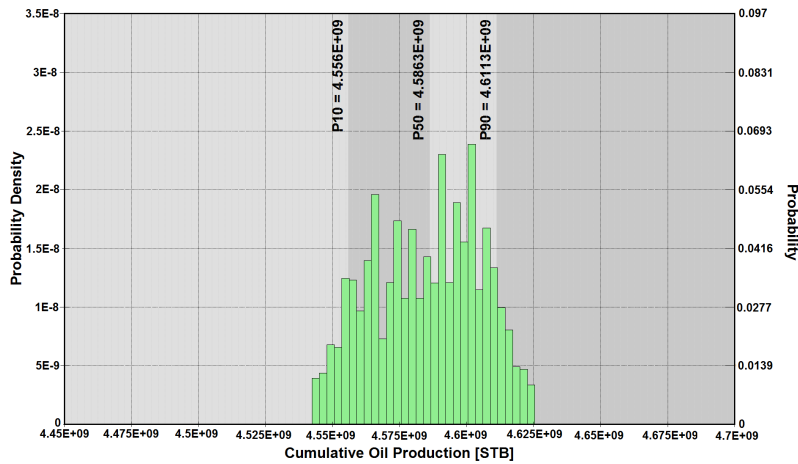


Figure 8.6: The Geological Uncertainty Quantification of Cumulative Oil production

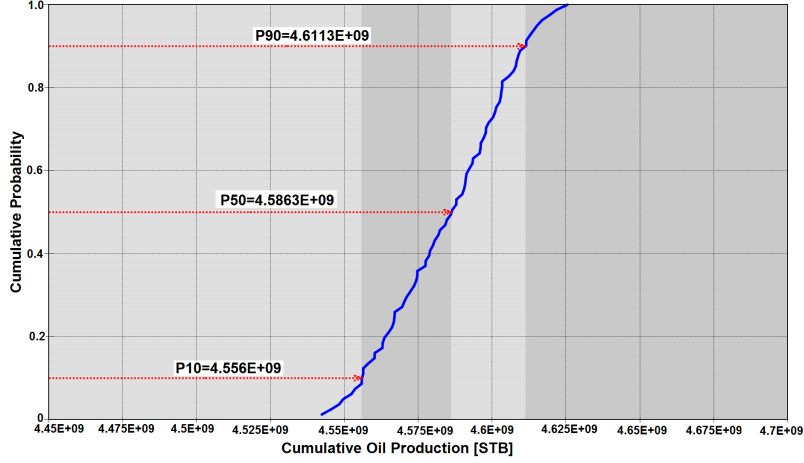


Figure 8.7: Cumulative Distribution Function of Cumulative Oil production

Form the previous two figures, the geological uncertainty has no effect on the reservoir flow response as the optimal robust solution has the highest amount of cumulative oil production (4.62547 million bbls). The input quantiles for the geological parameters that led to generate the less-likely, median, and most-likely outcomes are outlined in Table 8.1. The P80 level of reservoir permeability and the P70 level of the anisotropy ratio were both combined to obtain the most-likely reservoir flow response.

Table 8.1: Quantiles of Parameters in the Geological Uncertainty Quantification

Model	Permeability	Kv/Kh	CumOilProd [STB]
Less-likely	P30	P20	4.556E+09
Median	P40	P50	4.5863E+09
Most-likely	P80	P70	4.6113E+09

### 8.5.2 Production Uncertainty Quantification

The optimal robust solution was likewise integrated into the uncertainty assessment of the operational decision factors. The uncertain factors are minimum bottom hole pressure in production wells in addition to the injection, soaking, and production periods. Figure 8.8 and Figure 8.9 illustrate the reservoir flow response with respect to cumulative oil production and field flow rate for all the simulation runs conducted using proxy-based Monte-Carlo simulation, respectively.

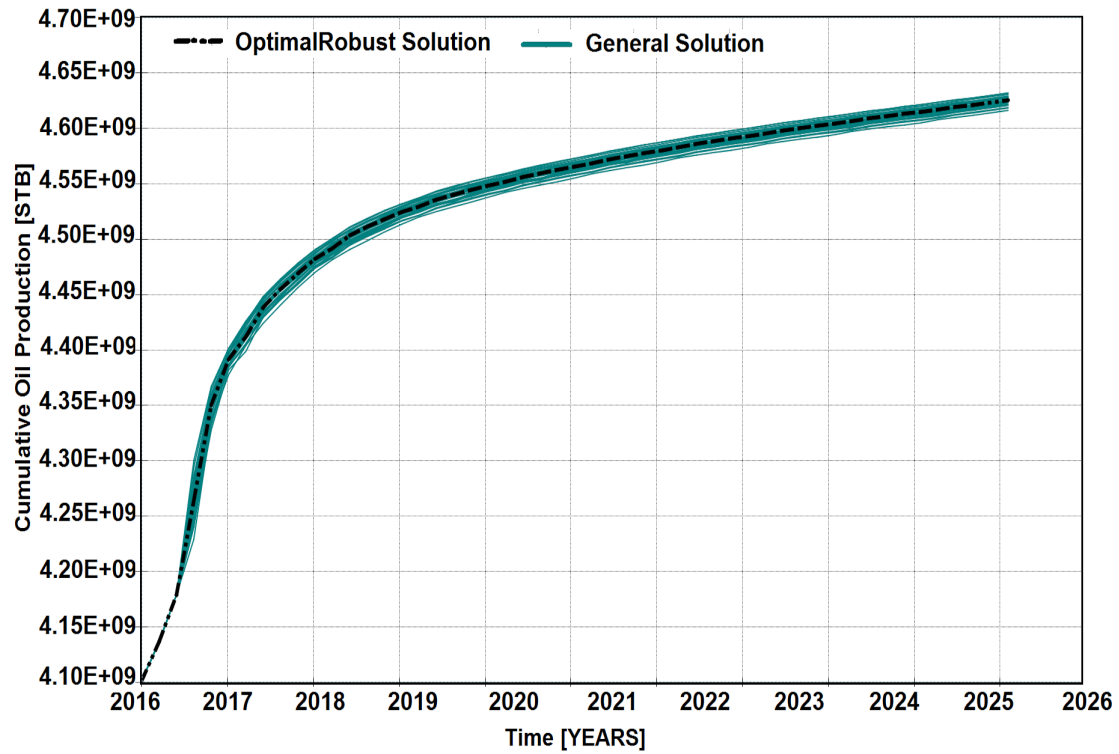


Figure 8.8: Reservoir Flow Responses Under Production Huff and Puff uncertainty

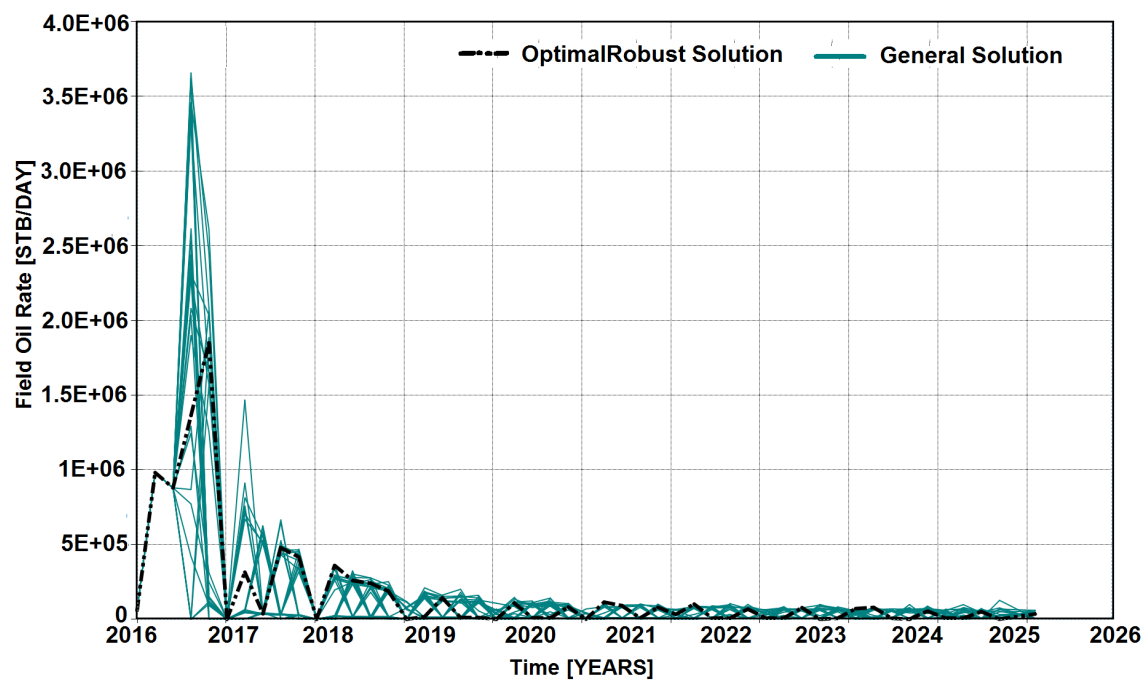


Figure 8.9: Reservoir Flow Responses Under Production Huff and Puff uncertainty

There was a little variability in the reservoir flow response among the base case, the optimistic model of geological uncertainty, and other simulated models. The difference between the resulting less-likely and most-likely objective functions is only 7.5 million barrels of oil, as illustrated in Figure 8.10 and Figure 8.11. The limited uncertainty for some production control variables led to slight increment in oil recovery, which is equal to 6.53 million barrels of oil over the optimal robust solution. The maximum cumulative oil production obtained by the production uncertainty assessment is 4.632 million bbls, whether the optimal robust solution was 4.62547 million bbls.

The P10, P50, and P90 of the four uncertain Huff-and-Puff parameters resulted in

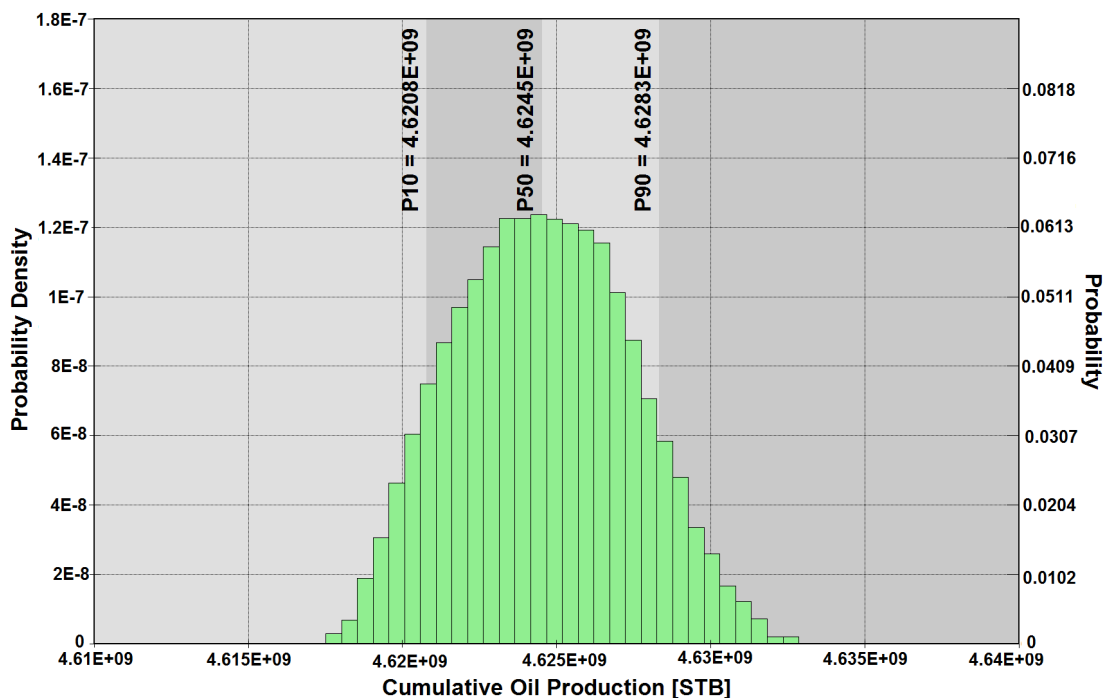


Figure 8.10: Cumulative Probability Distribution of Cumulative Oil Production under Huff and Puff Uncertainty

obtaining the less-likely, median, and most-likely objective functions in different impact levels, respectively. Figures 8.12, 8.13, 8.14, and 8.15 depict the probability distribution of the objective function (cumulative oil production) for the four uncertain parameters given the three model outcomes: less-likely, median, and most-likely. In these three figures, there is slight significant gap between the resulted reservoir flow response given the differ-

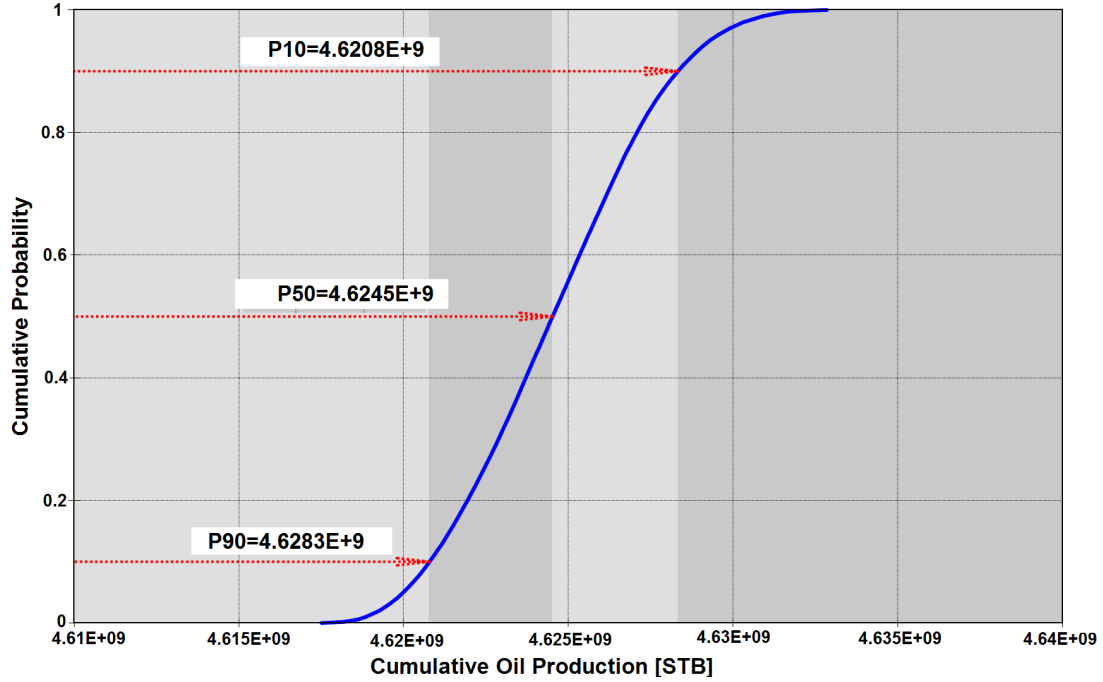
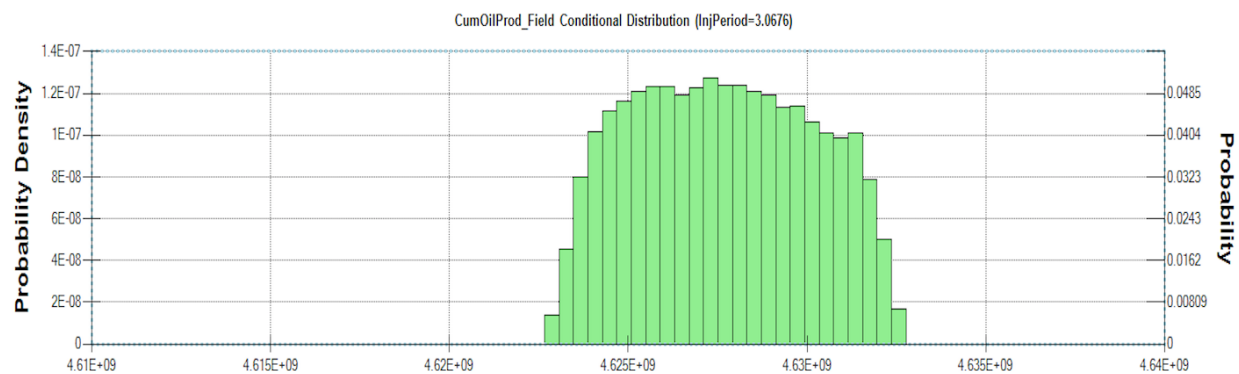


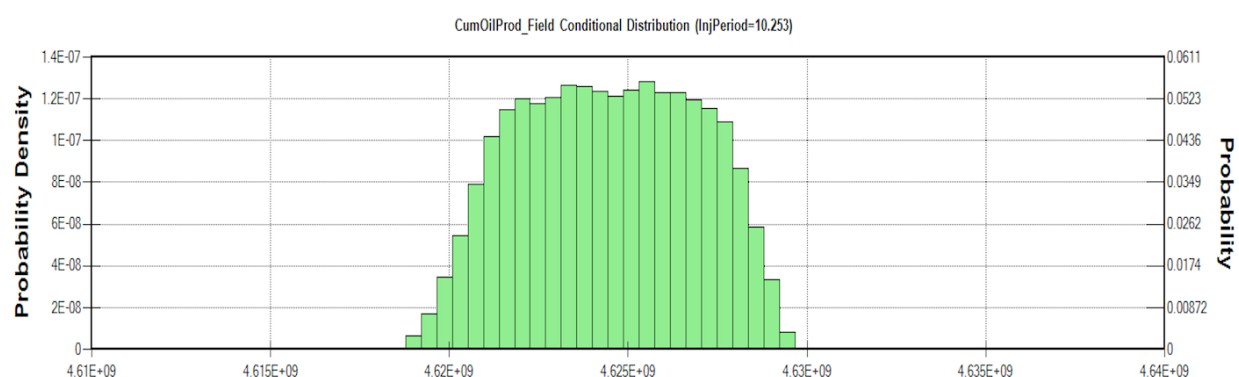
Figure 8.11: Cumulative Probability Distribution of Cumulative Oil Production under Huff and Puff Uncertainty

ent realizations of the injection, soaking, and production periods, as well as the minimum bottom hole pressure. More specifically, all the optimization processes of the Huff-and-Puff GAGD simulation resulted in impeding the optimal range the reservoir flow response as a function of a small uncertain range for the decision parameters. From the cumulative probability distributions for the four parameters, it can be inferred that the injection periods and minimum BHP are more sensitive parameters than soaking and production periods.

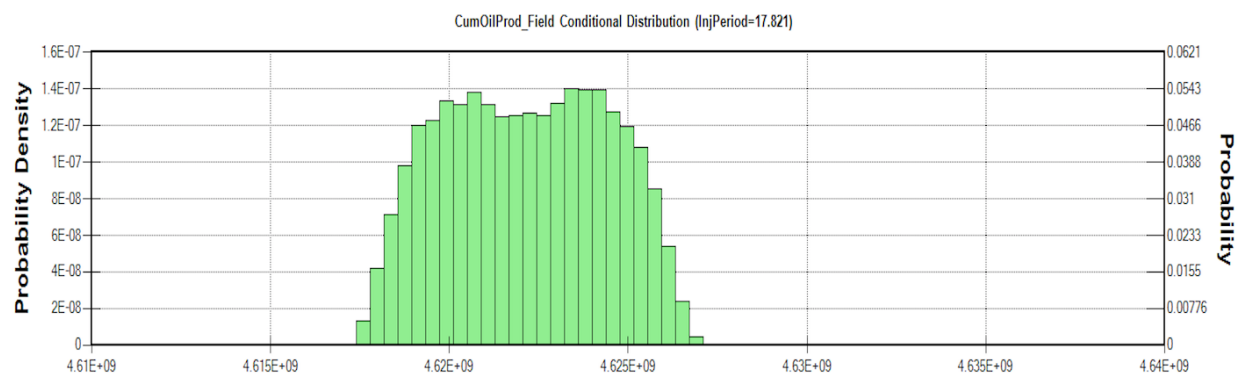




**Cumulative Oil Production [STB]**



**Cumulative Oil Production [STB]**



**Cumulative Oil Production [STB]**

Figure 8.12: Probability Distribution of Flow Response as a Function of Injection Periods

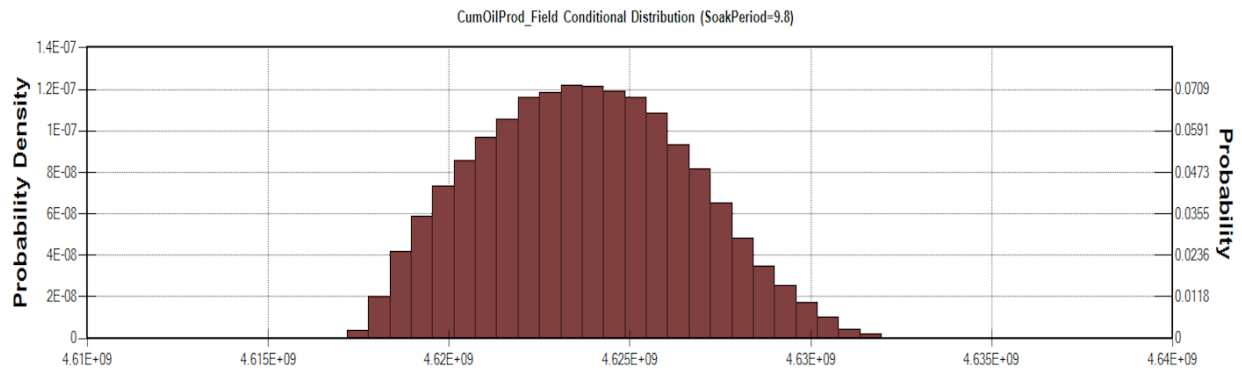
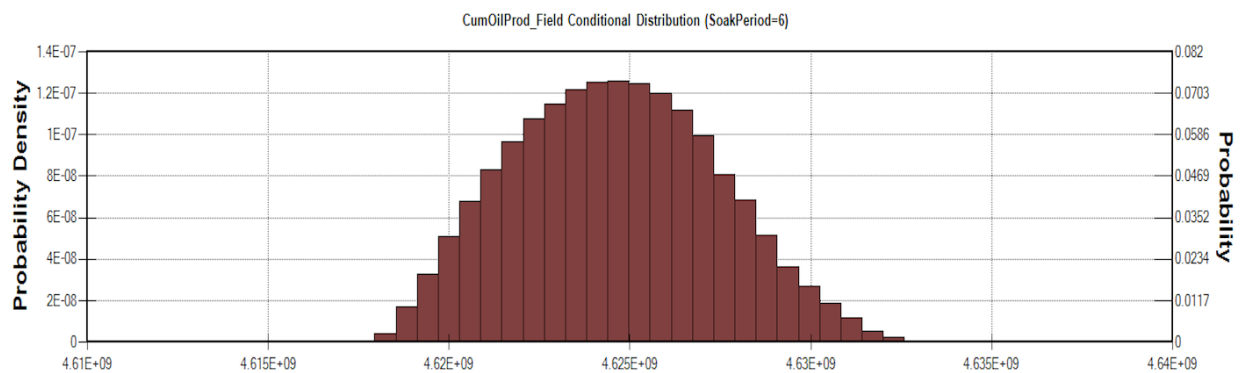
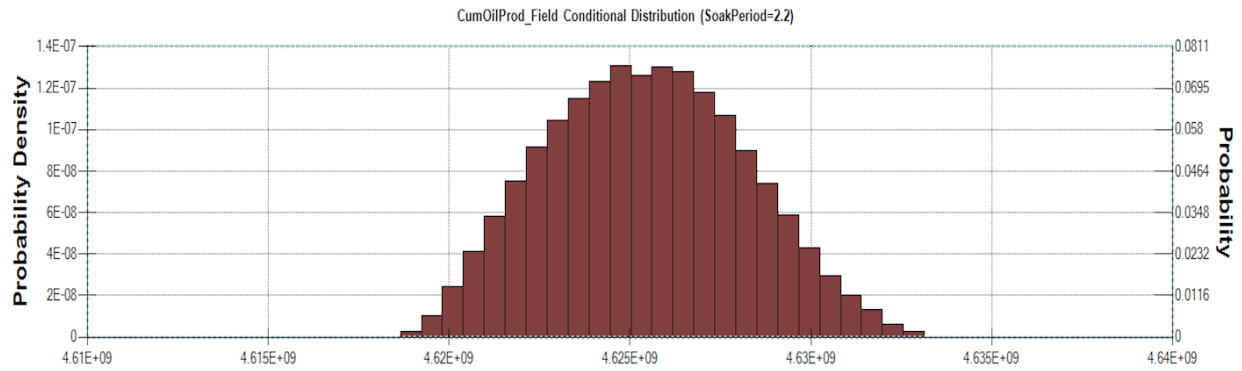
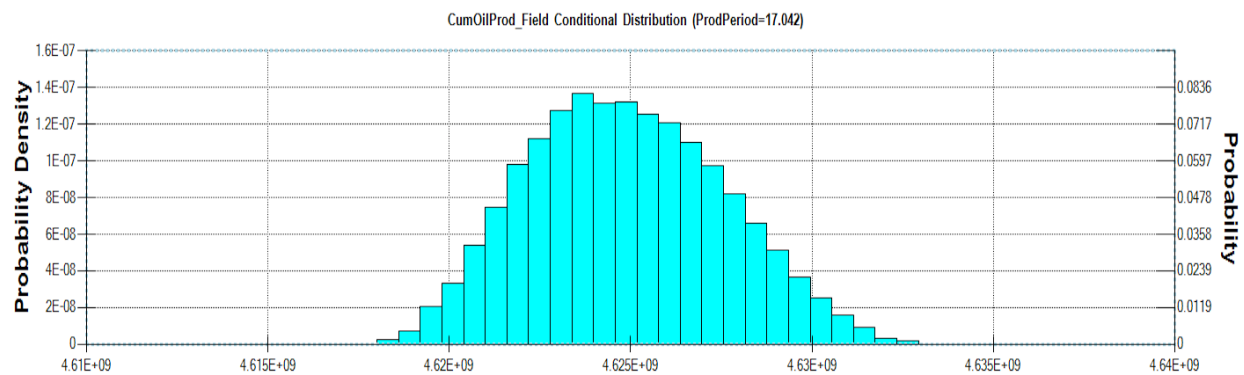
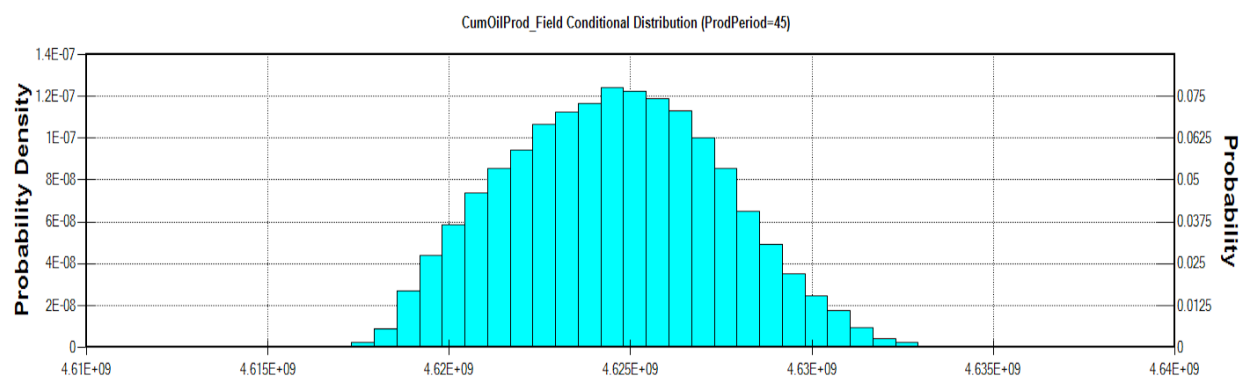


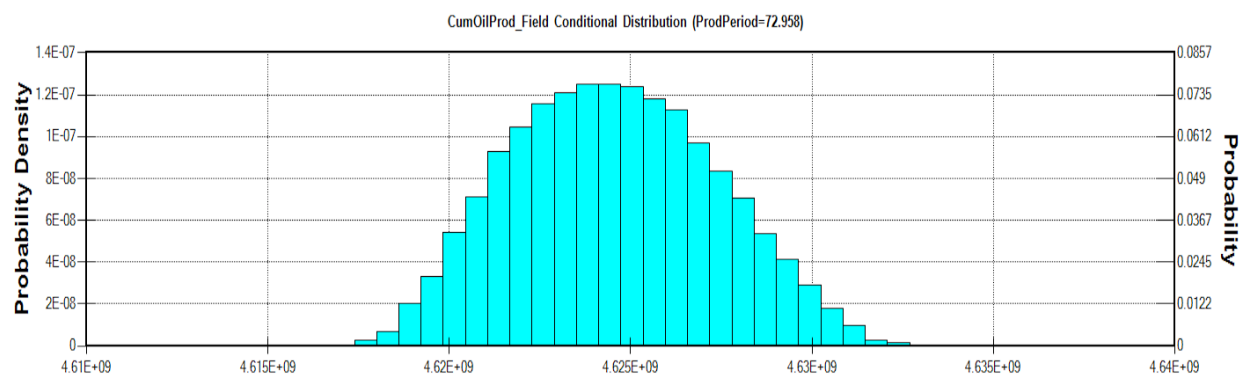
Figure 8.13: Probability Distribution of Flow Response as a Function of Soaking Periods



**Cumulative Oil Production [STB]**



**Cumulative Oil Production [STB]**



**Cumulative Oil Production [STB]**

Figure 8.14: Probability Distribution of Flow Response as a Function of Production Periods

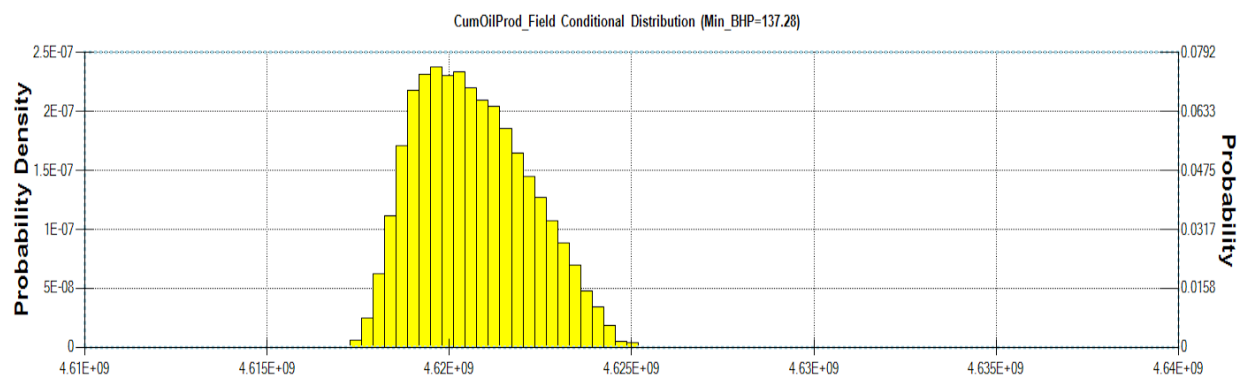
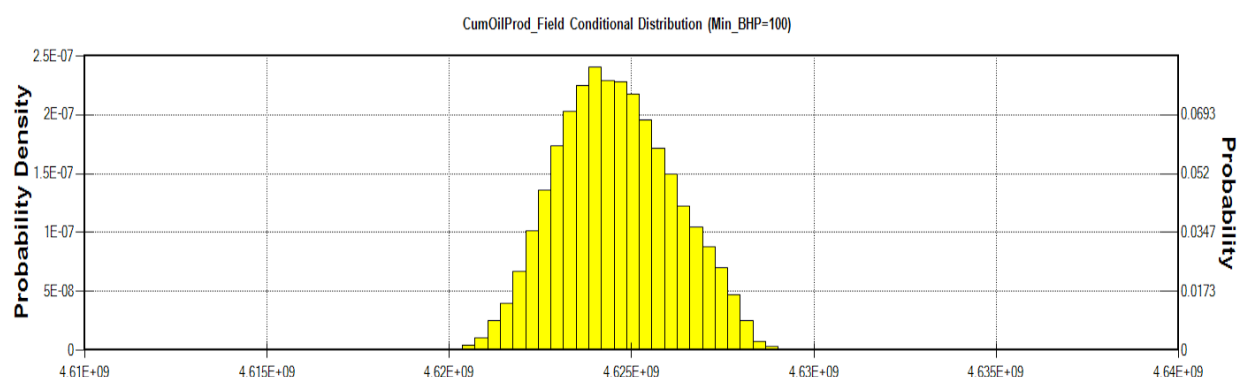
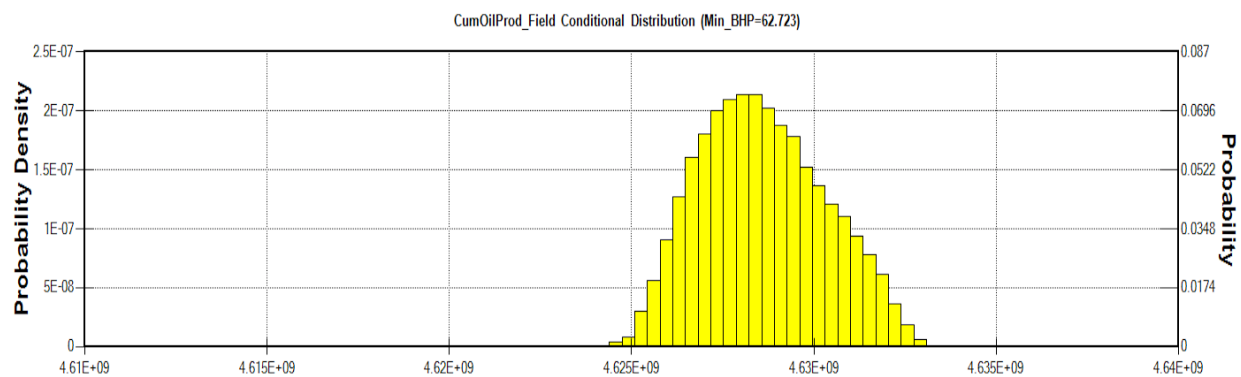


Figure 8.15: Probability Distribution of Flow Response as a Function of Minimum BHP

### 8.5.3 Heterogeneity and Anisotropy Effects

The same nine equiprobable stochastic images of spatial permeability variation, which were created in Chapter 5, were re-embedded to study the effect of geological heterogeneity on the Gas Assisted Gravity Drainage process performance. The geological heterogeneity was quantified by independently simulating the P10, P20, ..., P90 of the 3D spatial permeability variation in the compositional reservoir simulation to obtain the related cumulative oil production, field oil rate, and water cut.

Figures 8.16, 8.17, and 8.18 show the effect of geostatistical realizations on the GAGD process performance.

The permeability heterogeneity was noticed to have a significant impact on the GAGD process, especially in terms of cumulative oil production and field oil rate with limited effect on the water cut. Little variability in the produced water between the different models was obtained because water drains towards the bottom of payzone according to the gravity segregation.

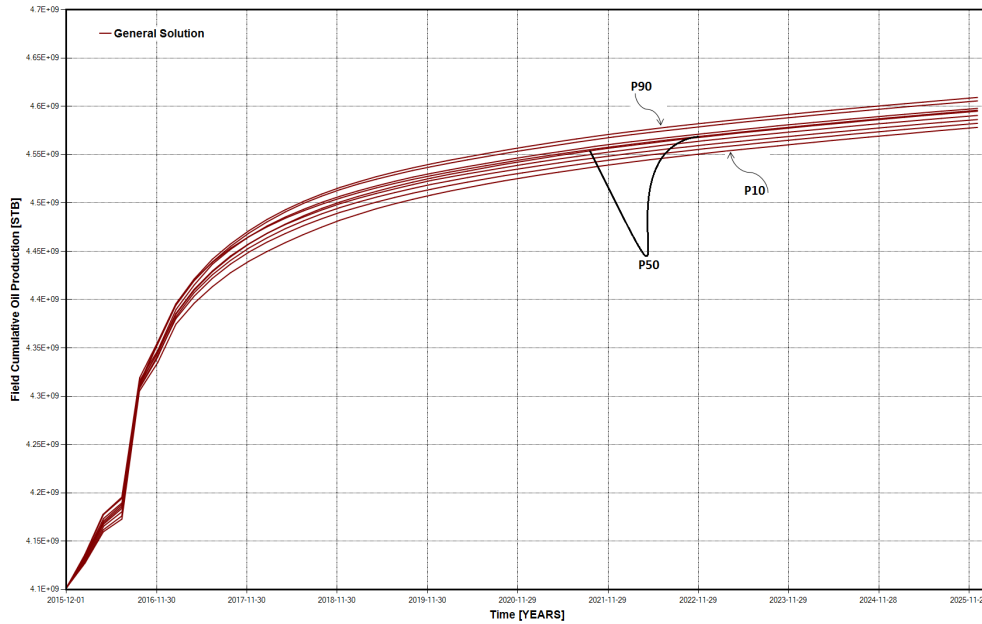


Figure 8.16: Effect of Permeability Variation on Field Cumulative Oil Production

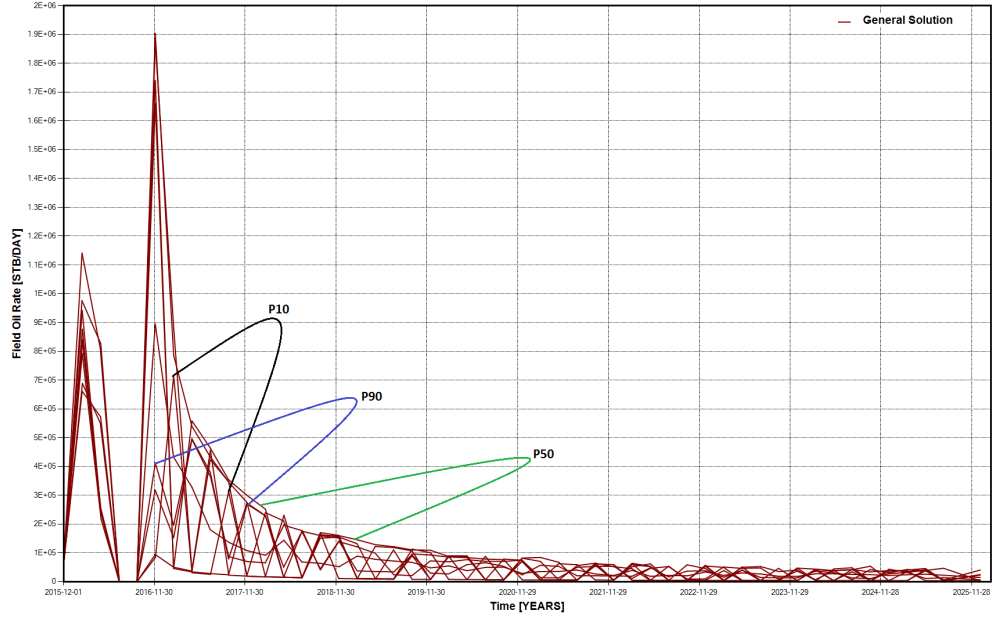


Figure 8.17: Effect of Permeability Variation on Field Oil Production Rate

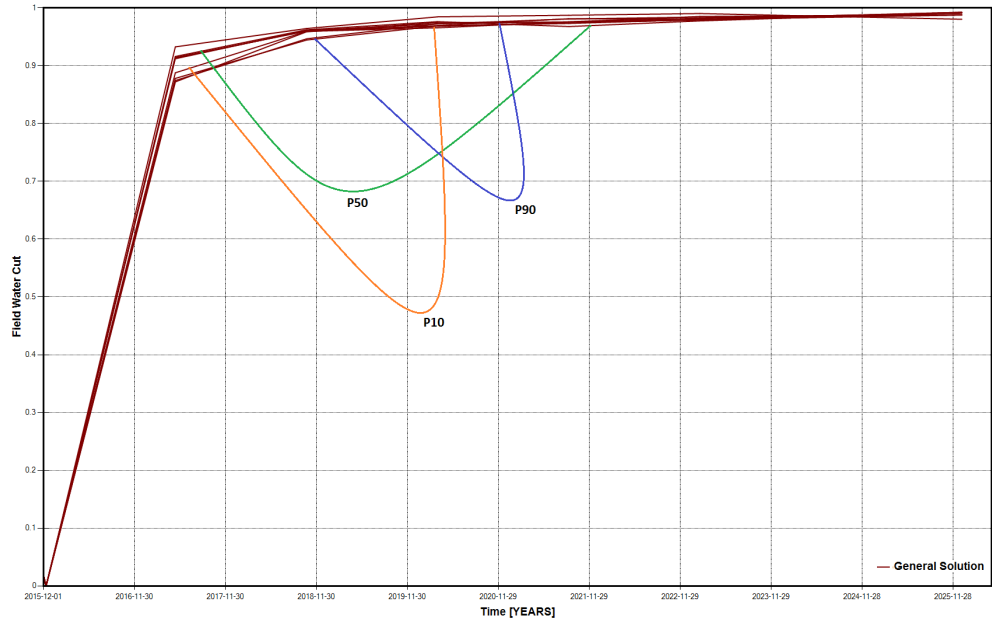


Figure 8.18: Effect of Permeability Variation on Field Water Cut

The 3D vertical permeability distribution was also adopted with nine anisotropy ratio levels to be independently simulated in the reservoir model in order to study the effect of permeability anisotropy on the GAGD process. Figures 8.19, 8.20, and 8.21 illustrate

the effect of permeability anisotropy on the GAGD process performance with regard to cumulative oil production, monthly oil rate, and field water cut. There is a significant impact of the permeability anisotropy on the GAGD process performance with respect to oil rate and cumulative production because of the main concept of GAGD process that considers vertical fluid movements from the vertical injection wells to the horizontal producers. Moreover, the high water cut values come from the infinite active edge water aquifer. Nevertheless, there is no significant difference between the produced water in different  $(K_v/K_h)$  values as the water is located at the bottom of payzone due to the gravity drainage.

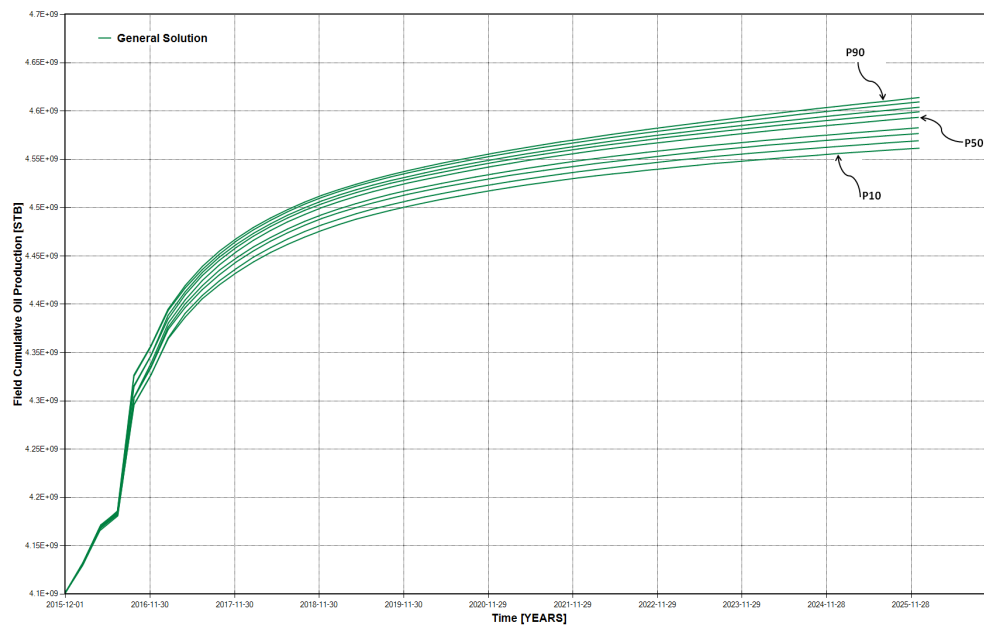


Figure 8.19: Effect of Permeability Anisotropy on Field Cumulative Oil Production

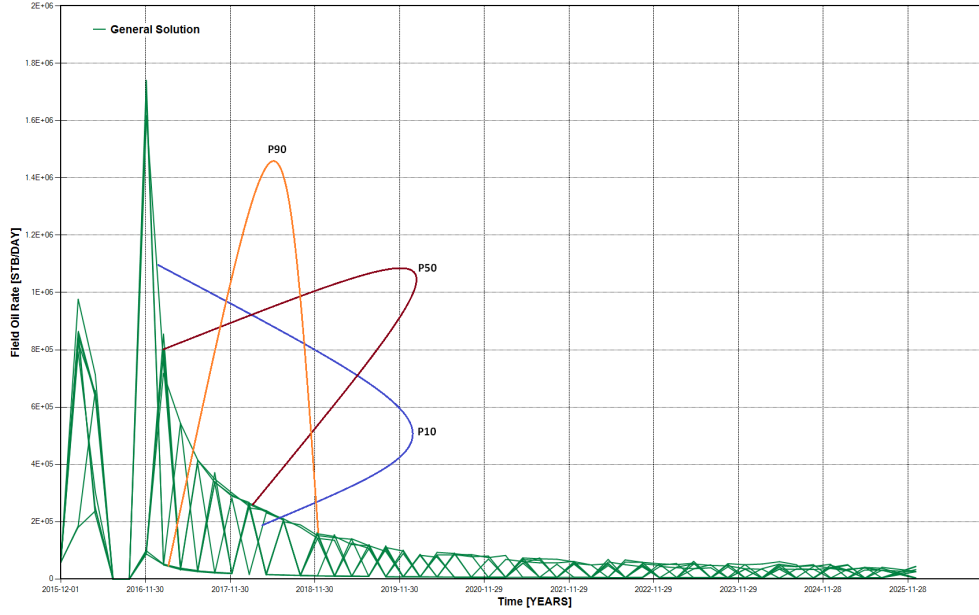


Figure 8.20: Effect of Permeability Anisotropy on Field Oil Production Rate

## 8.6 Summary and Significance

For uncertainty assessment, two different successive uncertainty quantifications were conducted based on the robust optimal case of the Huff-and-Puff GAGD process. The uncertainty was first quantified in terms of geological parameters. Nine realizations of permeability and anisotropy ratio were incorporated into the geological uncertainty assessment. Since the optimal robust solution included the optimal levels of permeability and anisotropy ratio, the geological uncertainty assessment has never led to obtain higher oil production than the optimal robust solution. Therefore, the optimal robust solution was then adopted for extra uncertainty with an optimization glance that includes further optimization of Huff-and-Puff operation parameters. These parameters are minimum bottom hole pressure in the injection wells in addition to the periods of injection, soaking, and production. The uncertainty results showed that there is limited uncertain space for the soaking and production periods. Nevertheless, there is wider uncertain space for the injection periods and minimum bottom hole pressure. As a result, the final optimal case of the Huff-and-Puff GAGD simulation led to increase the cumulative oil production about



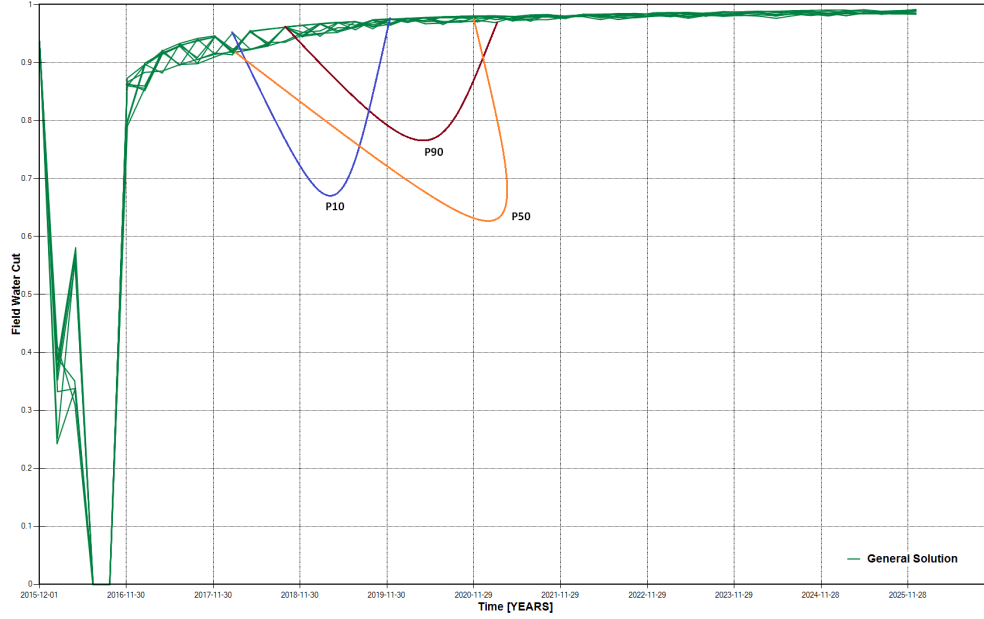


Figure 8.21: Effect of Permeability Anisotropy on Field Water Cut

only 6.53 million bbls larger than the optimal robust solution. Finally, incorporating the production uncertainties effects in the GAGD optimization processes is an efficient way to obtain the true optimal solution that provides a significant increment in oil recovery.

The final optimal case of the Huff-and-Puff GAGD simulation was considered for heterogeneity and anisotropy effects through independently integrating multiple reservoir realizations in the reservoir flow simulation. The effect of these realizations illustrates that the anisotropy ratio has higher impact on the reservoir flow response through the GAGD process than on the spatial permeability variation.

## 9. Conclusions and Recommendations

### 9.1 Conclusions

The reservoir uncertainties are very crucial elements to achieve reliable predictions and accomplish discreet decisions through the investments regarding future reservoir performance. Therefore, a full geostatistical reservoir characterization model for Lithofacies and petrophysical properties was conducted to capture the most realistic geological structure. Then, the geological model is adopted to build an accurate compositional reservoir simulation model in order to precisely evaluate the CO<sub>2</sub> flooding performance through the Gas Assisted Gravity Drainage (GAGD) process. Prior to performing the uncertainty assessment, the optimization of the GAGD process was carried out through several different approaches of reservoir flow response optimization under geological uncertainty.

The main findings, which were obtained from the entire research in this dissertation, are outlined below along with the dissertation work scope:

#### 9.1.1 Applied Multivariate Geostatistics

1. Different algorithms were adopted for the multivariate lithofacies classification and prediction. Kernel Support Vector Machine and Probabilistic Neural Networks were the most accurate lithofacies modeling because they provided nonlinear classification between the different lithofacies. More specifically, the two algorithms led to obtain the highest total percent correct of lithofacies classification predicted. However, the total percent correct of lithofacies classification predicted by linear discriminant

analysis was the lowest value as it produced a straight separating line between the nonlinear behavior system (lithofacies).

2. Smoothed Generalized Additive Model (sGAM) algorithm was adopted for petrophysical property modeling and prediction at missing value intervals and other wells in comparison to the generalized linear model (GLM). Smooth GAM led to obtain more accurate modeling than the GLM in terms of a higher coefficient of multiple determination ( $R^2$ ) and a lower root mean square prediction error ( $RMSPE$ ).
3. All the multivariate lithofacies and permeability modelings were conducted based the cross-validation assessment techniques.

### 9.1.2 Geostatistical Reservoir Modeling

1. The predicted distributions of discrete lithofacies and petrophysical properties were incorporated into building the 3D geostatistical reservoir model. In particular, the sequential indicator simulation (SISIM) and multiple-point geostatistics (MPS) were used for 3D lithofacies modeling of the main pay/ South Rumaila oil field. MPS algorithm led to obtain the most realistic fluvial system that reflected the fluvial depositional environment of the main pay/ Rumaila field, which has been described in the literature. Moreover, the multiple-point geostatistics algorithm was capable of capturing the real depositional environments as it adopts a training image, which integrates a synthetic lithofacies pattern into the facies modeling. On the other hand, the simulated lithofacies model through the SISIM captured the fluvial depositional environment less accurately than the multiple-point geostatistics because SISIM considers the variogram analysis, which, in some cases, has the same curves for different facies structures.
2. Given the MPS lithofacies model, the sequential Gaussian simulation (SGSIM) was adopted to build a 3D petrophysical property distribution. The petrophysical modeling was implemented given each lithotypes in order to capture the most realistic

geological heterogeneity as each lithotype has distinct property ranges and distributions.

3. The Leave-One-Out cross-valuation was adopted to examine the validity of the petrophysical properties modeling through the sequential Gaussian simulation. Based on the cross-plots, there were perfect matches between the predicted complete and the one well-missing datasets for all the available-data wells. The validated geostatistical reservoir model provided a solid basis for accurately implementation and evaluation of the performance of the GAGD process in the reservoir simulation. Finally, multiple reservoir stochastic images (realizations) were created and ranked to capture a geological uncertainty space.

### **9.1.3 Integrated Reservoir Modeling**

1. The fine-scale geostatistical lithofacies and petrophysical property model was up-scaled and incorporated into the compositional reservoir model for the GAGD process simulation. The upscaling model had an excellent matching with the fine-scale one with regard to histograms of the petrophysical and lithofacies distributions.
2. The coarse-scale model was adopted in the compositional reservoir model to simulate the Gas Assisted Gravity Drainage (GAGD) process in the main pay of South Rumaila oil field. A notable history matching was obtained in terms of the oil rates, cumulative production, water injection rate, and cumulative injection for the entire field and all the wells, which regularly distributed across the entire reservoir.
3. Given the fact that locations of production and injection wells serve as constraints for the lithofacies and petrophysical modeling, the perfect history matching provided a decent indication about the accuracy of the geostatistical model to capture the true fluvial lithofacies description.
4. In the 10 year prediction evaluation of the GAGD reservoir simulation, 20 vertical injectors were installed at the top two layers of the reservoir for CO<sub>2</sub> injection along

with a series of horizontal oil producers at the middle zones. The cumulative oil production at the beginning of the prediction period is 4.1 billion bbls. The cumulative oil production reaches to 4.255 billions bbls through the primary production at the end of the prediction period.

5. However, the GAGD process prediction resulted in obtaining approximately 4.4 billion bbls. Specifically, approximately a 150 million bbls was obtained from the base GAGD case over the primary case. The recovery factor in the base case, given the remaining oil is then 15%. Additionally, the obtained amount of oil in 10 years primary production can be obtained in only one year by the GAGD base case.
6. Two special cases were again implemented with a longer prediction period than the GAGD base case in order to show the effectiveness of the GAGD process to reach promising levels of oil recovery. These special cases represent the implementation of the GAGD process in immiscible and miscible for 25 prediction years. The recovery factor through immiscible and miscible special cases reach to 30% and 42%, correspondingly. More specifically, the cumulative oil production reaches to 4.7 billion bbls and 4.93 billion bbls in immiscible and miscible injection modes, respectively.
7. There is an absence of the gas breakthrough in the immiscible GAGD process implementation as the injected does not reach the production zone and accumulates at the top of the reservoir. On the other hand, there is a negligible amount of produced gas during the first five years of production period (forecasting period) because of the released gas through the vertical production tubing when the well flow pressure become less than the bubble-point pressure. Continuing gas injection leads to increase the reservoir pressure to reach more than 4500 psi that keeps the well flow pressure above bubble-point pressure and eliminates gas releasing and production.

All the studied three cases of the GAGD process evaluation discloses how the GAGD process is significant to increase oil recovery in real field-scale application.

#### 9.1.4 Sensitivity Analysis of the GAGD Process

1. In sensitivity analysis, the permeability and anisotropy ratio ( $Kv/Kh$ ) of the entire reservoir were determined as the most prominent reservoir parameters that impact the GAGD process performance. There was no effect for the reservoir porosity on the GAGD process because the porosity controls the fluid storage not fluid flow in the reservoir.
2. A detailed approach of sensitivity analysis was carried out with regard to determining the most influential petrophysical parameters in terms of the different layers, not the entire reservoir. The permeability of injection, below-production, and transition layers were found to be more influential in the GAGD process than the anisotropy ratio. However, anisotropy ratio is only influential in the production and below-production layers. The permeability and anisotropy ratio at the bottom (below-production) layers have significant effects because they have been fully flooded with the aquifer water, which supports the reservoir pressure. Furthermore, the horizontal permeability was more influential on the GAGD process performance at the top injection layers comparing with any other layer as the shale modeling impeded the vertical gas and oil flow based on the concept of GAGD process.
3. Bayesian Model Averaging (BMA) and stepwise elimination were adopted to eliminate the non-influential factors on the reservoir flow response through the GAGD process. The BMA was considered as a stochastic approach for the layer-based property parameter sensitivity analysis. Among 25 models, the BMA model was selected for the highest posterior probability that includes the most influential reservoir parameters with more accurate effect analysis, unlike the stepwise elimination that produces only one model. The stochastic model was essential for nonlinear physical system behavior that would have been inefficient to apply one solution of a linear regression.

4. The influential strength of various parameters were recognized by visualizing the effect of different factors on the field cumulative oil production. The validity and accuracy of the sensitivity analysis models were further checked through the non-parametric bootstrapping technique that computes the estimated ( $R^2$ ), confidence intervals, normal Q-Q plots, and standard error in the regression analysis.

#### **9.1.5 GAGD Process Optimization**

1. Oil recovery optimization of the GAGD process was conducted in three various ways through the Design of Experiments and proxy modeling approaches. The optimization approaches adopted the to generate different simulation jobs in order to reach the optimal solution.
2. In the first optimal case (OptimalCase1), the operational decision factors were sampled to generate training simulation runs to determine the optimal cumulative oil production through the immiscible CO2 flooding by the end of prediction period. The parameters are CO2 injection rate and maximum BHP in the injection wells, along with the maximum oil production rate, minimum BHP, skin factor, and maximum water cut in the horizontal producers. The OptimalCase1 led to obtain 4.6039 million bbls with increment of 212.5 million bbls of oil over the base GAGD case (360 Million bbls over the primary production case).
3. In OptimalCase1, the maximum bottom hole pressure in the injection wells (Max'BHP) was detected as the dominant factor that influences the GAGD process performance. Therefore, it was reconsidered for the OptimalCase2, to be annually changing in immiscible and miscible injection modes. The obtained cumulative oil recovery in this case approached 4.838 million bbls with 226.4 million bbls of oil higher than the OptimalCase1 and with 438.9 million bbls of oil over the GAGD base case. However, the miscible injection mode of GAGD process required high ranges of CO2 injection pressure, which could be economically impractical in the field.

4. Simulation and optimization Huff-and-Puff GAGD processes were conducted in an immiscible CO<sub>2</sub> flooding mode to test the effectiveness of cycling injection/production operations in the GAGD process. The base case of the Huff-and-Puff GAGD process simulation led to trigger the cumulative oil production to be 142.71 million bbls larger than the 1st base case of the GAGD process.
5. The first optimal Huff-and-Puff GAGD process was accomplished based on a nominal reservoir model (single realization). The optimization procedure included sampling the Huff-and-Puff parameters of the injection, soaking, production periods, number of cycles, as well as the minimum bottom hole pressure in the producers. In nominal optimization, the increment in oil recovery was 62.26 million bbls larger than the Huff-and-Puff base case.
6. Further, the less-likely, median, and most-likely of the ranked realizations were incorporated in robust Huff-and-Puff optimization of the GAGD process. The optimal solution increased the cumulative oil production by 28.7 million bbls of oil larger than the nominal optimization case.
7. In particular, the periodic injection, soaking, and production activities led to trigger the oil production during the prediction period. Additionally, incorporating the geological uncertainties in the immiscible Huff-and-Puff GAGD process resulted in obtaining higher oil recovery than the optimal case of immiscible production control optimization case (OptimalCase1). More specifically, the cumulative oil production of 4.62547 million bbls predicted from the Huff-and- Puff optimization process was higher than the 4.6039 million bbls obtained from the production control optimization case.
8. The DECE optimization algorithm was conducted in Huff-and-Puff GAGD simulation, as it was faster to reach the optimal solution than latin hypercube plus proxy



model. DECE algorithm removed the weak levels that negatively affected the objective function values. Therefore, it avoided from trapping in local optima.

#### **9.1.6 Uncertainty Assessments of the GAGD Process**

1. For further uncertainty assessment, two different successive uncertainty quantification approaches were achieved based on the robust optimal case of the Huff-and-Puff GAGD process. Firstly, the uncertainty was quantified in terms of geological parameters. Nine realizations of permeability and anisotropy ratio were incorporated into the reservoir simulator to assess the geological uncertainty. The most-likely flow response had the same cumulative oil production of the optimal robust solution.
2. The previous optimal robust solution was again adopted for another uncertainty assessment with an optimization glance of the Huff-and-Puff process parameters. The parameters were minimum bottom hole pressure in the production wells, as well as the periods of injection, soaking, and production. The uncertainty results showed that there was limited uncertain space for the soaking, and production periods, unlike the wider uncertain space for the injection periods and minimum bottom hole pressure in the production wells. The production uncertainties led to reduce the risk in decision making through all the future reservoir performance scenarios.
3. The final optimal case of Huff-and-Puff GAGD simulation was considered for heterogeneity and anisotropy effect through independently integrating multiple reservoir realizations in the reservoir flow simulation. The anisotropy ratio has a higher effect on the reservoir flow response through than the spatial permeability variation because of the vertical-based fluid flow in the GAGD process.

## 9.2 Recommendations

Although an extensive work was performed in this dissertation on the GAGD process evaluation, many other research trends are recommended to conduct, as outlined below:

1. More static reservoir parameters could be incorporated in a sensitivity analysis to determine the most influential constraint factors of the GAGD process.
2. The reservoir fluid parameters such as PVT data, relative permeability, and capillary pressure, etc, could also be combined in the sensitivity analysis of the GAGD process.
3. In this research, only the proxy-based optimizations approaches were used at production control and injection pressure optimization cases. It would be also to consider the proxy model for Huff-and-Puff optimization under a geological uncertainty. However, such research would be time consuming with much computation.
4. Building a reliable economic analysis model by modern softwares could provide more efficient optimization of GAGD process in terms of calculating the Net Present Value (NPV).
5. For a fast implement the geostatistical characterization modeling, next step would be conducting bayesian kriging algorithm for petrophysical property distribution to be incorporated for uncertainty assessment of GAGD process performance.
6. The GAGD reservoir simulation in this dissertation was conducted based on the multiple-point geostatistics, which is more realistic than other geostatistical algorithms. Therefore, GAGD would be implemented based on the sequential indicator simulation for lithofacies modeling to provide a comparison between the two algorithm in terms of history matching and an uncertainty quantification.
7. An optimization of injection/production well locations would be attained to determine the optimal future reservoir performance through the GAGD process.

# References

- [1] Alabert, F. G. and G. J. Massonnat. (1990). Heterogeneity in a Complex Turbiditic Reservoir: Stochastic Modelling of Facies and Petrophysical Variability. SPE-20604-MS paper presented at the SPE Annual Technical Conference and Exhibition, (23-26 September), New Orleans, Louisiana.
- [2] Alabert, F.G. and V. Modot. (1992). Stochastic Models of Reservoir Heterogeneity: Impact on Connectivity and Average Permeabilities. SPE-24893-MS paper presented at the SPE Annual Technical Conference and Exhibition, (4-7 October), Washington, D.C.
- [3] Al-Ameri, T. K., A. J. Al-Khafaji, and J. Zumberge. (2009). Petroleum system analysis of the Mishrif reservoir in the Ratawi, Zubair, North and South Rumaila oil fields, southern Iraq. *GeoArabia*: 91-108.
- [4] Al-Ansari, R.. (1993). The petroleum Geology of the Upper sandstone Member of the Zubair Formation in the Rumaila South. Geological Study, Ministry of Oil, Baghdad, Iraq.
- [5] Al-Mudhafar, W. and Y Gebrai. (2014). Using Multiple Linear Models for Permeability Estimation and Modeling in a Well of Sandstone Reservoir. An Extended Abstract presented at the 76th EAGE Conference and Exhibition, (16-19 June), Amsterdam, The Netherlands.
- [6] Al-Mudhafar, W. J. and W. da Silva. (2015). OTC-26143-MS paper presneted at the OTC Brasil, (27-29 October), Rio de Janeiro, Brazil.
- [7] Al-Mudhafar, W. and A. Al Isawi. (2015). Multivariate Multiple Linear Regression for Jointly Modeling of Petrophysical Properties. Extended Abstract presented at the 77th EAGE Conference and Exhibition Incorporating SPE EUROPEC, (1-4 June), Madrid, Spain.
- [8] Al-Mudhafar, W. J. and A. K. Al-Khazraj. (2016). Non-Parametric Adaptive Regression Splines for Multisource Permeability Modeling in a Sandstone Oil Reservoir. OTC-26431-MS presented at the Offshore Technology Conference Asia, (22-25 March), Kuala Lumpur, Malaysia.

- [9] Al-Mudhafar, W. (2014). Using Generalized Linear Regression Of Multiple Attributes For Modeling And Prediction The Formation Permeability In Sandstone Reservoir. OTC-25158-MS presented at the Offshore Technology Conference, (4-8 May), Houston, Texas, USA.
- [10] Al-Mudhafar, W. and A. Rostami. (2014). Comparative Applied Multivariate Geostatistical Algorithms For Formation Permeability Modeling. JFES-1401, Extended abstract presented at Japan Formation Evaluation Symposium-SPWLA, (1-2 October), Tokyo, Japan.
- [11] Al-Mudhafar, W. (2015). Integrating Bayesian Model Averaging for Uncertainty Reduction in Permeability Modeling. OTC-25646-MS presented at the Offshore Technology Conference, (4-7 May), Houston, Texas, USA.
- [12] Al-Mudhafar, W. J. and D. N. Rao. (2015). Optimization of Gas Assisted Gravity Drainage (GAGD) Process in a Heterogeneous Sandstone Reservoir: Field-Scale Study. SPE-174579-MS presented at the SPE Enhanced Oil Recovery Conference, (11-13 August), Kuala Lumpur, Malaysia.
- [13] Al-Mudhafar, W. and M. Bondarenko. (2015). Integrating K-means Clustering Analysis and Generalized Additive Model for Efficient Reservoir Characterization. Extended abstract presented at the 77th EAGE Conference and Exhibition, (1-4 June), Madrid, Spain.
- [14] Al-Mudhafar, W. (2014). Integrating Markov Chains for Bayesian Estimation of Vertical Facies Sequences through Linear Discriminant Analysis. Extended abstract presented at the 76th EAGE Conference and Exhibition, (16-19 June), Amsterdam, The Netherlands.
- [15] Al-Mudhafar, W. (2015). Integrating Component Analysis and Classification Techniques for Comparative Prediction of Continuous and Discrete Lithofacies Distributions. OTC-25806-MS presented at the Offshore Technology Conference, (4-7 May), Houston, Texas, USA.
- [16] Al-Mudhafar, W. (2016). Applied Geostatistical Reservoir Characterization in R: Review and Implementation of Rock Facies Classification and Prediction Algorithms-Part I. OTC-26947-MS presented at the Offshore Technology Conference, (2-5 May), Houston, Texas, USA.
- [17] Al-Mudhafar, W. and L. Mohamed. (2015). Incorporating Lithofacies Classification and well logs into Statistical Learning Algorithms for Comparative Multisource Permeability Modelling. SPE-175776-MS paper presented at the SPE North Africa Technical Conference and Exhibition, (14-16 September), Cairo, Egypt.

- [18] Al-Mudhafer, W. J. (2013). A Practical Economic Optimization Approach with Reservoir Flow Simulation for Infill Drilling in A Mature Oil Field. SPE-164612-MS paper SPE 130054 presented at the SPE North Africa Technical Conference and Exhibition, (15-17) April, Cairo, Egypt.
- [19] Al-Mudhafer, W. J., M. S. Al Jawad, and D. A. Al-Shamaa. (2010). Using Optimization Techniques for Determining Optimal Locations of Additional Oil Wells in South Rumaila Oil Field,” paper SPE 130054 presented at the CPS/SPE International Oil and Gas Conference and Exhibition, (8–10 June), Beijing, China.
- [20] Al-Muhailan, M., I. Hussain, H. Maliekkal, O. Ghoneim, P. Nair, and M. Fayed. (2013). New HTHP Cutter Technology Coupled with FEA-Based Bit Selection System Improves ROP by 60% in Abrasive Zubair Formation. IPTC-17122, paper presented at the International Petroleum Technology Conference, (26–28 March), Beijing, China.
- [21] Al Naqib, K.M. (1967) Geology of the Arabian Peninsula Southwestern Iraq. Paper 560-G, U.S. Geological Survey Professional. United States Government Printing Office, Washington.
- [22] Amudo, C., T. Graf, R. R. Dandekar, and J. M. Randle. (2009). The Pains and Gains of Experimental Design and Response Surface Applications in Reservoir Simulation Studies. SPE-118709-MS paper presented at the SPE Reservoir Simulation Symposium, (2-4 February), Houston, Texas.
- [23] Arpat, G. B. and J. Caers. (2005). A multiple-scale, pattern-based approach to sequential simulation. in Geostatistics Banff 2004, edited by O. Leuathong, and C. V. Deutsch, pp. 255–264, Springer, Dordrecht, Netherlands.
- [24] Ates, H., A. Bahar, S. El-Abd, M. Kelkar, and A. Datta-Gupta. (2003). Ranking and Upscaling of Geostatistical Reservoir Models Using Streamline Simulation: A Field Case Study. SPE-81497-MS paper presented at the Middle East Oil Show, (9-12 June), Bahrain.
- [25] Al-Obaidi, R.Y. (2009). Identification of Palynozones and Age Evaluation of Zubair Formation, Southern Iraq. Journal of Al-Nahrain University, 12(3): 16-22.
- [26] Al-Obaidi, R.Y. (2010). Determination of Palynofades to Assess Depositional Environments and Hydrocarbons Potential, Lower Cretaceous, Zubair Formation South Iraq. Journal of College of Education 5: 163-174.
- [27] Aziz, K. and A. Settari. (1979). *Petroleum reservoir simulation*. Chapman and Hall.
- [28] Ballin, P. R., K. Aziz, A.G. Journel, and L. Zuccolo. (1993). Quantifying the impact of geologic uncertainty on reservoir performance forecasts. SPE-25238-MS paper presented at the SPE Symposium on Reservoir Simulation, (28 February-3 March), New Orleans, Louisiana.

- [29] Begg, S. H., A. Kay, E. R. Gustason, and P. F. Angert. (1996). Characterization of a Complex Fluvial-Deltaic Reservoir for Simulation. SPE Formation Evaluation. 11(03), 47-154.
- [30] Behrens, R. A., M. K. MacLeod, T. T. Tran, and A. C. Alimi. (1998). Incorporating seismic attribute maps in 3D reservoir models. SPE Reservoir Evaluation and Engineering. 1(02), 122-126.
- [31] Ben-Hur, A., and J. Weston. (2010). *A user's guide to support vector machines*. In Data mining techniques for the life sciences. Humana Press.
- [32] Bhat, C. R. (2001). Quasi-Random Maximum Simulated Likelihood Estimation of the Mixed Multinomial Logit Model. Transportation Research Part B: Methodological, (35)7, 677-693.
- [33] Box, George E. P. and D. R. Cox. (1964). An analysis of transformations. Journal of the Royal Statistical Society, Series B 26 (2): 211-252. JSTOR 2984418. MR 192611.
- [34] Box, G.E., J.S. Hunter, and W.G Hunter. (2005). *Statistics for Experimenters: Design, Innovation, and Discovery*. 2nd Edition. John Wiley and Sons, INC.
- [35] Brockett, P. L., S. L. Chuang, and U. Pitaktong. (2014). Generalized Additive Models and Nonparametric Regression. Predictive Modeling Applications in Actuarial Science: Volume 1, Predictive Modeling Techniques, 367.
- [36] Brown, P. J. and T. F. Vannucci. (2002). Bayes model averaging with selection of regressors. Journal of the Royal Statistical Society (Series B: Statistical Methodology), 64(3), 519 - 536.
- [37] Burges, C. J. C. (1998). A Tutorial on Support Vector Machines for Pattern Recognition. Submitted to Data Mining and Knowledge Discovery, <http://svm.research.belllabs.com/SVMdoc.html>.
- [38] Caers, J. (2000). Direct sequential indicator simulation. In the 6th International Geostatistics Congress, South Africa, April 10-14, 2000. 12p.
- [39] Caers, J. (2005). *Petroleum Geostatistics*. Society of Petroleum Engineers, Richardson.
- [40] Caers, J. and T. Zhang. (2004). Multiple-point geostatistics: a quantitative vehicle for integrating geologic analogs into multiple reservoir models. AAPG Mem 80:383-394.
- [41] Cai, Z. and P.J. Hicks Jr. (1999). 3D Conditional Simulation of Porosity For a Heterogeneous Core. Journal of Canadian Petroleum Technology, 38(01).
- [42] Carnell, R. (2015). Package 'lhs'. Available at: <https://cran.r-project.org/web/packages/lhs/lhs.pdf>.

- [43] Carpenter, C. (2013). Produced-Water-Reinjection Design and Uncertainties Assessment. *Journal of Petroleum Technology*, 12(65):129-131. doi:10.2118/1213-0129-JPT.
- [44] Chasset, Pierre-Olivier. (2015). Package 'pnn'. <https://cran.r-project.org/web/packages/pnn/pnn.pdf>, retrieved Feb. 2015.
- [45] Cheung, V. and K. Cannons. (2003). An Introduction to Probabilistic Neural Networks. Online. <http://www.psi.toronto.edu/~vincent/research/presentations/PNN.pdf>.
- [46] CMG. (2011). Manual of Computer Modelling Group's Software.. Computer Modelling Group, Calgary, Canada.
- [47] Cruz, P.S. (2000). *Reservoir Management Decision-Making In The Presence Of Geological Uncertainty*. Stanford, USA. 239p. PhD Dissertation, Department of Petroleum Engineering, Stanford University.
- [48] Cui, H., A. Stein, and D.E. Myers. (1995). Extension of spatial information, Bayesian kriging and updating of prior variogram parameters. *Environmetrics*4,373-384.
- [49] Damsleth, E. (1994). Mixed Reservoir Characterization Methods. SPE-27969-MS paper presented at the University of Tulsa Centennial Petroleum Engineering Symposium, (29-31 August), Tulsa, Oklahoma.
- [50] Damsleth, E., B. T. Charlotte, H. Omre, and H. H. Haldorsen. (1992). *Journal of Petroleum Technology*, 44(04), 402-486.
- [51] David, M. (1986). Using Kriging To Map The Ocean Floor. OTC-5274-MS paper presented at the Offshore Technology Conference, (5 May), Houston, Texas.
- [52] Denney, D. (2010). Pros and Cons of Applying a Proxy Model as a Substitute for Full Reservoir Simulations. Society of Petroleum Engineers. doi:10.2118/0710-0041-JPT.
- [53] Deutsch, C. V. and A. G. Journel. (1998). *GSLIB. Geostatistical Software Library and User's Guide*. New York.
- [54] Deutsch, C.V. and L. Wang. (1996). Hierarchical object-based stochastic modeling of fluvial reservoir. *Mathematical Geology*, v. 28, no.7, p. 857-880.
- [55] Deutsch, C. V. and P. W. Cockerham. (1994). Practical considerations in the application of simulated annealing to stochastic simulation. *Mathematical Geology*. 26(1), 67-82.
- [56] De Vries, L., J. Carrera, O. Falivene, O. Gratacos, and L. Slooten. (2009). Application of multiple point geostatistics to non-stationary images. *Math. Geosci.*, 41(1), 29-42.
- [57] Diggle, J. Peter and P. J. Ribeiro Jr. (2007). *Model-based Geostatistics*. New York: Springer.

- [58] Diggle, Peter J. and Lophaven, Soren. (2004). Bayesian Geostatistical Design. Johns Hopkins University, Dept. of Biostatistics Working Papers. Working Paper 42. <http://biostats.bepress.com/jhubiostat/paper42>.
- [59] Diggle, P. and S. Lophaven. (2006). Bayesian Geostatistical Design. Scandinavian Journal of Statistics 33,53-64.
- [60] Ding, Y. and P. Lemonnier. (1995). Use of Corner Point Geometry in Reservoir Simulation. International Meeting on Petroleum Engineering, Beijing, China, 14-17 November.
- [61] Dobson, A. J. (1990). *An Introduction to Generalized Linear Models*. Chapman and Hall, London.
- [62] Doyen, P.M., L.D. Den Boer, and W.R. Pillet. (1996). Seismic Porosity Mapping in the Ekofisk Field Using a New Form of Collocated Cokriging. SPE-36498-MS paper presented at the SPE Annual Technical Conference and Exhibition, (6-9 October), Denver, Colorado.
- [63] Efron, B. (2004). Large-Scale Simultaneous Hypothesis Testing. Journal of the American Statistical Association 99 (465): 96.
- [64] Fanchi, J.R. (1997). *Principles of Applied Reservoir Simulation*. Houston, Tex, Gulf Pub.
- [65] Fedutenko, E., C. Yang, Card, C., and L. Nghiem. (2013). Optimization of SGAD Process Accounting for Geological Uncertainties Using Proxy Models. Proceedings of CSPG/CSEG/CWLS GeoConvention. (6-12 May), Calgary, AB, Canada.
- [66] Fisher, R. (1936). The Use of Multiple Measurements in Taxonomic Problems. Annals of Eugenics 7 (2): 179–188. doi:10.1111/j.1469-1809.1936.tb02137.x.
- [67] Friedman, J., T. Hastie, and R. Tibshirani.(2000). Special invited paper. Additive logistic regression: A statistical view of boosting. Annals of statistics, 337-374.
- [68] Gomez, J. and R. Srivastava. (1990). ISIM3D: an ANSI-C three dimensional multiple indicator conditional simulation: Computer and Geosciences, 16, p. 395–410.
- [69] Gringarten, E. and C. V. Deutsch. (1999). Methodology for variogram interpretation and modeling for improved reservoir characterization. In SPE annual technical conference (pp. 355-367).
- [70] Goovearts, P. (1997). *Geostatistics for Natural Resources Evaluation*. Applied Geostatistics Series. 483 pp.
- [71] Guisan, A., T. C. Edwards Jr, and T. Hastie.(2002). Generalized linear and generalized additive models in studies of species distributions: setting the scene. Ecological modeling, 157, no.2, 89-100.



- [72] Guyaguler, B., and R. N. Horne. (2001). Uncertainty Assessment of Well Placement Optimization. SPE-71625-MS presented at the SPE Annual Technical Conference and Exhibition, (30 September-3 October), New Orleans, Louisiana, USA. doi:10.2118/71625-MS.
- [73] Guyon and A. Elisseeff. (2003). An introduction to variable and feature selection. *Journal of Machine Learning Research*, 3: 1157–1182.
- [74] Haldosen, H. H., and D. W. Chang. (1986). Notes on stochastic shales; from outcrop to simulation model, in L. W. Lake, and H. B. Carroll, eds., *Reservoir characterization*: London, Academic Press, p. 445 - 485.
- [75] Haldorsen, H. and L. Lake. (1984). A new approach to shale management in field-scale models. *Old SPE J*, 24(4): 447–457.
- [76] Harris, G.D., R.W. Wellner, V. Catterall, S. Kairo, C. Liu, and Y. Chen. (2012). Stratigraphy and Depositional Environment of the Upper Zubair Sandstone (“Main Pay”), West Qurna 1 Field, Iraq. IR17, paper presented at the EAGE Workshop on Iraq: Hydrocarbon Exploration and Field Development, (29 April – 2 May), Istanbul, Turkey.
- [77] Hastie, T. J. and R. J. Tibshirani. (1990). *Generalized additive models*. (Vol. 43). CRC Press.
- [78] Hastie T, Tibshirani R, Friedman J. H. (2001). *The elements of statistical learning : data mining, inference, and prediction*. Springer series in statistics New York, Springer 2001, xvi, 533.
- [79] Hegstad, B. K. and O. Henning. (2001). Uncertainty in Production Forecasts Based on Well Observations, Seismic Data, and Production History. *SPE Journal*, 6(01), 409-424.
- [80] Hoeting, J. A., D. Madigan, A. E. Raftery, and C. T. Volinsky. (1999). Bayesian model averaging: A tutorial. *Statistical Science*, 14(4), 382 – 401.
- [81] Hothorn, T. and B. S. Everitt. (2009). *A handbook of statistical analyses using R*. CRC Press.
- [82] Hove, K., G. Olsen, S. Nilsson, M. Tonnesen, and A. Hatloy. (1992). From Stochastic Geological Description to Production Forecasting in Heterogeneous Layered Reservoirs. SPE-24890-MS paper presented at the SPE Annual Technical Conference and Exhibition, (4-7 October), Washington, D.C.
- [83] Isaaks, E. (1990). *The application of Monte Carlo methods to the analysis of spatially correlated data*. PhD Thesis, Stanford University, Stanford, CA, USA.
- [84] Isaaks, E. H. and R.M. Srivastava. (1989). *An Introduction to Applied Geostatistics*. Oxford University Press.

- [85] Jin, R., W. Chen, and A. Sudjianto. (2005). An Efficient Algorithm for Constructing Optimal Design of Computer Experiments. *Journal of Statistical Planning and Inferences*, 134(1), 268-287.
- [86] Journel, A. G. (1989). Geostatistical Characterization of Reservoir Heterogeneities. In *SEG Annual Meeting*, (29 October-2 November), Dallas, Texas.
- [87] Journel, A. G. (1990). Geostatistics for reservoir characterization. SPE-20750-MS, presented at the SPE Annual Technical Conference and Exhibition, (23-26 September), New Orleans, Louisiana.
- [88] Journel, A. G. and F. G. Alabert. (1990). New method for reservoir mapping. *Journal of Petroleum technology*, 42(02), 212-218.
- [89] Journel, A. G., and J. J. Gomez-Hernandez. (1993). Stochastic imaging of the Wilmington clastic sequence. *SPE Formation Evaluation*, 8(01), 33-40.
- [90] Kass, R. E. and A. E. Raftery. (1995). Bayes Factors. *Journal of the American Statistical Association*, 90(430), 773 – 795.
- [91] Kaas, I., T. Svanes, J. C. Van Wagoner, G. Hamar, S. Jorgenvag, P. I. Skarnes, and O. Sundt. (1994). The use of high resolution sequence stratigraphy and stochastic modelling to reservoir management of the Ness Formation of the Statfjord Field, offshore Norway. ed Johnson S. D. (Department of Earth Sciences University of Liverpool), pp 57–58.
- [92] Kabir, C. S., M. C. H. Chien, and J. L. Landa. (2003). Experiences With Automated History Matching. Society of Petroleum Engineers. doi:10.2118/79670-MS.
- [93] Kabir, C. S., N. I. Mohammed, and M. K. Choudhary. (2007). Lessons Learned From Energy Models: Iraq's South Rumaila Case Study. Paper SPE 105131, presented at the SPE Middle East Oil and Gas Show and Conference, (11-14 March) Manama, Bahrain.
- [94] Kalla, S., and C. D. White. (2007). Efficient Design of Reservoir Simulation Studies for Development and Optimization. *SPERE* 629-637.
- [95] Kalogerakis, N. (1994). An Efficient Procedure for the Quantification of Risk in Forecasting Reservoir Performance. SPE-27569-MS paper presented at the European Petroleum Computer Conference, (15-17 March), Aberdeen, United Kingdom.
- [96] Krivoruchko, K. (2012). Empirical Bayesian Kriging-Esri: Redlands, CA, USA. <http://www.esri.com/news/arcuser/1012/empirical-byesian-kriging.html>.
- [97] Kulkarni, M. M. and D. N. Rao. (2006). Analytical Modeling of the Forced Gravity Drainage GAGD Process. Paper AIChE 72361, presented at the AIChE Annual Meeting, (12-17 November), San Francisco, CA, USA.

- [98] Lacentre, P. E. and P. M. Carrica. (2003). A Method To Estimate Permeability on Uncored Wells Based on Well Logs and Core Data. SPE-81058-MS, paper presented at the SPE Latin American and Caribbean Petroleum Engineering Conference, (27-30 April), Port-of-Spain, Trinidad and Tobago.
- [99] Lallier, F., S. Vignau, and H. Kombrink. (2013). The use of Geophysical Data in MPS Facies Simulation in a Seismically Tuned Reservoir - a new Approach Based on the Direct Sampling Algorithm. SPE-166029-MS, paper presented at the SPE Reservoir Characterization and Simulation Conference and Exhibition, (16-18 September), Abu Dhabi, UAE. <http://dx.doi.org/10.2118/166029-MS>.
- [100] Landa, J.L. (2001). Technique to Integrate Production and Static Data in a Self-Consistent Way. SPE-71597 presented at the 2001 SPE Annual Technical Conference and Exhibition, New Orleans, LA, 30 September – 3 October.
- [101] Landa, J. L. and R. N. Horne. (1997). A Procedure to Integrate Well Test Data, Reservoir Performance History and 4-D Seismic Information into a Reservoir Description. SPE-38653-MS paper presented at the SPE Annual Technical Conference and Exhibition, (5-8 October), San Antonio, Texas.
- [102] Lazic, Z. R. (2006). *Design of experiments in chemical engineering*. Wiley-Vch.
- [103] Lee, S. H. and A. Datta-Gupta.. (1999). Electrofacies Characterization and Permeability Predictions in Carbonate Reservoirs: Role of Multivariate Analysis and Nonparametric Regression. SPE-56658 paper presented at the SPE Annual Technical Conference and Exhibition, (3-6 October), Houston, Texas, USA.
- [104] Le, N. D. and J. V. Zidek. (1992). Interpolation with uncertain spatial covariances: a Bayesian alternative to kriging. *Journal of Multivariate Analysis*. 43(2),351–374.
- [105] Lenth R. (2009). Response-surface methods in R, using RSM. *Journal of Statistical Software*, 32(7):1–17.
- [106] Li, K. and R. N. Horne. (2003). Prediction of Oil by Gravity Drainage. SPE-84184-MS, paper presented at the SPE ATCE, (5-8 October), Denver, CO.
- [107] Li T., S. H. Zhu, and M. Ogihara. (2006). Using discriminant analysis for multi-class classification: an experimental investigation. *Knowl. Inf. Syst.* 10, 453–472.
- [108] Liu, Y. and A. G. Journel. (2005). Improving sequential simulation with a structured path guided by information content: *Math. Geol.*, v. 38, p. 945–964.
- [109] Liu, Y., A. Harding, A. G. Journel, and R. Gilbert. (2005). A workflow for multiple-point geostatistical simulation, in Leuangthong, O., and Deutsch, C. V., eds., *Geostatistics Banff 2004*, vol. 1: Springer, Dordrecht, p. 245–254.

- [110] Liu, Y., A. Harding, W. Abriel, and S. Strebelle. (2004). Multiple-point simulation integrating wells, three-dimensional seismic data, and geology. AAPG bulletin, 88(7), 905-921.
- [111] Liu, N., S. Betancourt, and D. S. Oliver. (2001). Assessment of Uncertainty Assessment Methods. SPE-71624-MS presented at the SPE Annual Technical Conference and Exhibition, (30 September-3 October), New Orleans, Louisiana, USA. doi:10.2118/71624-MS.
- [112] MacDonald, A. C., T. H. Høye, P. Lowry, T. Jacobsen, J. O. Aasen and A. O. Grindheim. (1991). Stochastic flow unit modelling of a North Sea coastal-deltaic reservoir. Extended Abstract presented at the 6th European Symposium on Improved Oil Recovery (IOR 1991), (21 May).
- [113] Mathisen, T., S. H. Lee, and A. Datta-Gupta. (2003). Improved Permeability Estimates in Carbonate Reservoirs Using Electrofacies Characterization: A Case Study of the North Robertson Unit, West Texas. SPE Reservoir Evaluation and Engineering Journal: 176-184.
- [114] Mahmoud, T. N. and D. N. Rao. (2007). Mechanisms and Performance Demonstration of the Gas-Assisted Gravity-Drainage Process Using Visual Models. Paper SPE 110132, presented at the SPE Annual Technical Conference and Exhibition, (11-14 November), Anaheim, California, USA.
- [115] Mahmoud, T.N. and D. N. Rao.. (2008). Range of Operability of Gas-Assisted Gravity Drainage Process. Paper SPE 113474, presented at the SPE Symposium on Improved Oil Recovery, (20-23 April), Tulsa, Oklahoma, USA.
- [116] Massonnat, G.J., F.G. Alabert, and C.B. Giudicelli. (1992). Anguille Marine, a Deepsea-Fan Reservoir Offshore Gabon: From Geology to Stochastic Modelling. SPE-24709-MS paper presented at the SPE Annual Technical Conference and Exhibition, (4-7 October), Washington, D.C.
- [117] McKay, M.D., R.J. Beckman, and W.J. Conover. (1979). A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code. Technometrics (JSTOR Abstract) (American Statistical Association) 21 (2): 239-245. doi:10.2307/1268522.
- [118] Mikes, D. and C.R. Geel. (2006). Standard facies models to incorporate all heterogeneity levels in a reservoir model: Marine and Petroleum Geology v.23, p. 943-959.
- [119] Minka, T. P. (2005). A statistical learning/pattern recognition glossary. Online: <http://alumni.media.mit.edu/~tpminka/statlearn/glossary>.
- [120] Mohammed, W. J, M. S. Al Jawad, and D A. Al-Shamaa. (2010). Reservoir Flow Simulation study for a Sector in Main Pay-South Rumaila Oil Field. SPE-126427-MS paper presented at the SPE Oil and Gas India Conference and Exhibition, (20-22 January), Mumbai, India.

- [121] Montgomery, D. C. (1997). *Design and Analysis of Experiments*. 5th Edition. John Wiley and Sons, INC.
- [122] Montgomery, D. C. and G. C. Runger. (2003). *Applied Statistics and Probability for Engineers*. 3rd Edition. John Wiley and Sons, INC.
- [123] Moore, W., Z. Ma, J. Urdea, and T. Bratton. (2011). Uncertainty Analysis in Well Log and Petrophysical Interpretations. In book: *Uncertainty Analysis and Reservoir Modeling*, Chapter: 2, Publisher: AAPG, Editors: Y. Zee Ma and Paul LaPointe, pp.17-28.
- [124] Nashawi, I. S. and A. Malallah. (2010). Permeability Prediction from Wireline Well Logs Using Fuzzy Logic and Discriminant Analysis. SPE-133209 paper presented at the SPE Asia Pacific Oil and Gas Conference and Exhibition, (18-20 October), Brisbane, Queensland, Australia.
- [125] Nashawi, I. S., and A. Malallah. (2009). Improved Electrofacies Characterization And Permeability Predictions In Sandstone Reservoirs Using a Data Mining And Expert System Approach. PETROPHYSICS: 250-268.
- [126] Neuman, S. P. (2003). Maximum likelihood Bayesian averaging of uncertain model predictions. *Stochastic Environmental Research and Risk Assessment* 17(5), 291-305.
- [127] Okano, H., G. E. Pickup, M. A. Christie, S. Subbey, and M. Sambridge. (2006). Quantification of Uncertainty Due to Sub-Grid Heterogeneity in Reservoir Models. SPE-100223-MS paper presented at the SPE Europec/EAGE Annual Conference and Exhibition, (12-15 June), Vienna, Austria.
- [128] Oliver, D. S. (1996). Multiple Realizations of the Permeability Field from Well Test Data. SPEJ 1(2): 145-154. SPE-27970-PA.
- [129] Omre, H. (1987). Bayesian kriging-Merging observations and qualified guesses in kriging. *Mathematical Geology*. 19(1), 25-39.
- [130] Omre, H. and K. B. Halvorsen. (1989). Bayesian Kriging-merging observations and qualified guesses in Kriging. *Mathematical Geology*. 21(7), 767-786.
- [131] Overeem, I. (2008). Geological modeling introduction: Lecture, Community Surface Dynamic Modeling System, University of Colorado at Boulder.
- [132] Paidin, W.R., P. Mwangi, and D. N. Rao. (2010). Economic Evaluation Within the Scope of the Field Development and Application of the Gas-Assisted Gravity Drainage (GAGD) Process in an Actual Northern Louisiana Field. Paper SPE 129723, presented at the SPE Hydrocarbon Economics and Evaluation Symposium, (8-9 March), Dallas, Texas, USA.
- [133] Pasternack, E. (2009). Uncertainty in Petrophysical Evaluation. *Search and Discovery Article*, No. 120011.

- [134] Perez, H. H., A. Datta-Gupta, and S. Mishra. (2005). The Role of Electrofacies, Lithofacies, and Hydraulic Flow Units in Permeability Predictions from Well Logs: A Comparative Analysis Using Classification Trees. *SPE Reservoir Evaluation and Engineering Journal*: 143-155.
- [135] Petrel. (2014). Software Manual. Schlumberger Information Solutions.
- [136] Pires, A.M. and J.A. Branco. (2010). Projection-pursuit approach to robust linear discriminant analysis. *Journal of Multivariate Analysis* 101 2464–2485.
- [137] Pyrcz, M. J. and C. V. Deutsch. (2014). *Geostatistical reservoir modeling*. Oxford university press.
- [138] Raftery, A. E. (1995). Bayesian model selection in social research. *Sociological Methodology*, 25,111–163.
- [139] Raftery, A. E. (1996). Approximate Bayes factors and accounting for model uncertainty in generalized linear models. *Biometrika*, 83(2), 251 – 266.
- [140] Raftery, A. E., J. Hoeting, C. Volinsky, I. Painter, and K. Y. Yeung. (2010). Package 'BMA'.
- [141] Raftery, A. E., D. Madigan, and J. A. Hoeting. (1997). Bayesian model averaging for linear regression models. *Journal of the American Statistical Association*, 92(437), 179-191.
- [142] Rao, D. N., S. C. Ayirala, and M. M. Kulkarni. (2004). Development of Gas-Assisted Gravity Drainage (GAGD) Process for Improved Light Oil Recovery. Paper SPE 89357, presented at the SPE/DOE Symposium on Improved Oil Recovery, (17-21 April), Tulsa, Oklahoma, USA.
- [143] Rao, D.N. (2012). *Gas-Assisted Gravity Drainage Process for Improved Oil Recovery*. United States patent 8,215,392 B2, July 2012.
- [144] Rao, D. N., and et al. (2006). *Development and Optimization of Gas Assisted Gravity Drainage (GAGD) process for improved light oil recovery*. United States Department of Energy Research Proposal, Report Number 15323R15, Final Technical Report, prepared for National Energy Technology Laboratory, United States Department of Energy.
- [145] Rao, D. N., S. C. Ayirala, A. A. Abe, and W. Xu. (2006). Impact of Low-Cost Dilute Surfactants on Wettability and Relative Permeability. Paper SPE 99609 presented at the 2006 SPE/DOE Symposium on Improved Oil Recovery, (April 22-26), Tulsa, Oklahoma, USA.
- [146] Ripley, B. (1996). *Pattern Recognition and Neural Networks*. Cambridge, Mass. Cambridge Univ. Press.

- [147] Rivoirard, J. (2005). Concepts and Methods of Geostatistics. In Space, Structure and Randomness (17-37). Springer New York.
- [148] Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics*, 6, 461 – 464.
- [149] Seifert, D. and J. L. Jensen. (1999). Using Sequential Indicator Simulation as a Tool in Reservoir Description: Issues and Uncertainties. *Mathematical Geology*, Vol. 31, No. 5, 1999.
- [150] Stein, M. (1987). Large Sample Properties of Simulations Using Latin Hypercube Sampling. *Technometrics*. 29, 143–151.
- [151] Silva, F. P. T., A. A. Ghani, A. Al Mansoori, and A. Bahar. (2002). Rock Type Constrained 3D Reservoir Characterization and Modeling. SPE 78504 paper presented at the Abu Dhabi International Petroleum Exhibition and Conference, (13-16 October), Abu Dhabi, UAE.
- [152] Specht, D. F. (1988). Probabilistic neural networks for classification, mapping, or associative memory. In *Neural Networks, IEEE International Conference*. (pp. 525-532).
- [153] Srivastava, R. M. (1994). An overview of stochastic methods for reservoir characterization, in Yarush, J. M., and Chambers, R. L., eds., *AAPG computer applications in geology no. 3, Stochastic modeling and geostatistics: Principles, methods and case studies: American Association of Petroleum Geologists, Tulsa, OK*, p. 3–16.
- [154] Stocki, R. (2005). A method to improve design reliability using optimal Latin hypercube sampling. *Computer Assisted Mechanics and Engineering Sciences*. 12, 87–105.
- [155] Stoyan, D., W. S. Kendall, and J. Mecke. (1987). *Stochastic geometry and its applications*. New York, John Wiley and Sons, 345 p.
- [156] Strebelle, S. (2003). New multiple-point statistics simulation implementation to reduce memory and CPU-demand, in *Proceedings to the IAMG*, (7–12 September), Portsmouth, UK.
- [157] Strebelle, S. and A. G. Journel. (2001). Reservoir modeling using multiple-point statistics. SPE-71324-MS paper presented at the SPE Annual Technical Conference and Exhibition, (30 September-3 October), New Orleans, Louisiana.
- [158] Subbey S., C. Mike, and M. Sambridge. (2003). A Strategy for Rapid Quantification of Uncertainty in Reservoir Performance Prediction. SPE-79678-MS paper presented at the SPE Reservoir Simulation Symposium, (3-5 February), Houston, Texas.

- [159] Taware, S. V., A. G. Taware, A. K. Sinha, A. Jamkhindikar, R. Talukdar, and A. Datta-Gupta. (2008). Integrated Permeability Modeling Using Wireline Logs, Core and DST Data in a Deepwater Reservoir. SPE 113599 paper presented at the SPE Indian Oil and Gas Technical Conference and Exhibition, (4-6 March), Mumbai, India.
- [160] Teh, W., G. P. Willhite, and J. H. Doveton. (2012). Improved Reservoir Characterization in the Ogallah Field using Petrophysical Classifiers within Electrofacies. SPE-154341 paper presented at the SPE Improved Oil Recovery Symposium, (14-18 April), Tulsa, Oklahoma, USA.
- [161] Tiab, D. and E. C. Donaldson. (2004). *Petrophysics: Theory and Practice of Measuring Reservoir Rock and Fluid Transport Properties*. 2nd Edition. Elsevier.
- [162] Vapnik, V. (1996). *The Nature of Statistical Learning Theory*. Springer, New York.
- [163] Van Essen, G., M. Zandvliet, P. Van den Hof, O. Bosgra, and J. D. Jansen. (2006). Robust Waterflooding Optimization of Multiple Geological Scenarios. SPE-102913-MS presented at the SPE Annual Technical Conference and Exhibition, ( 24-27 September), San Antonio, Texas, USA. doi:10.2118/102913-MS.
- [164] Walker, R. G. (1992). Facies, facies models and modern stratigraphic concepts: Facies Models, edited by Walker, R.G., and N.P. James, Geological Association of Canada, p. 1-14.
- [165] Wasserman, L. (2000). Bayesian model selection and model averaging. *Journal of Mathematical Psychology* 44(1), 92-107.
- [166] Wells, M., D. Kitching, D. Finucane, and B. Kostic. (2013). An Integrated Description of the Stratigraphy and Depositional Environment of the "Main Pay" Member of the Zubair Formation, Rumaila, Iraq. IR21, paper presented at the Second EAGE Workshop on Iraq, (15 - 18 September), Dead sea, Jordan.
- [167] Wen, Xian-Huan, C. V. and A. S. Cullick. (1997). High Resolution Reservoir Models Integrating Multiple-Well Production Data. SPE-38728-MS paper presented at the SPE Annual Technical Conference and Exhibition, (5-8 October), San Antonio, Texas.
- [168] White, C.D., B.J. Willis, K. Narayanan, and S.P. Dutton. (2001). Identifying and Estimating Significant Geologic Parameters with Experimental Design. SPEJ 311-324.
- [169] White, C.D., and S.A. Royer. (2003). Experimental Design as a Framework for Reservoir Studies. Paper SPE 79676 presented at the Reservoir Simulation Symposium, Houston, Texas, USA.



- [170] Woodhouse, R. (2003). *Statistical Regression Line-Fitting in the Oil and Gas Industry*. Tulsa, Oklahoma: PennWell.
- [171] Wu J, T. Zhang , and A. Boucher. (2007). Non-stationary multiple-point geostatistical simulations with region concept. In: Proceedings of the 20th SCRF meeting, Stanford, CA, USA.
- [172] Xu, W., T.T. Tran, R.M. Srivastava, and A.G. Journel. (1992). Integrating Seismic Data in Reservoir Modeling: The Collocated Cokriging Alternative. SPE-24742-MS presented at the SPE Annual Technical Conference and Exhibition, (4-7 October), Washington, DC.
- [173] Yang, C., Card, C., Nghiem, L. X., and Fedutenko, E. (2011). Robust optimization of SAGD operations under geological uncertainties. In SPE Reservoir Simulation Symposium. Society of Petroleum Engineers.
- [174] Yang, C., L.X. Nghiem, C. Card, and M. Breimeier. (2007). Reservoir Model Uncertainty Quantification through Computer-Assisted History Matching. SPE Annual Technical Conference and Exhibition, Anaheim, California, USA.
- [175] Ye, M., S.P. Neuman, and P.D. Meyer. (2004). Maximum likelihood Bayesian averaging of spatial variability models in unsaturated fractured tuff. Water Resources Research, 40, W05113, doi:10.1029/2003WR002557.
- [176] Yerramilli, S. S., R. C. Yerramilli, N. Vedanti, M. K. Sen, and R. P. Srivastava. (2013). Integrated Reservoir Characterization of an Unconventional Reservoir Using 3D Seismic and Well Log Data: A Case Study of Balol Field, India. SEG-2013-0548 Extended Abstract presented at the SEG Annual Meeting, (22-27 September), Houston, Texas, USA.
- [177] Yeten, B., A. Castellini, B. Guyaguler, and W.H. Chen. (2005). A Comparison Study on Experimental Design and Response Surface Methodologies. Paper SPE 93347 presented at the SPE Reservoir Simulation Symposium, Woodlands, TX.
- [178] Zhang, G. (2003). *Estimating Uncertainties in Integrated reservoir Studies*. PhD Dissertation, Texas A&M University.
- [179] Zhang, T. (2008). Incorporating geological conceptual models and interpretations into reservoir modeling using multiple-point geostatistics: Earth Science Frontiers, v.15, no.1, p. 26-35.
- [180] Zhang, T., S. Bombarde, S. B. Strebelle, and E. Oatney. (2006). 3D Porosity Modeling of a Carbonate Reservoir Using Continuous Multiple-Point Statistics Simulation. SPE Journal, 3(11) 375-379.

## A. Bivariate Statistical Analysis

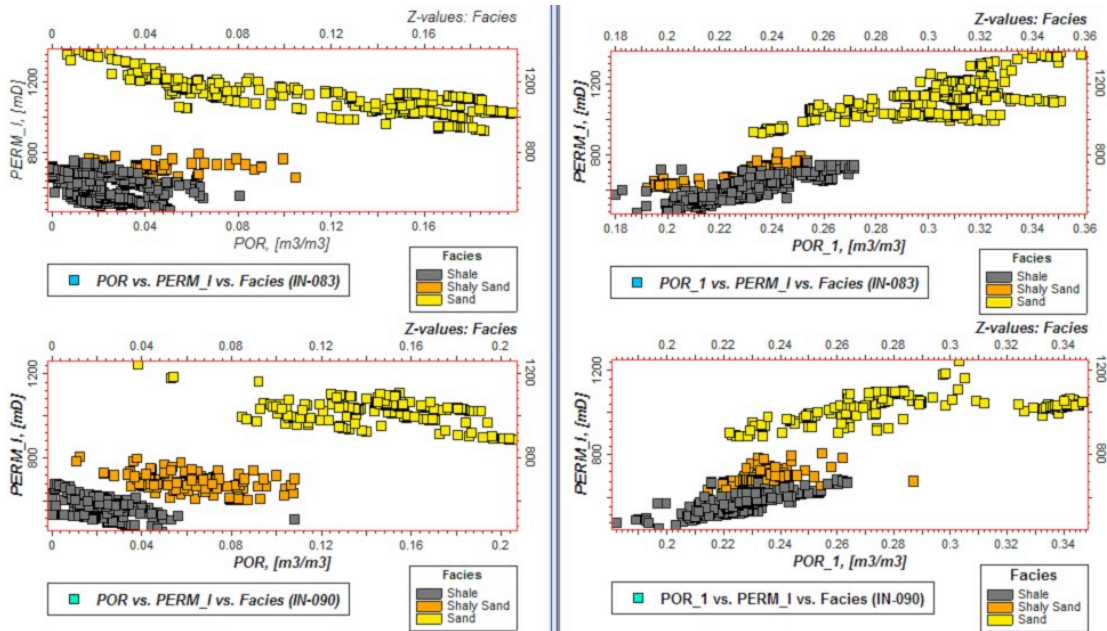


Figure A1: Scatter Plots of the Log and Core Porosity for wells IN-083 & IN-090

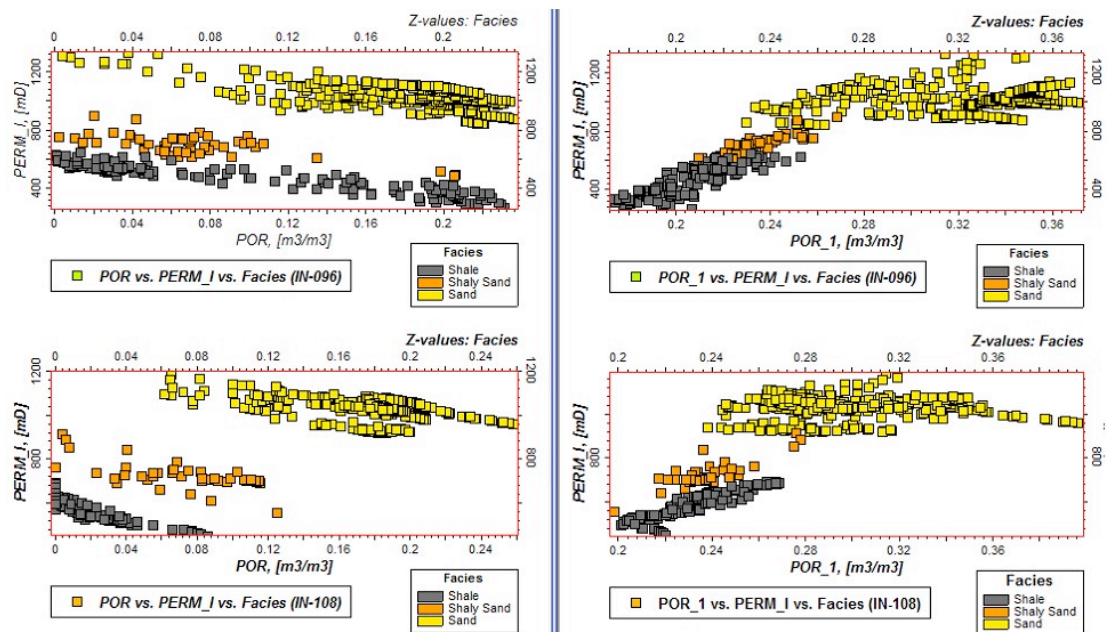


Figure A2: Scatter Plots of the Log and Core Porosity for wells IN-096 & IN-108

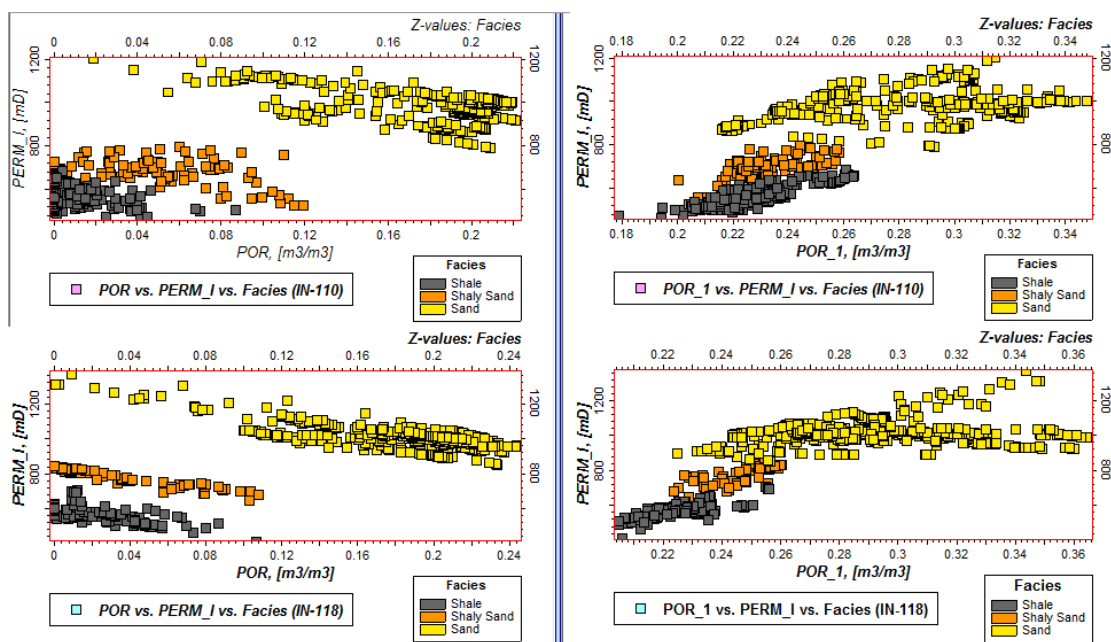


Figure A3: Scatter Plots of the Log and Core Porosity for wells IN-110 & IN-118

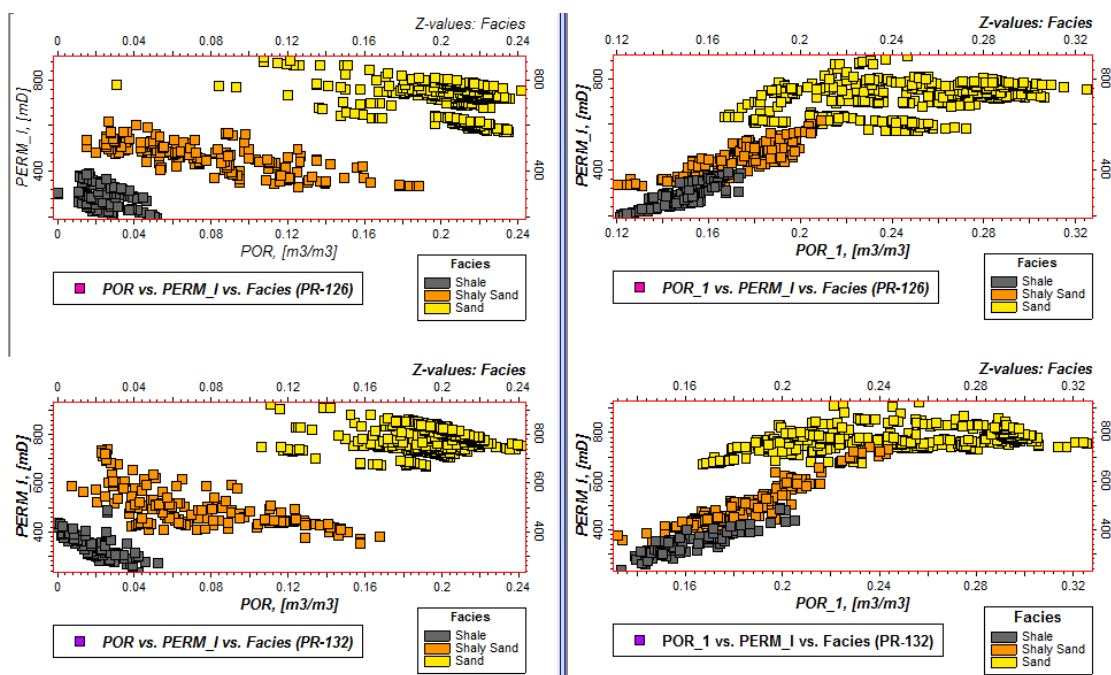


Figure A4: Scatter Plots of the Log and Core Porosity for wells IN-126 & IN-132

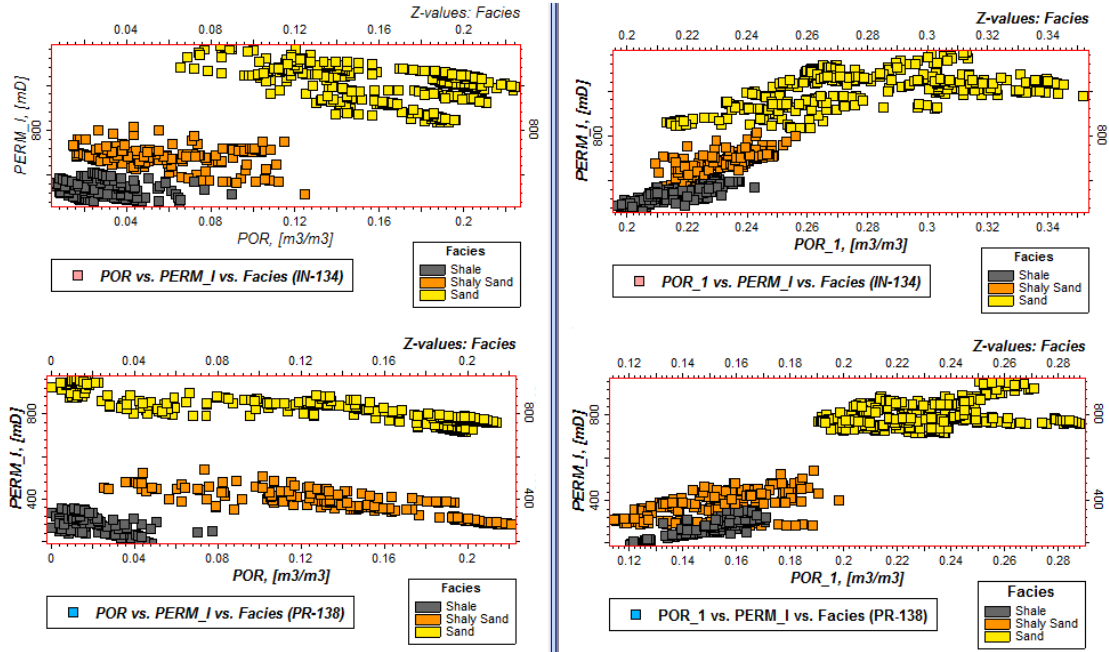


Figure A5: Scatter Plots of the Log and Core Porosity for wells IN-134 & PR-138

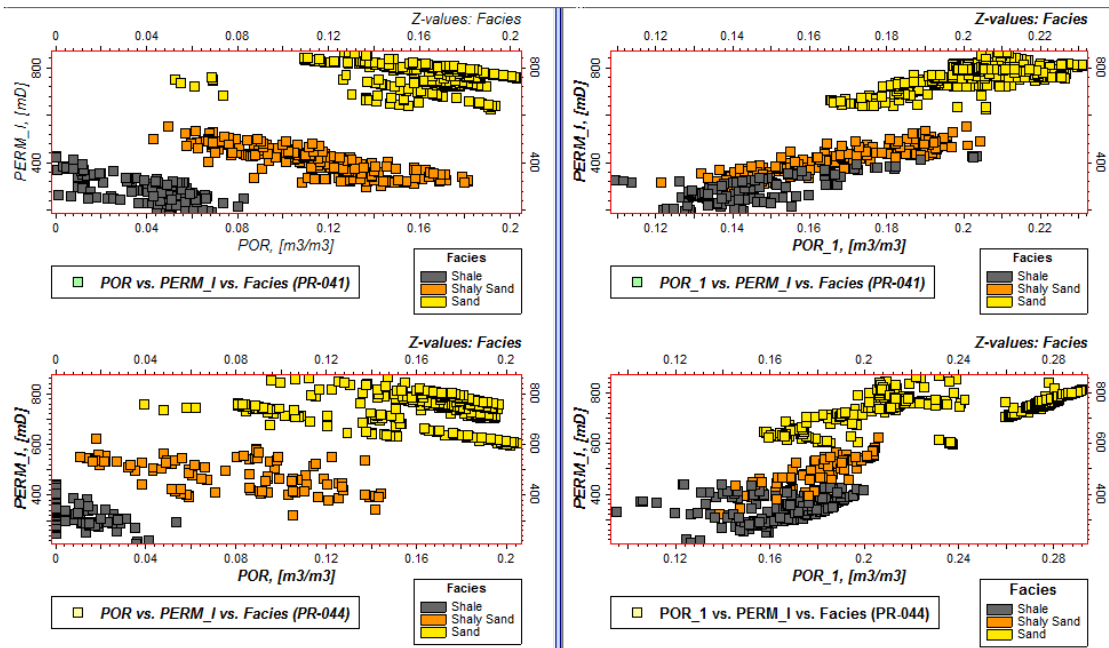


Figure A6: Scatter Plots of the Log and Core Porosity for wells PR-041 & PR-044

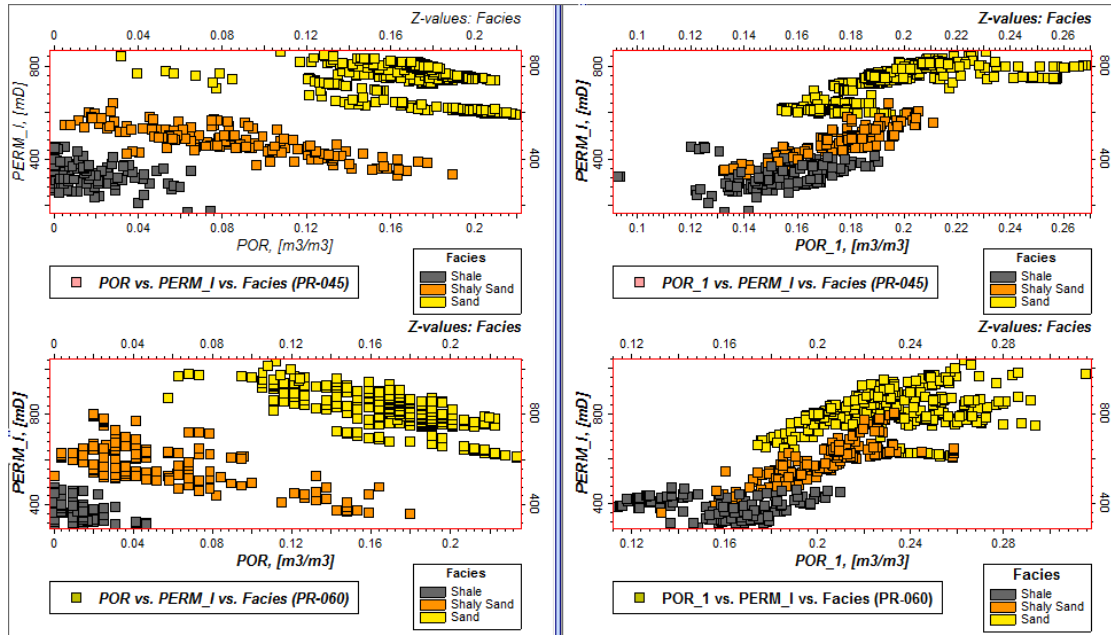


Figure A7: Scatter Plots of the Log and Core Porosity for wells PR-045 & PR-060

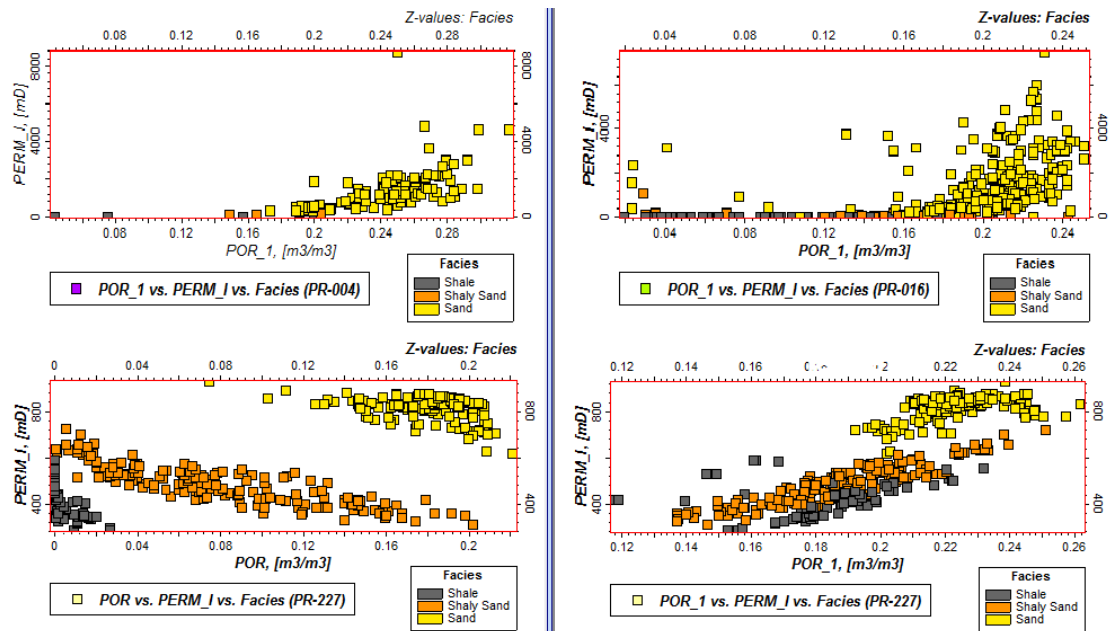


Figure A8: Scatter Plots of the Log and Core Porosity for wells PR-004, PR-016, & PR-227



## B. Multivariate Petrophysical Properties

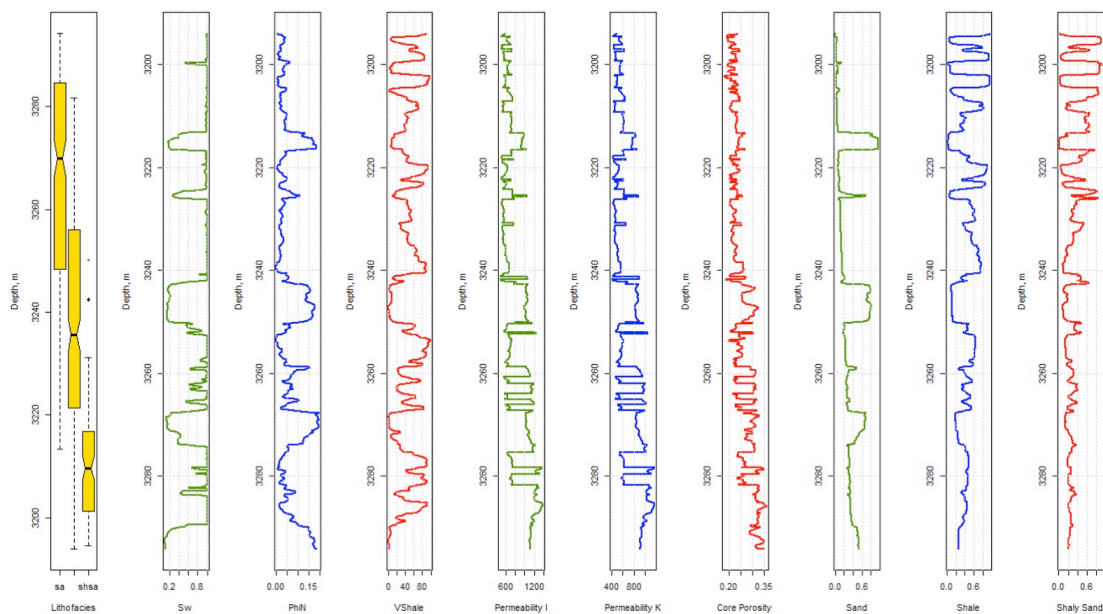


Figure B1: Well Logs, & Lithofacies and Petrophysical Prediction for Well IN-083

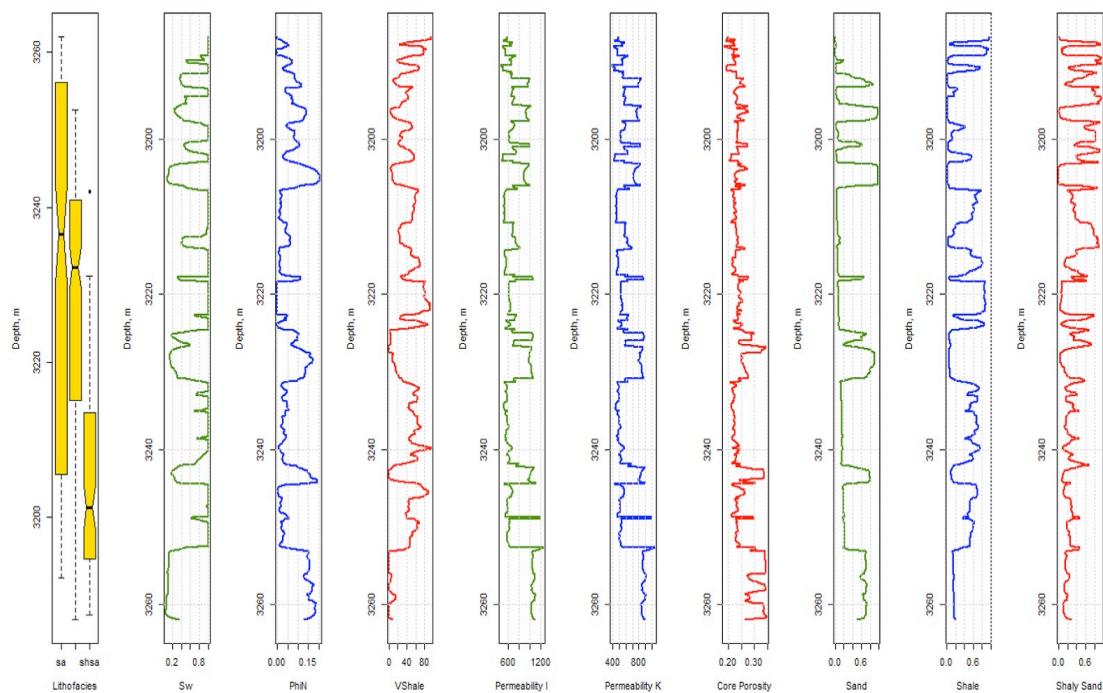


Figure B2: Well Logs & Lithofacies and Petrophysical Prediction for Well IN-090

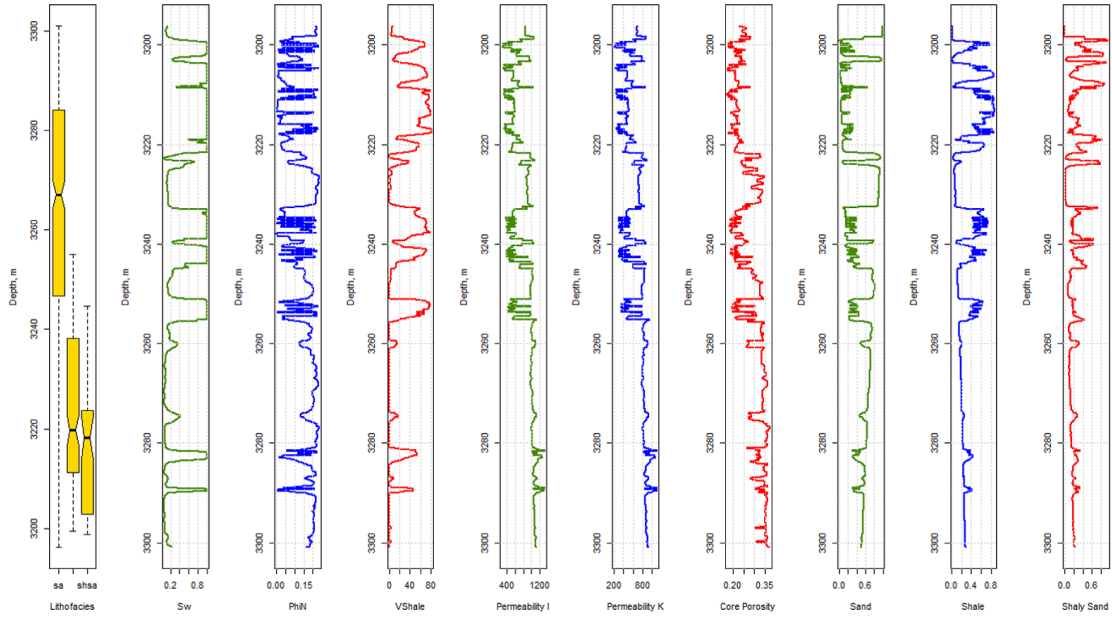


Figure B3: Well Logs & Lithofacies and Petrophysical Prediction for Well IN-096

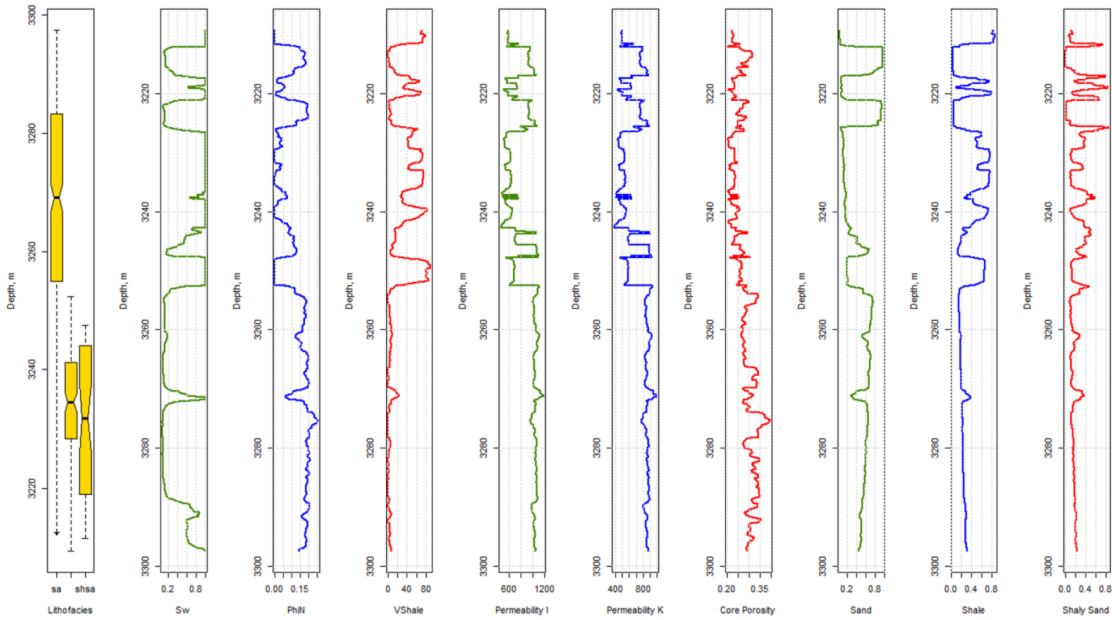


Figure B4: Well Logs & Lithofacies and Petrophysical Prediction for Well IN-108

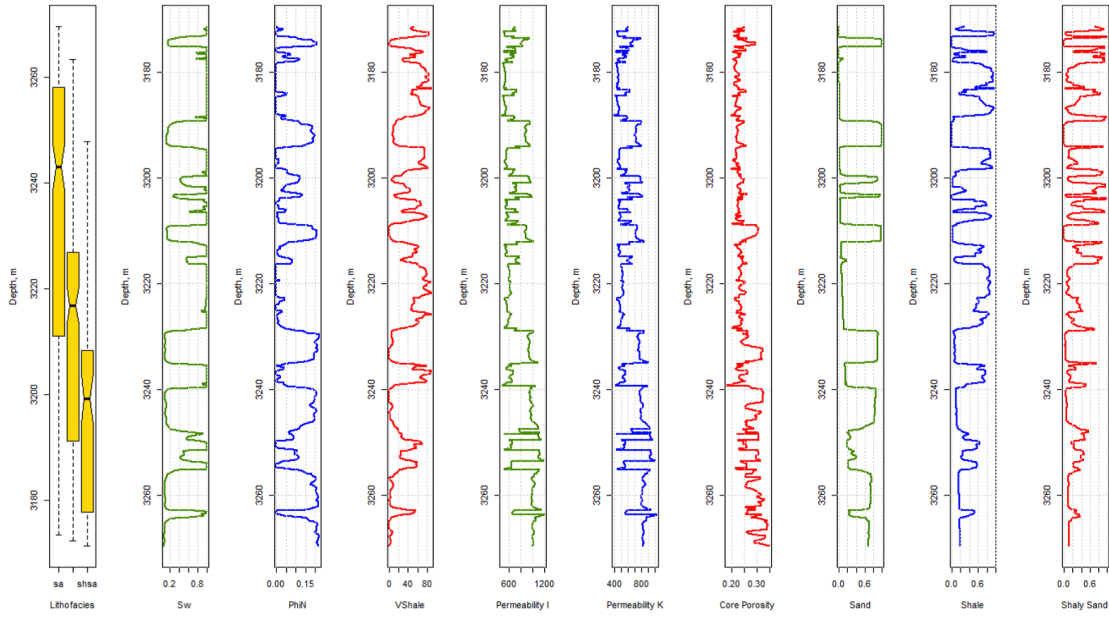


Figure B5: Well Logs & Lithofacies and Petrophysical Prediction for Well IN-110

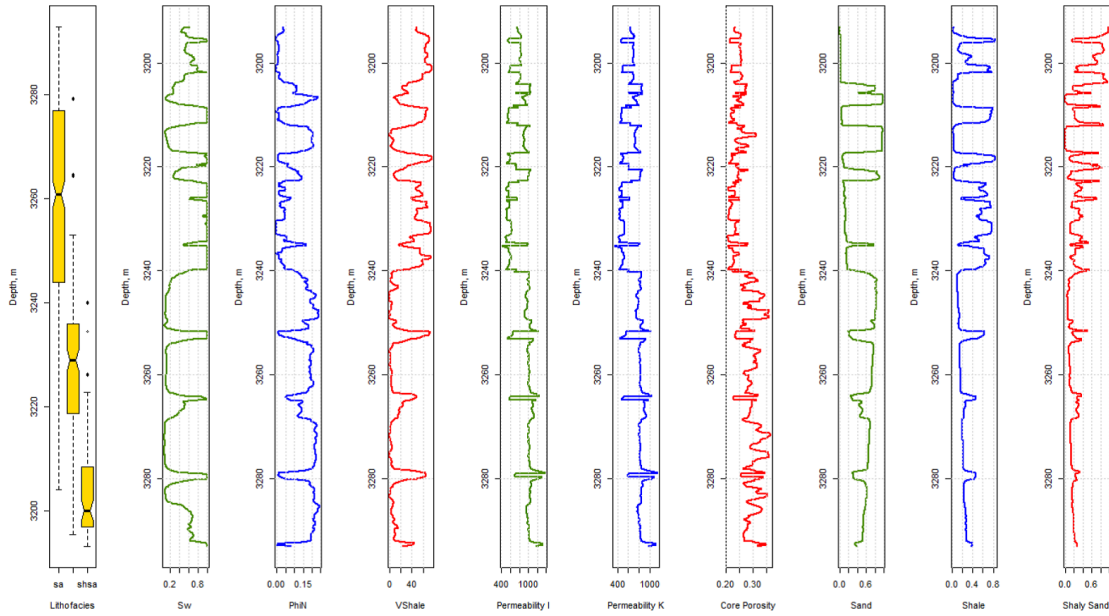


Figure B6: Well Logs & Lithofacies and Petrophysical Prediction for Well IN-118



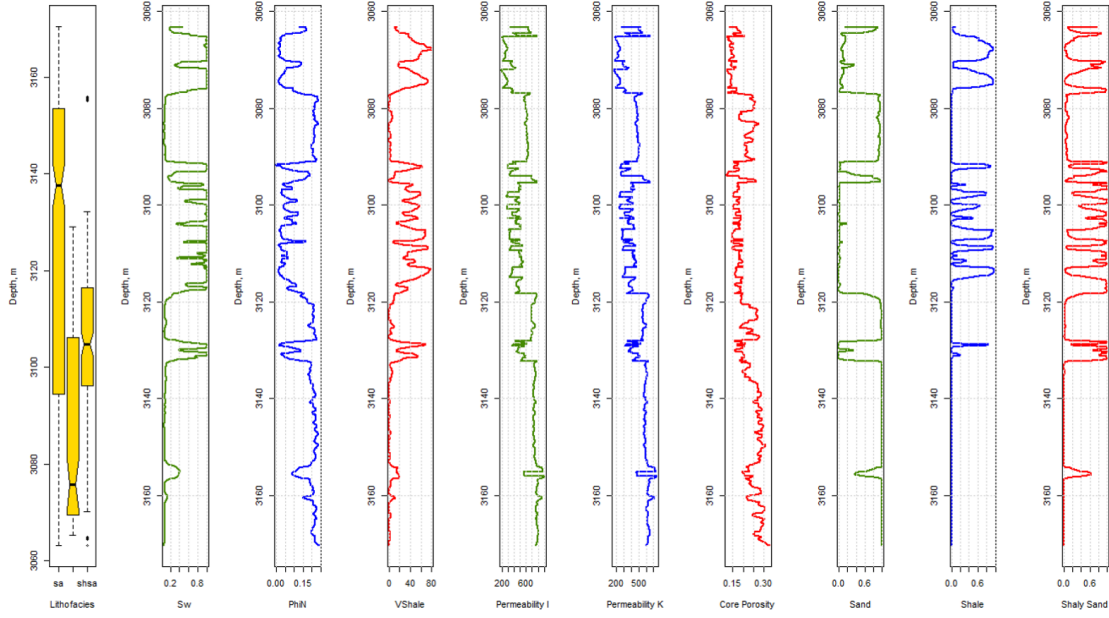


Figure B7: Well Logs & Lithofacies and Petrophysical Prediction for Well IN-126

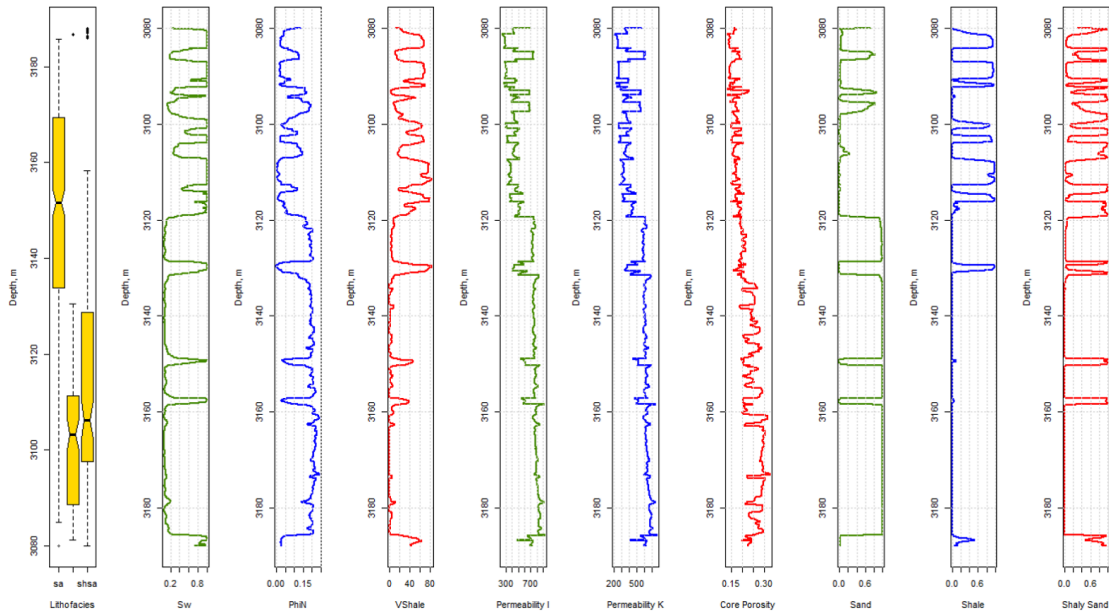


Figure B8: Well Logs & Lithofacies and Petrophysical Prediction for Well IN-132

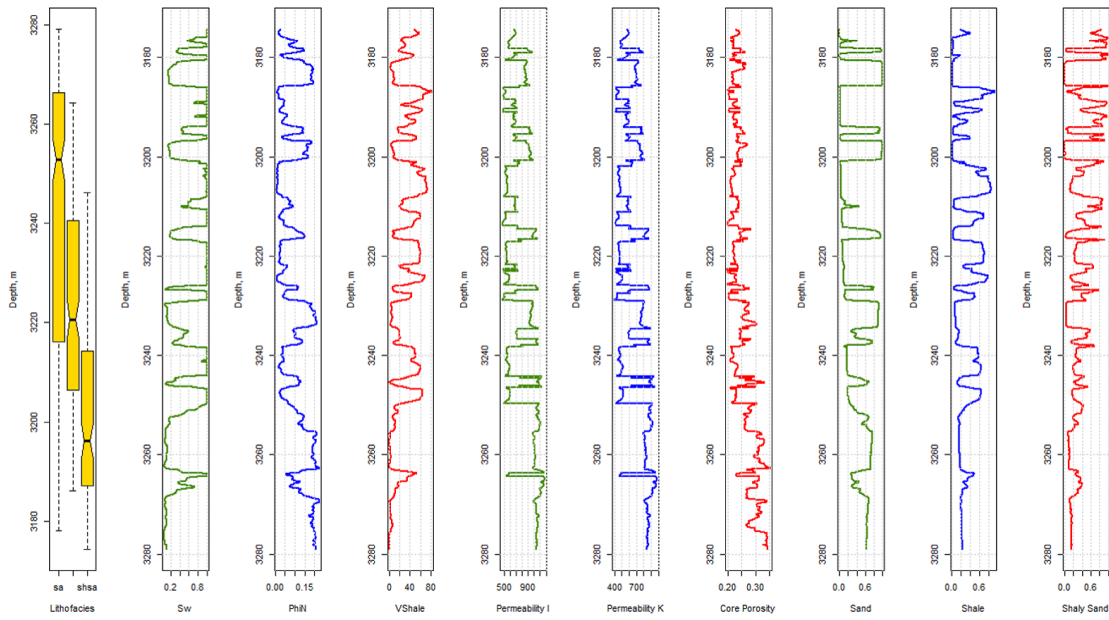


Figure B9: Well Logs & Lithofacies and Petrophysical Prediction for Well IN-134

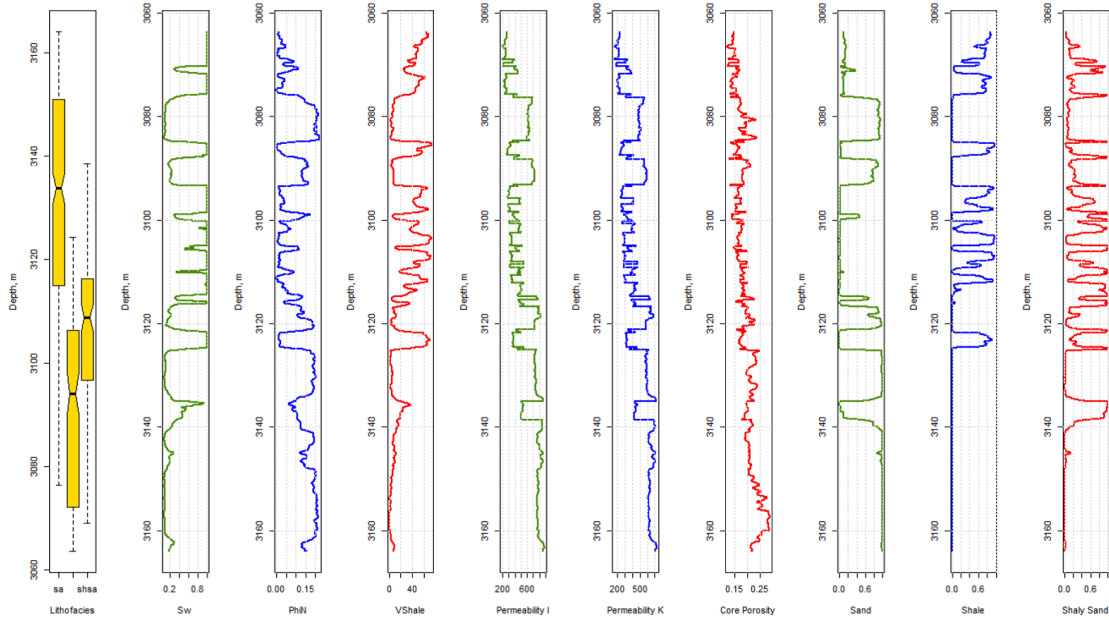


Figure B10: Well Logs & Lithofacies and Petrophysical Prediction for Well PR-002

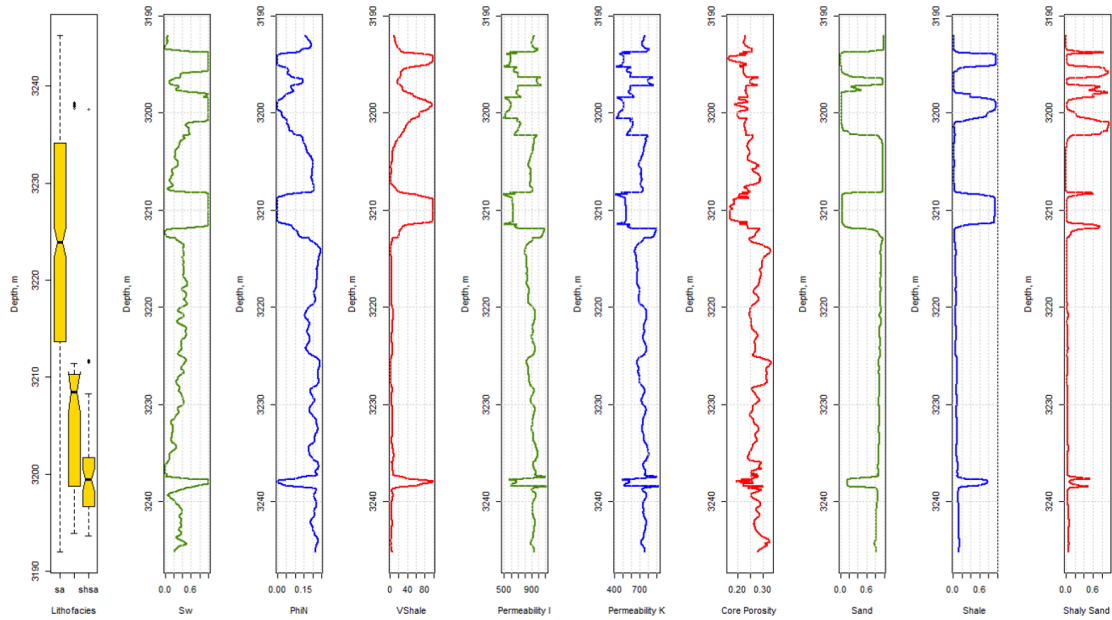


Figure B11: Well Logs & Lithofacies and Petrophysical Prediction for Well PR-004

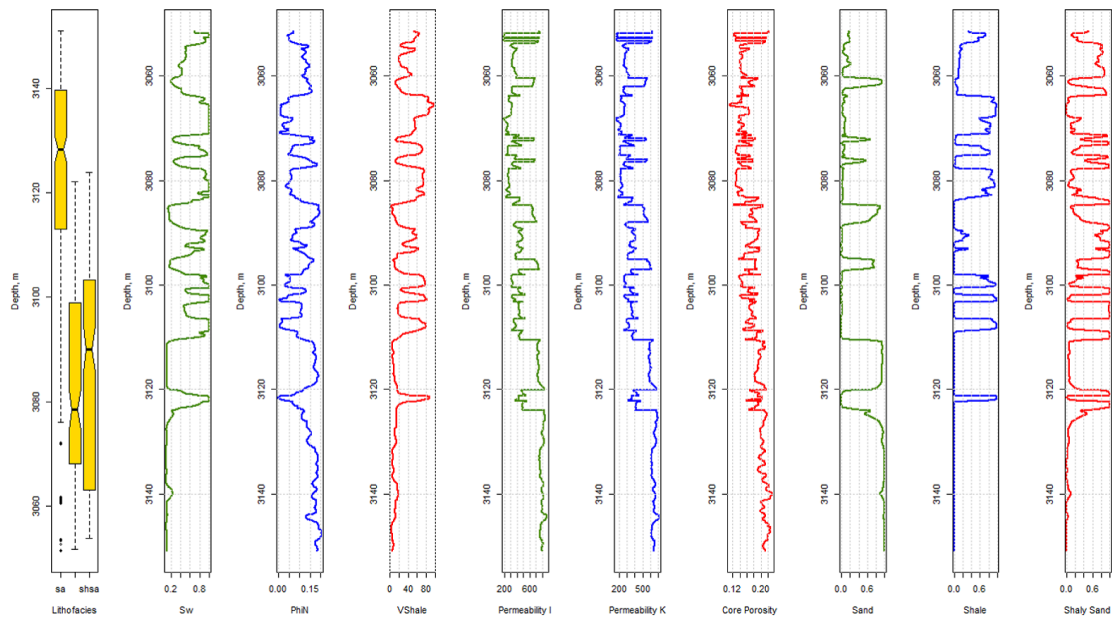


Figure B12: Well Logs & Lithofacies and Petrophysical Prediction for Well PR-016

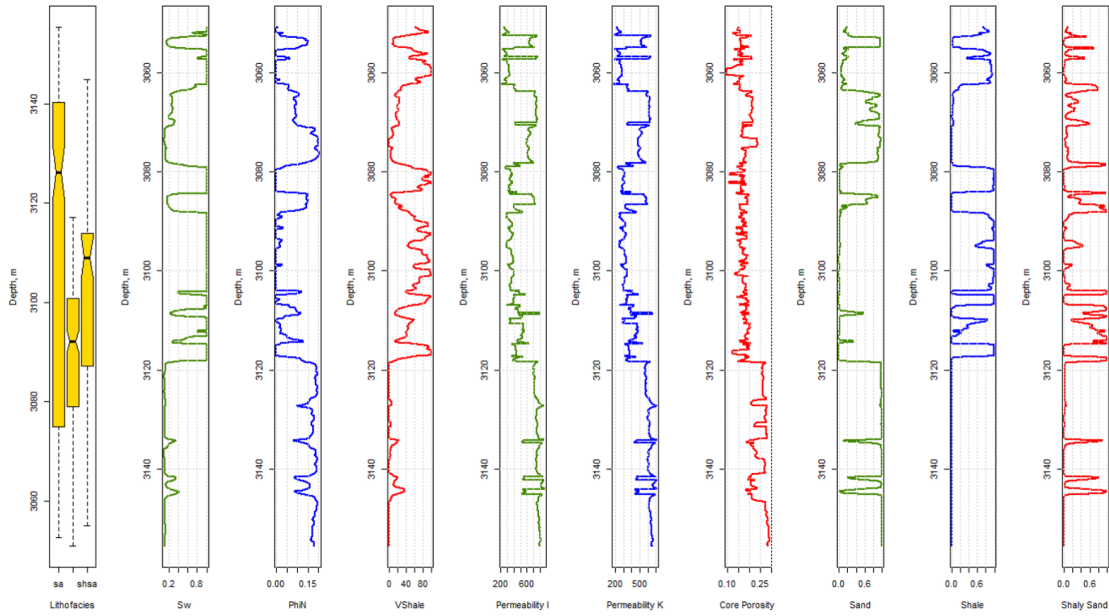


Figure B13: Well Logs & Lithofacies and Petrophysical Prediction for Well PR-041

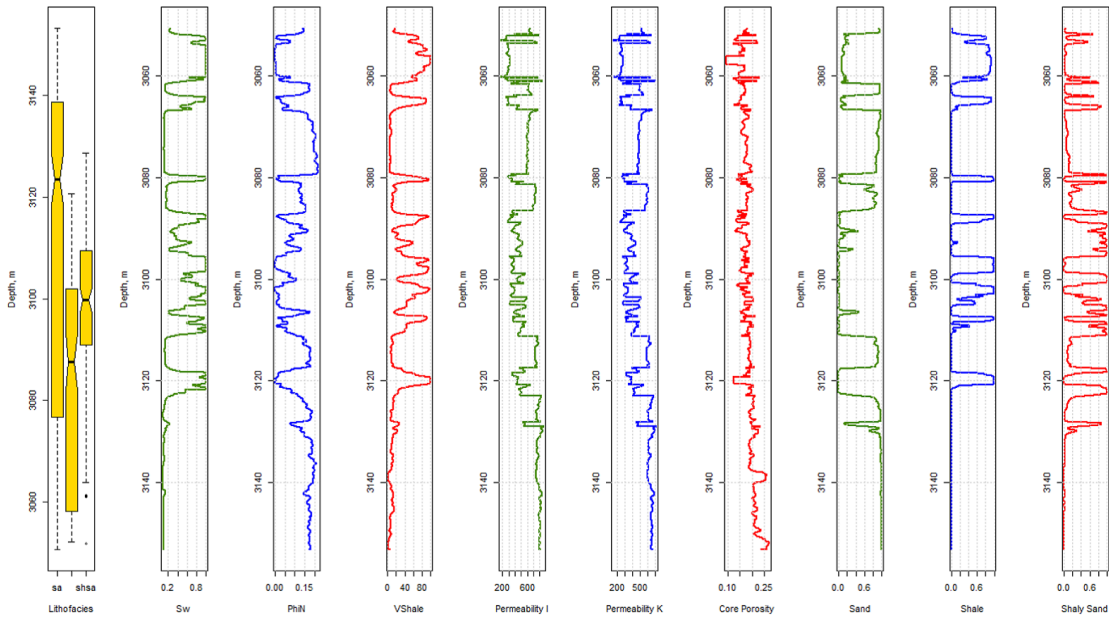


Figure B14: Well Logs & Lithofacies and Petrophysical Prediction for Well PR-045

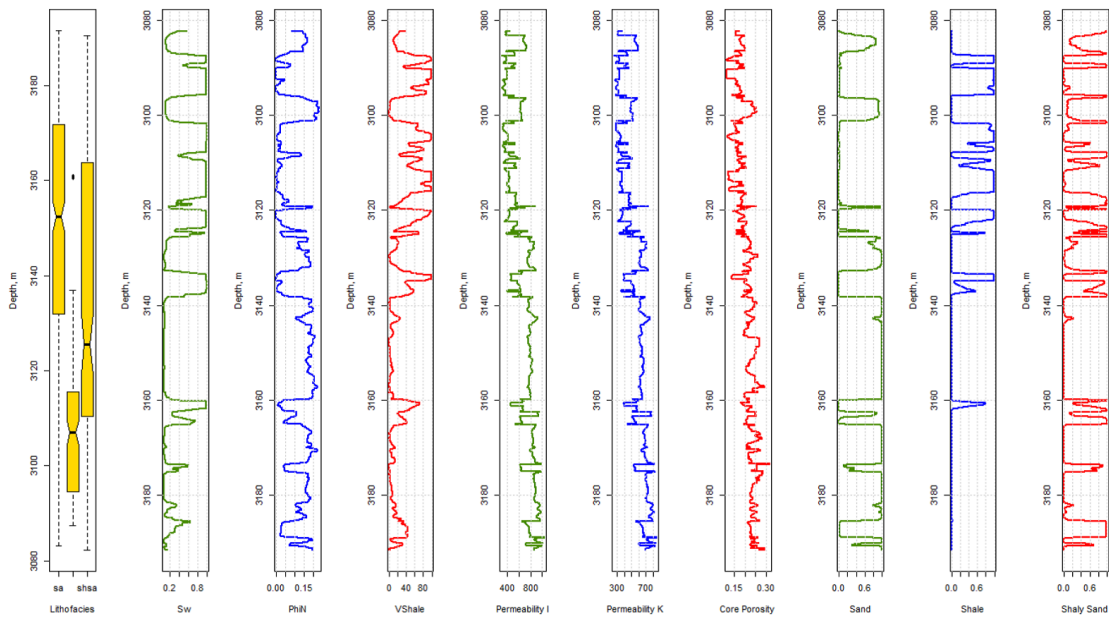


Figure B15: Well Logs & Lithofacies and Petrophysical Prediction for Well PR-060

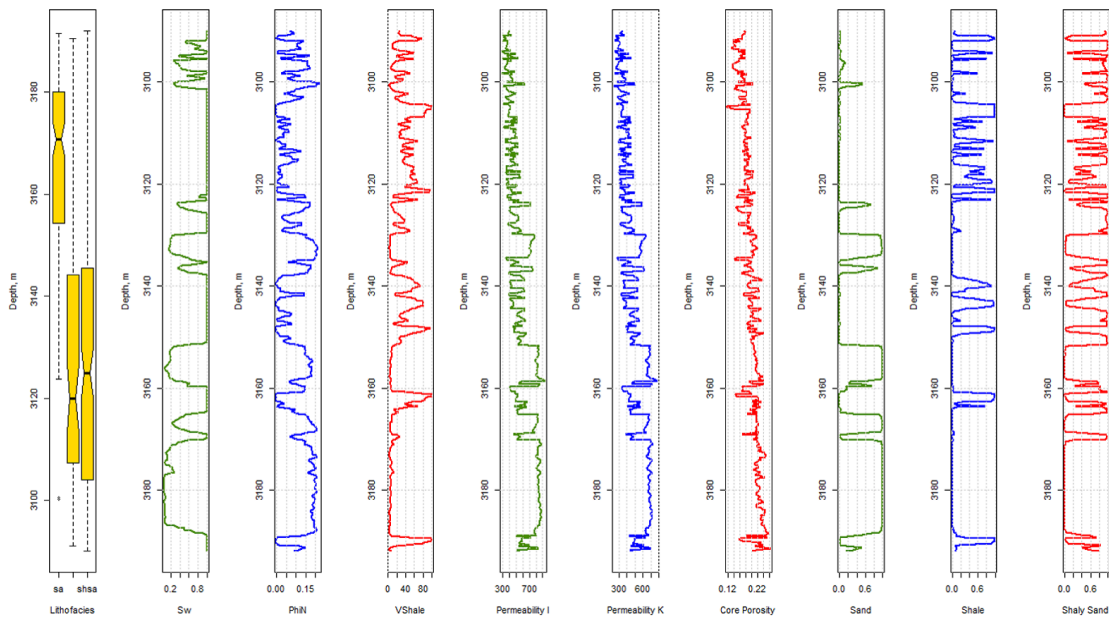


Figure B16: Well Logs & Lithofacies and Petrophysical Prediction for Well PR-158



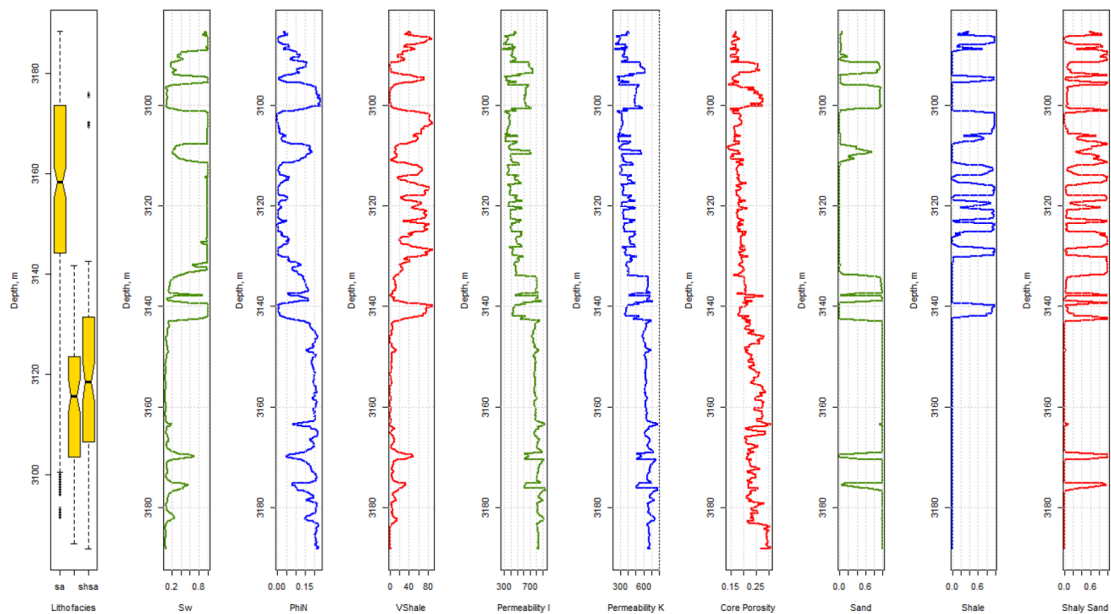


Figure B17: Well Logs & Lithofacies and Petrophysical Prediction for Well PR-227

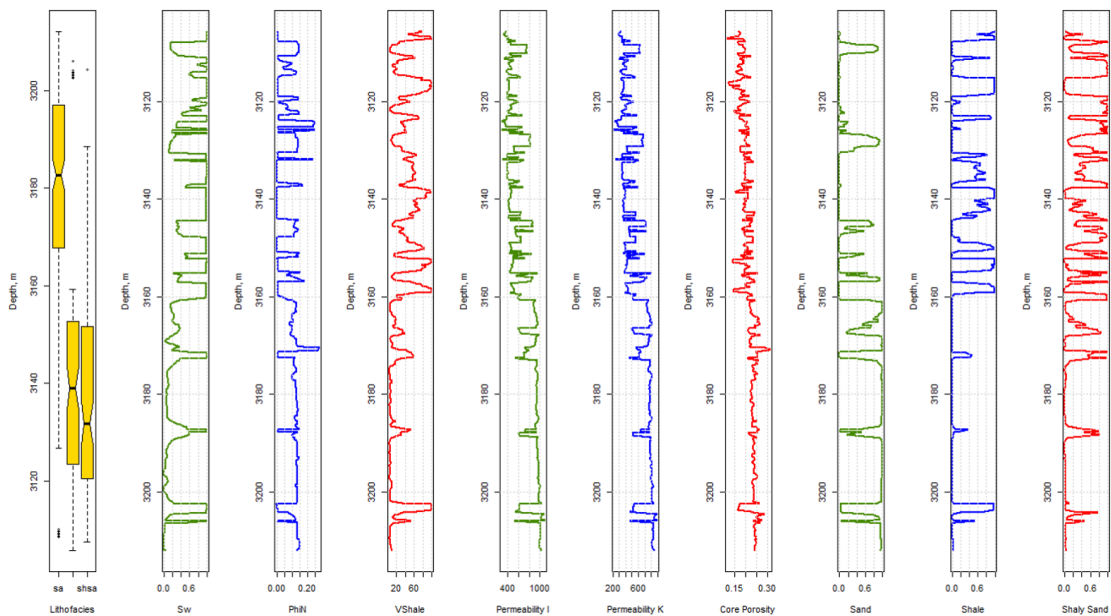


Figure B18: Well Logs & Lithofacies and Petrophysical Prediction for Well PR-244

## C. Geostatistical Property Simulation

In Figure C1, the three top left plots represent the indicator variogram for the Sand, Shaly Sand, and Shale in horizontal direction. The three top right plots represent the indicator variogram for the Sand, Shaly Sand, and Shale in  $45^\circ$  direction. The three bottom left plots represent the indicator variogram for the Sand, Shaly Sand, and Shale in vertical direction. The three bottom right plots represent the indicator variogram for the Sand, Shaly Sand, and Shale in  $135^\circ$  direction.

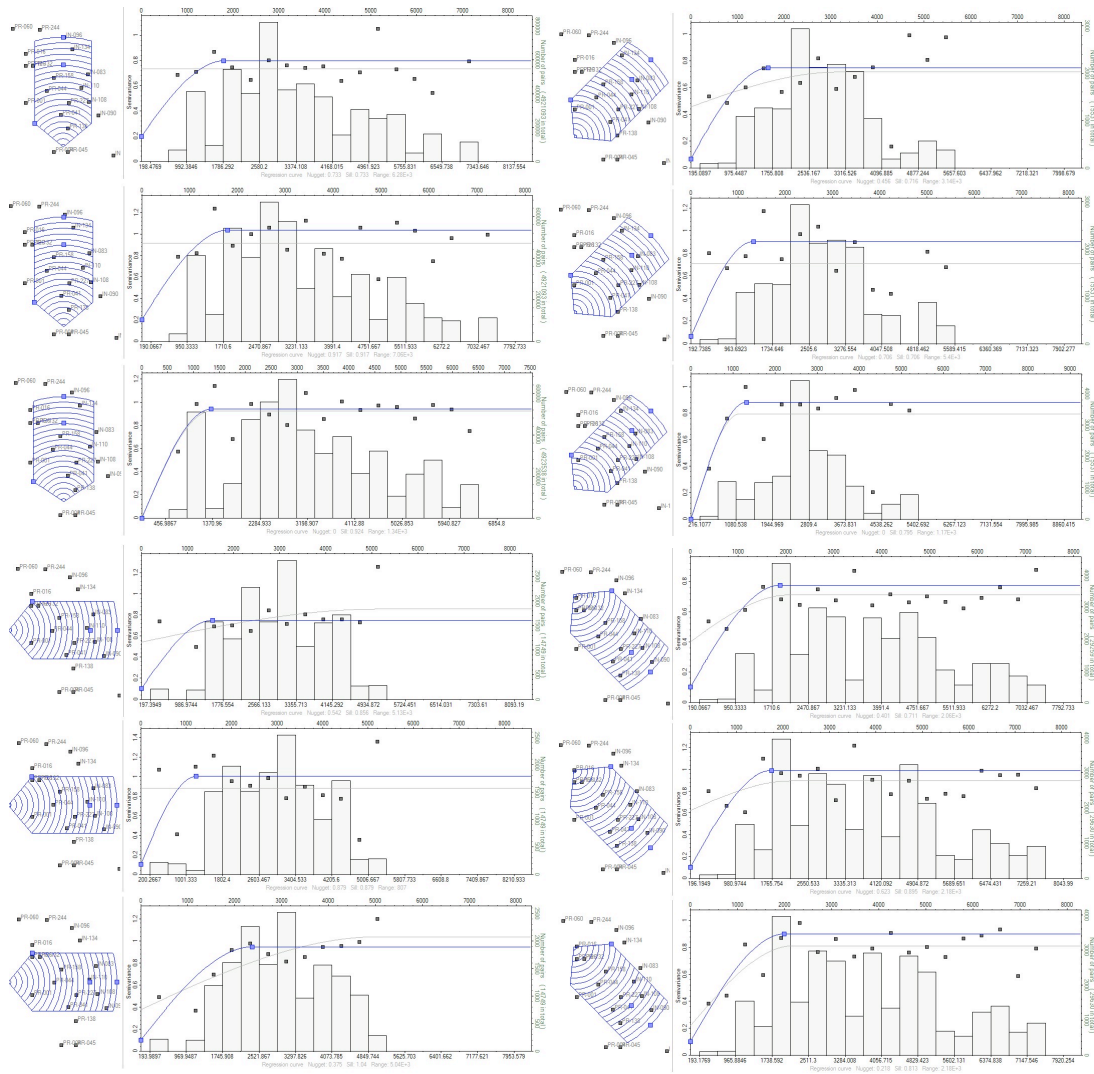


Figure C1: Indicator Variogram for the three Facies in Four Directions.

In Figure C2, The three top left plots represent the porosity variogram given the facies in horizontal direction. The three top right plots represent the porosity variogram given the facies in  $45^\circ$  direction. The three bottom left plots represent the porosity variogram given the facies in vertical direction. The three bottom right plots represent the porosity variogram given the facies in  $135^\circ$  direction.

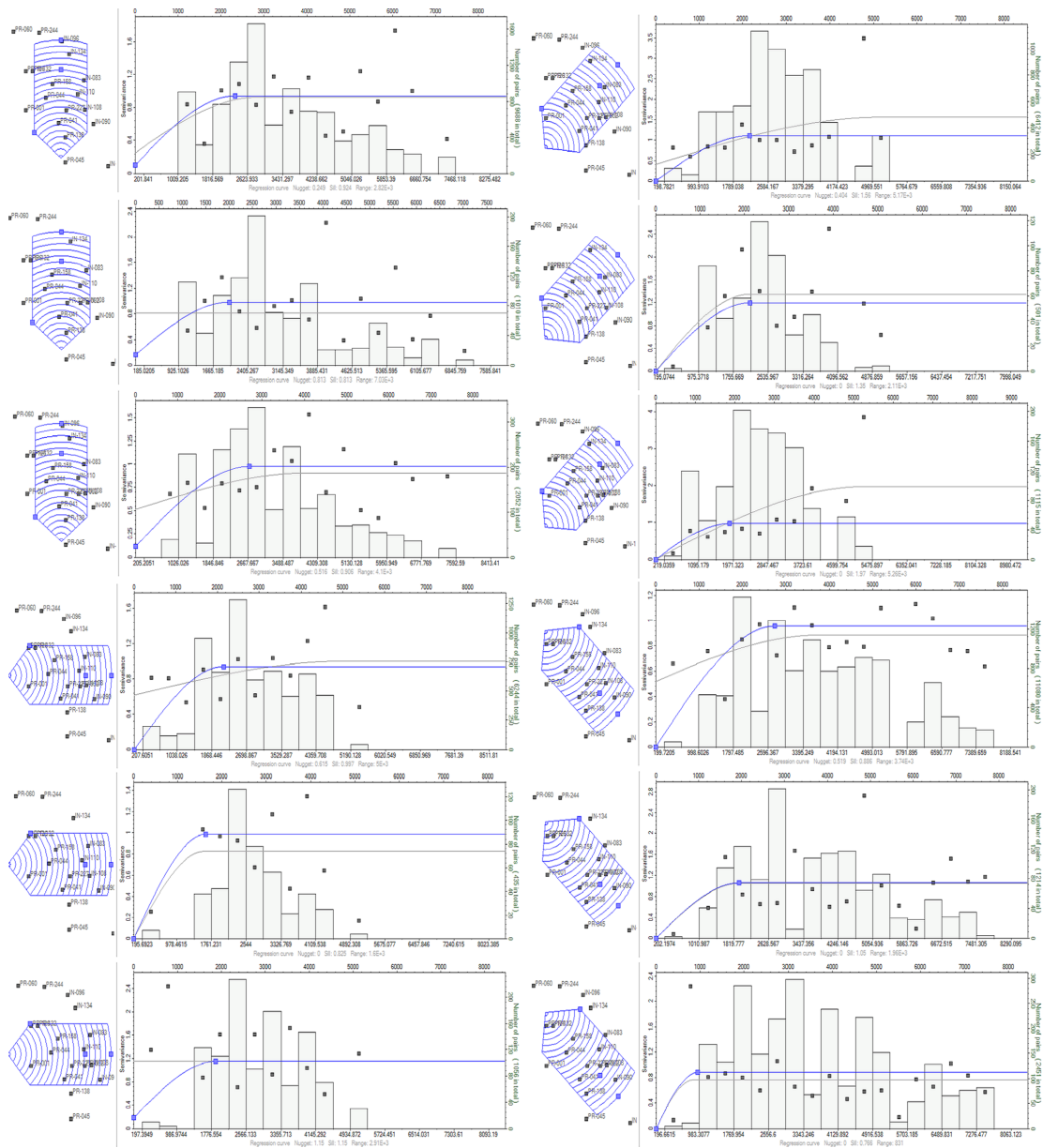


Figure C2: Porosity Variogram Given the Facies in Four Directions.



In Figure C3, The three top left plots represent the horizontal permeability variogram given the facies in horizontal direction. The three top right plots represent the horizontal permeability variogram given the facies in  $45^\circ$  direction. The three bottom left plots represent the horizontal permeability variogram given the facies in vertical direction. The three bottom right plots represent the horizontal permeability variogram given the facies in  $135^\circ$  direction.

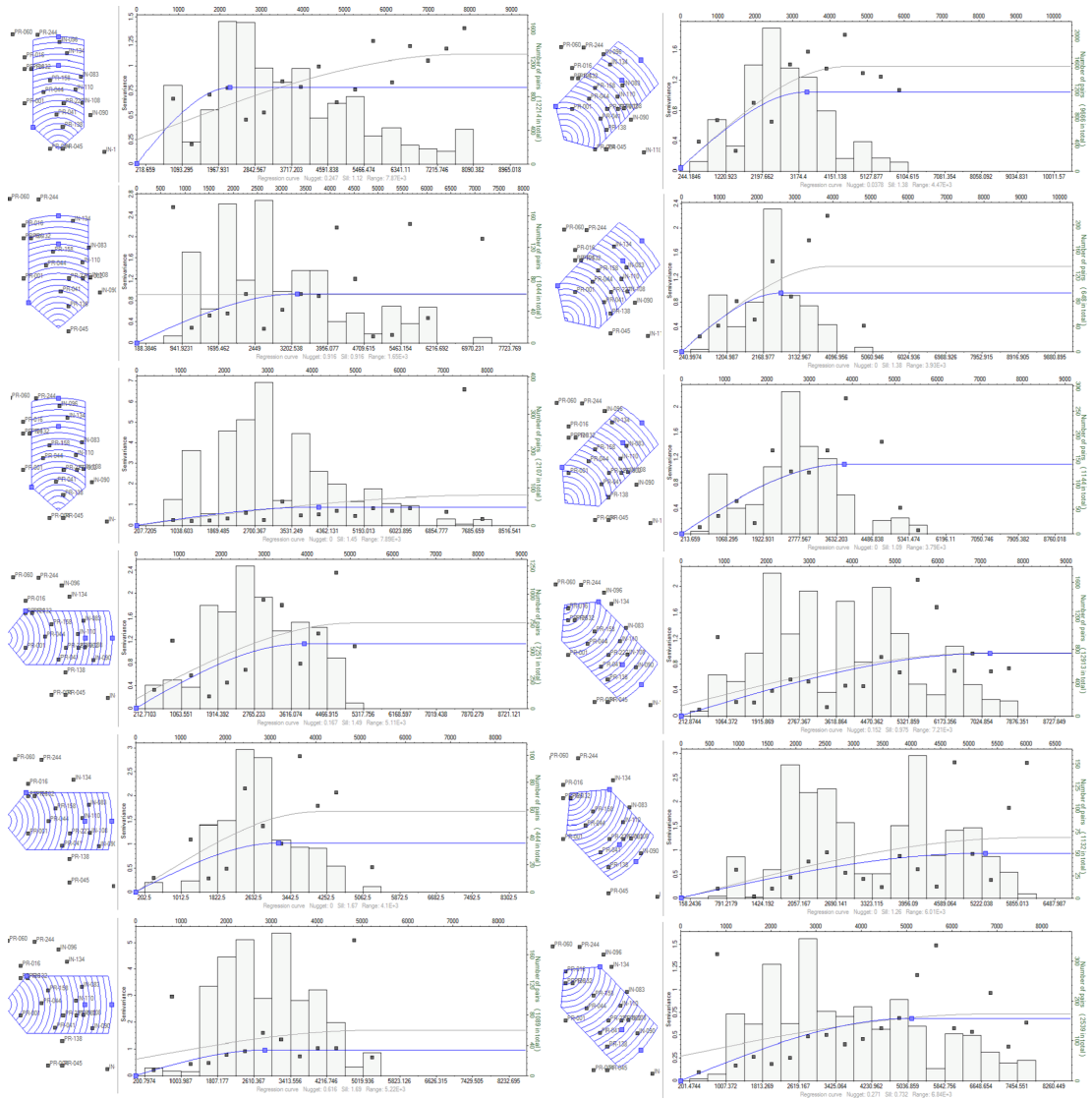


Figure C3: Horizontal Permeability Variogram Given the Facies in Four Directions

In Figure C4, The three top left plots represent the vertical permeability variogram given the facies in horizontal direction. The three top right plots represent the vertical permeability variogram given the facies in  $45^\circ$  direction. The three bottom left plots represent the vertical permeability variogram given the facies in vertical direction. The three bottom right plots represent the vertical permeability variogram given the facies in  $135^\circ$  direction.

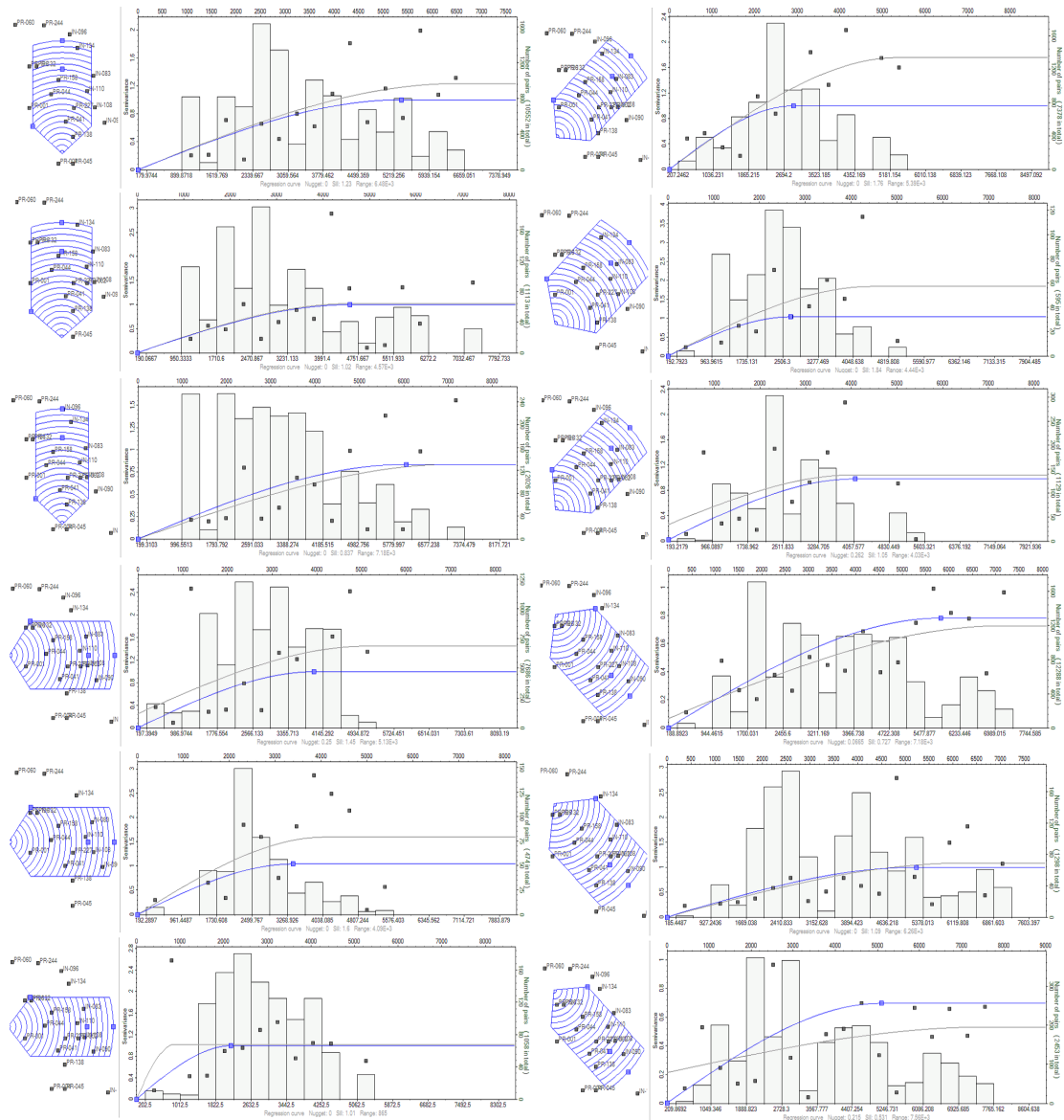


Figure C4: Vertical Permeability Variogram Given the Facies in Four Directions.

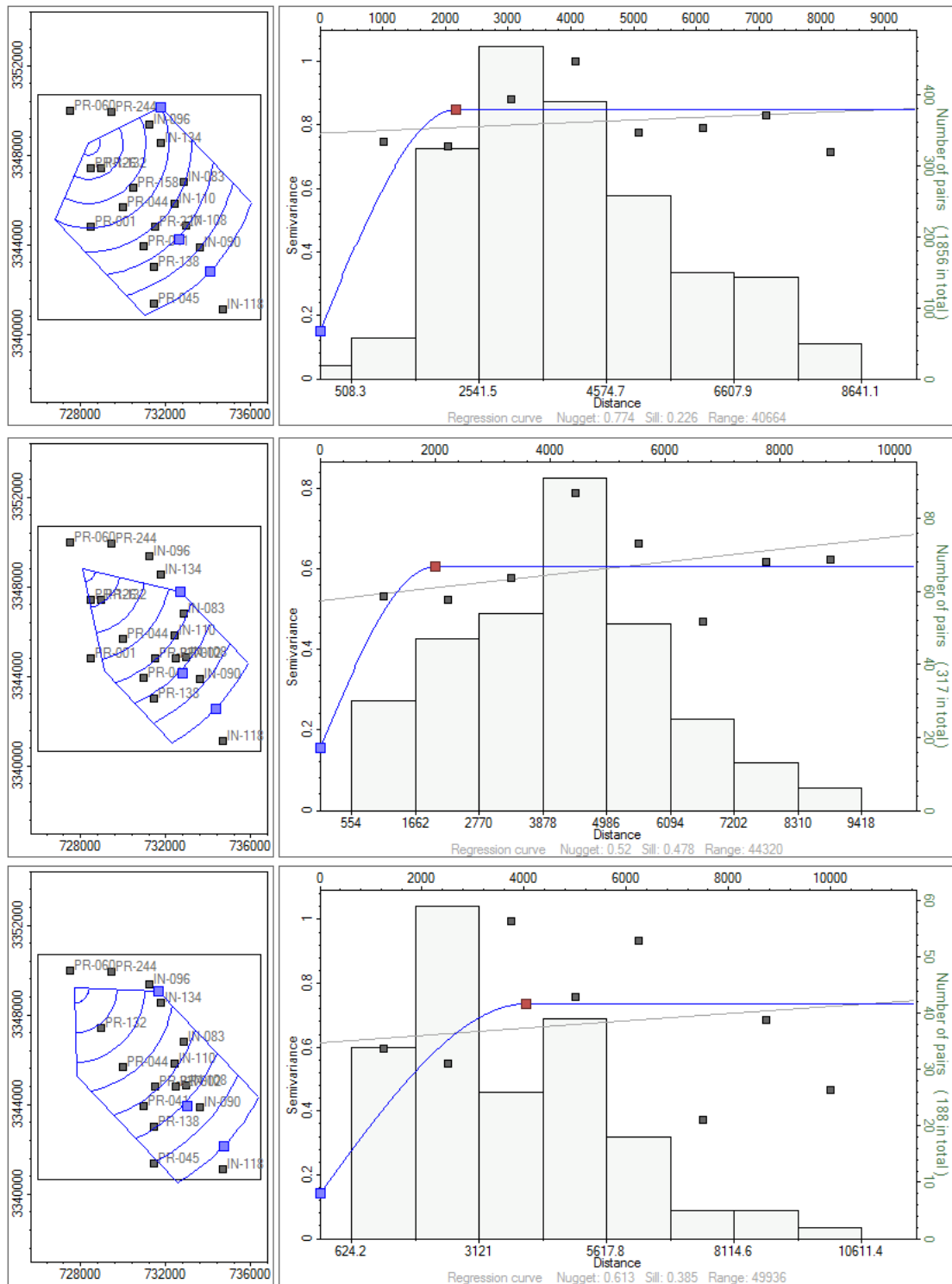


Figure C5: Log Porosity Variograms Given Sand (top), Shaly Sand (middle), and Shale (bottom)

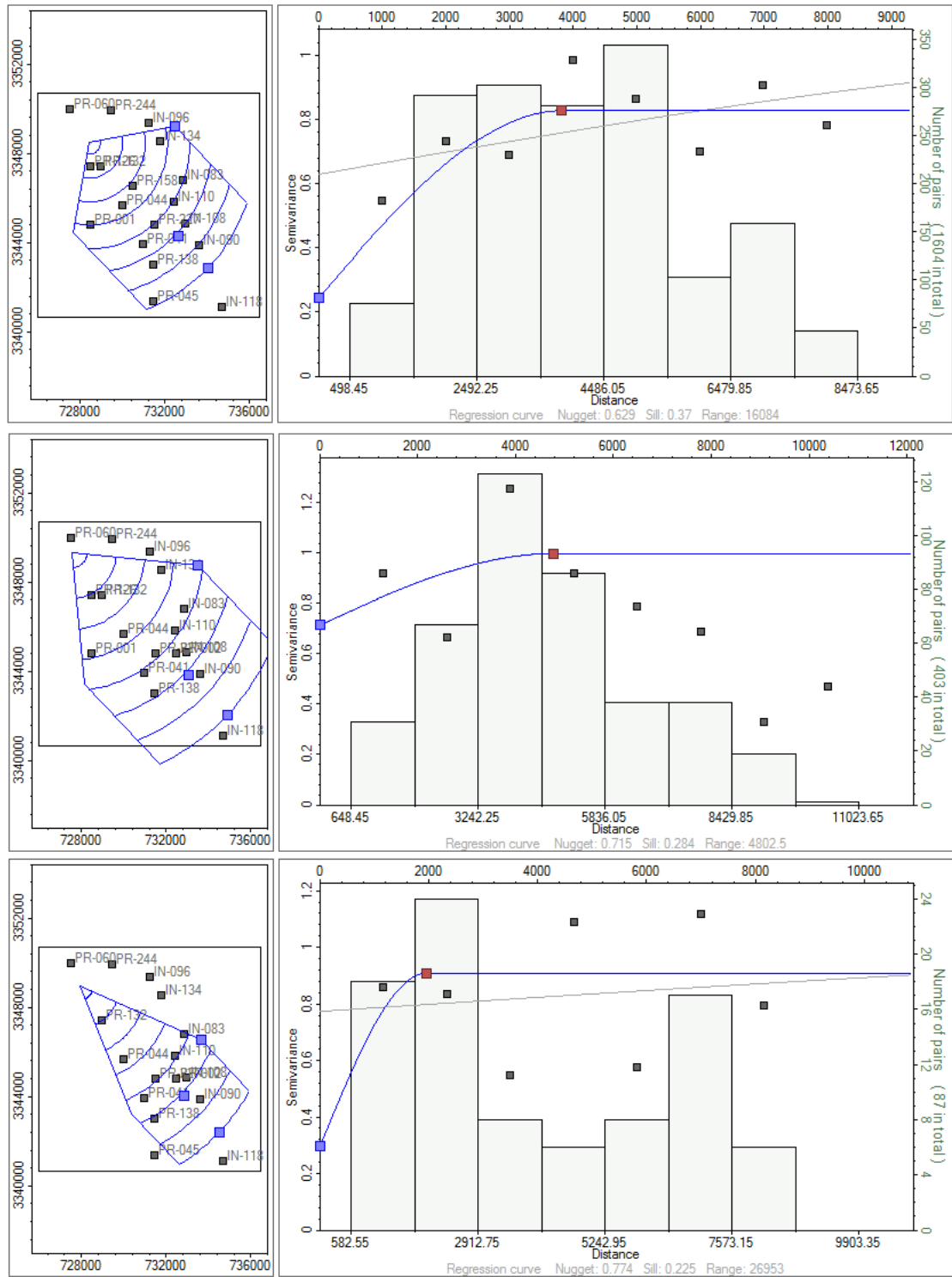


Figure C6: Horizontal Permeability Variograms Given Sand (top), Shaly Sand (middle), and Shale (bottom)

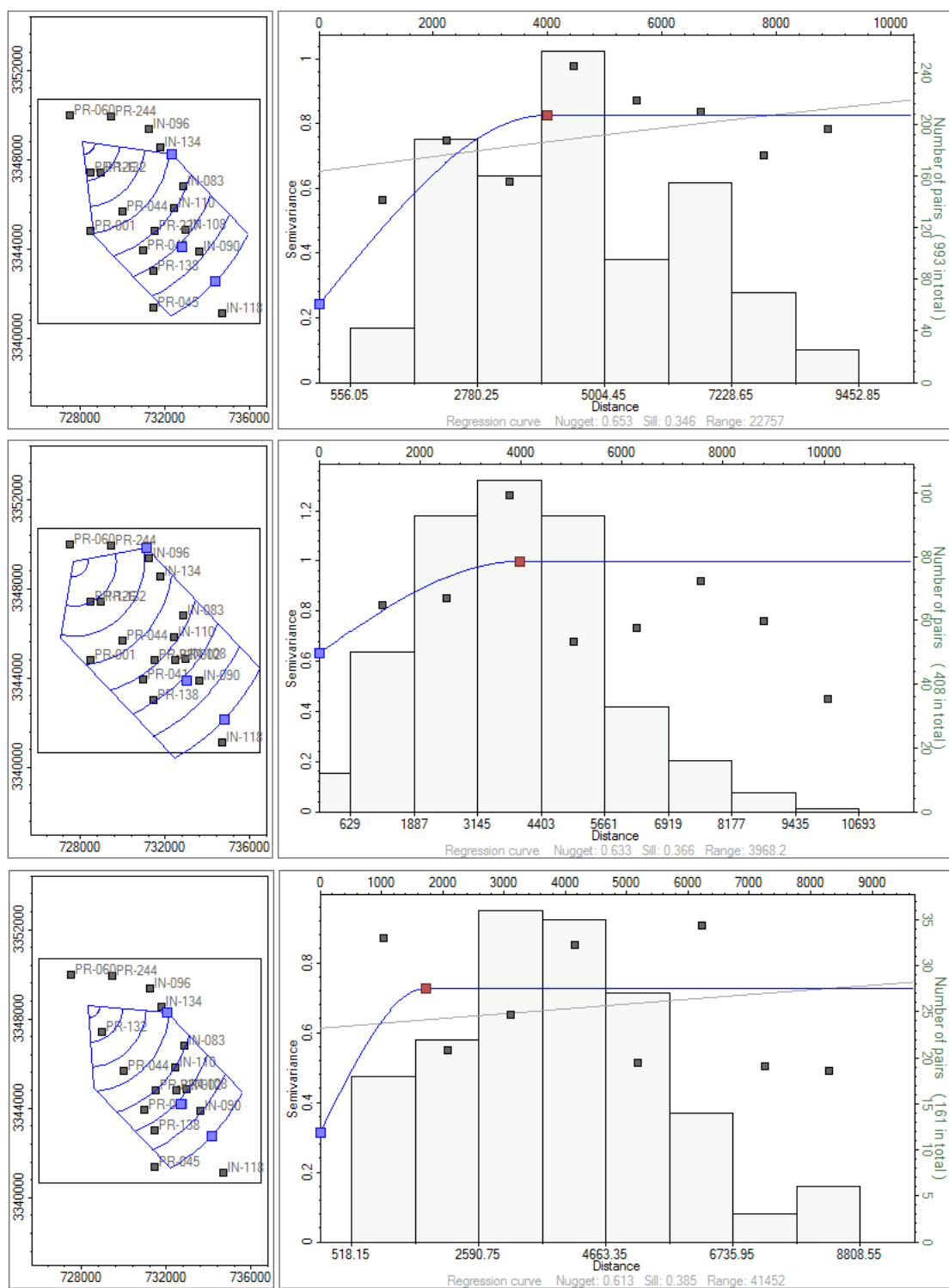


Figure C7: Vertical Permeability Variograms Given Sand (top), Shaly Sand (middle), and Shale (bottom)

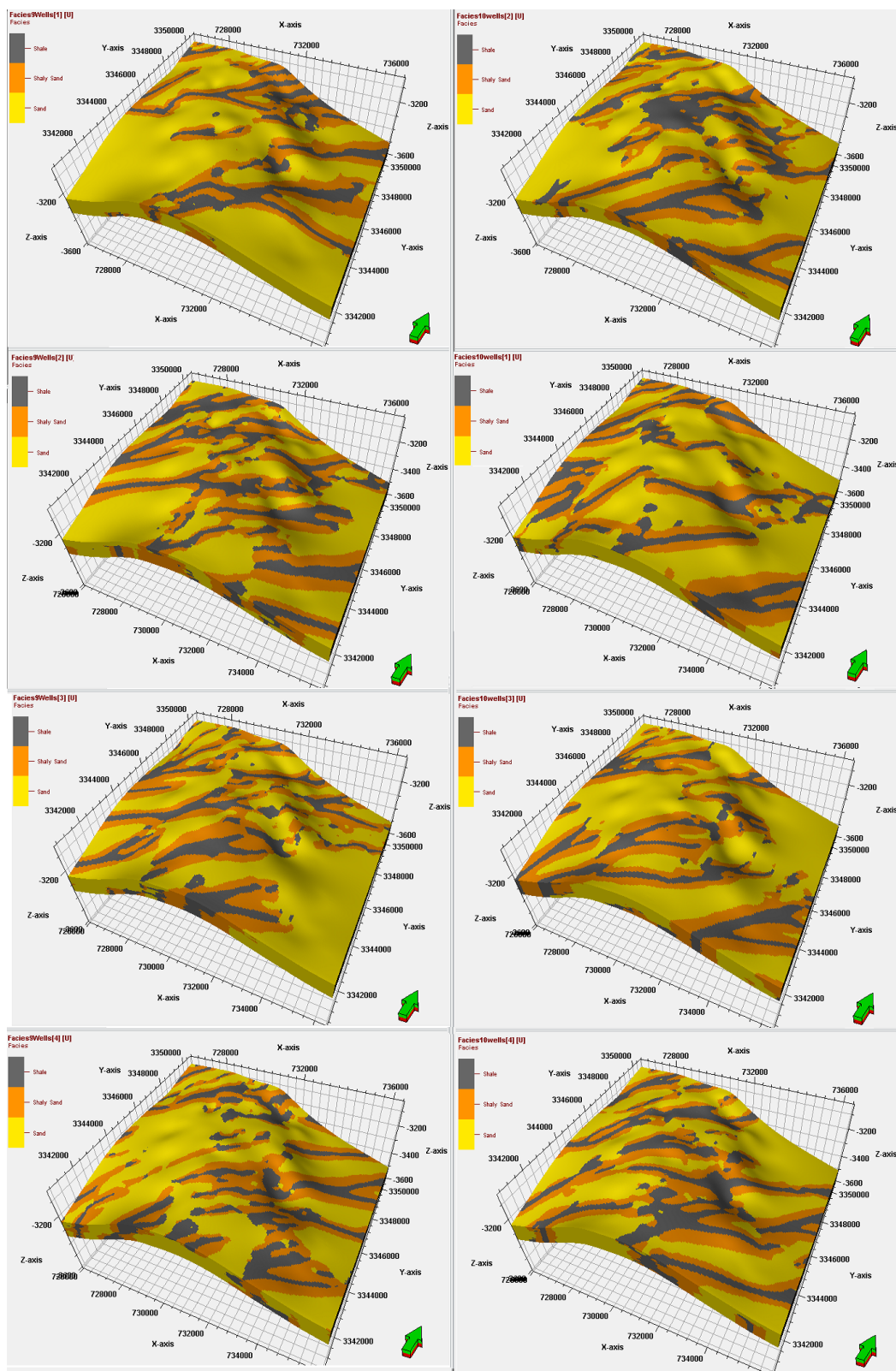


Figure C8: 1st Four Set of MPS Realizations given the Two Subset Well Groups: LHS: 9 Wells & RHS: 10 Wells



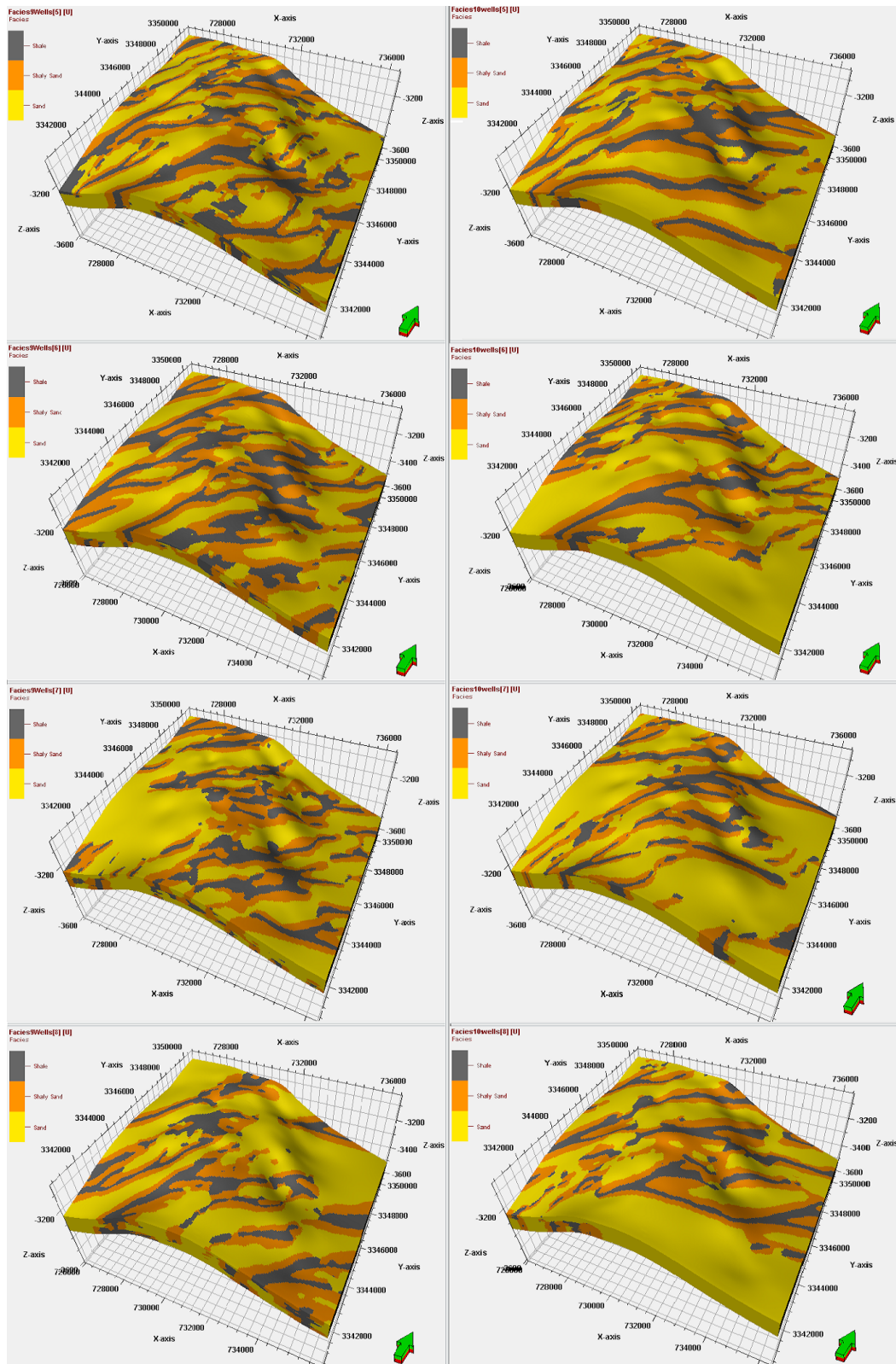


Figure C9: 2nd Four Set of MPS Realizations given the Two Subset Well Groups: LHS: 9 Wells & RHS: 10 Wells

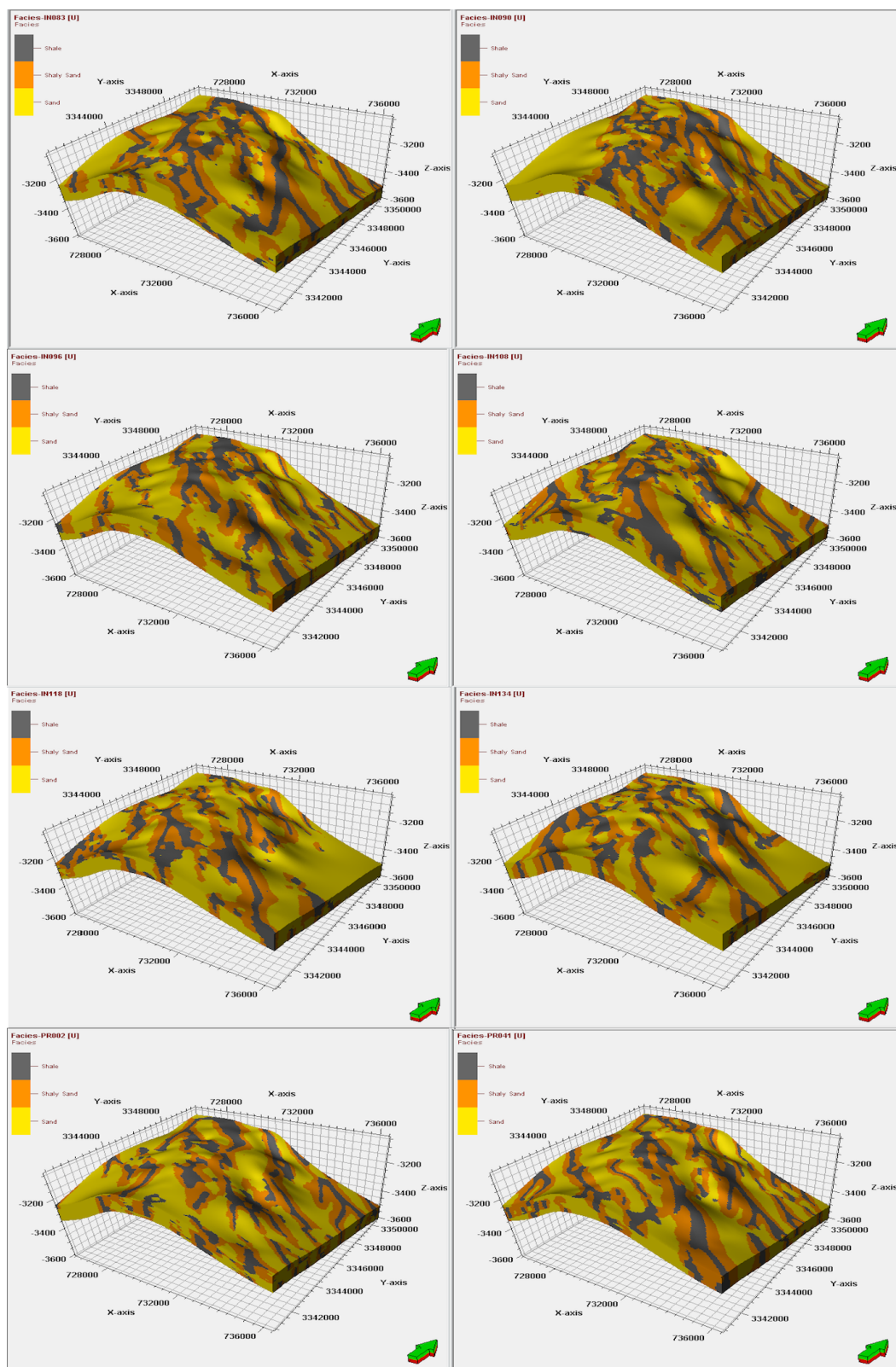


Figure C10: Leave-One-Out Cross Validation of MPS Facies Simulation-Wells Group1



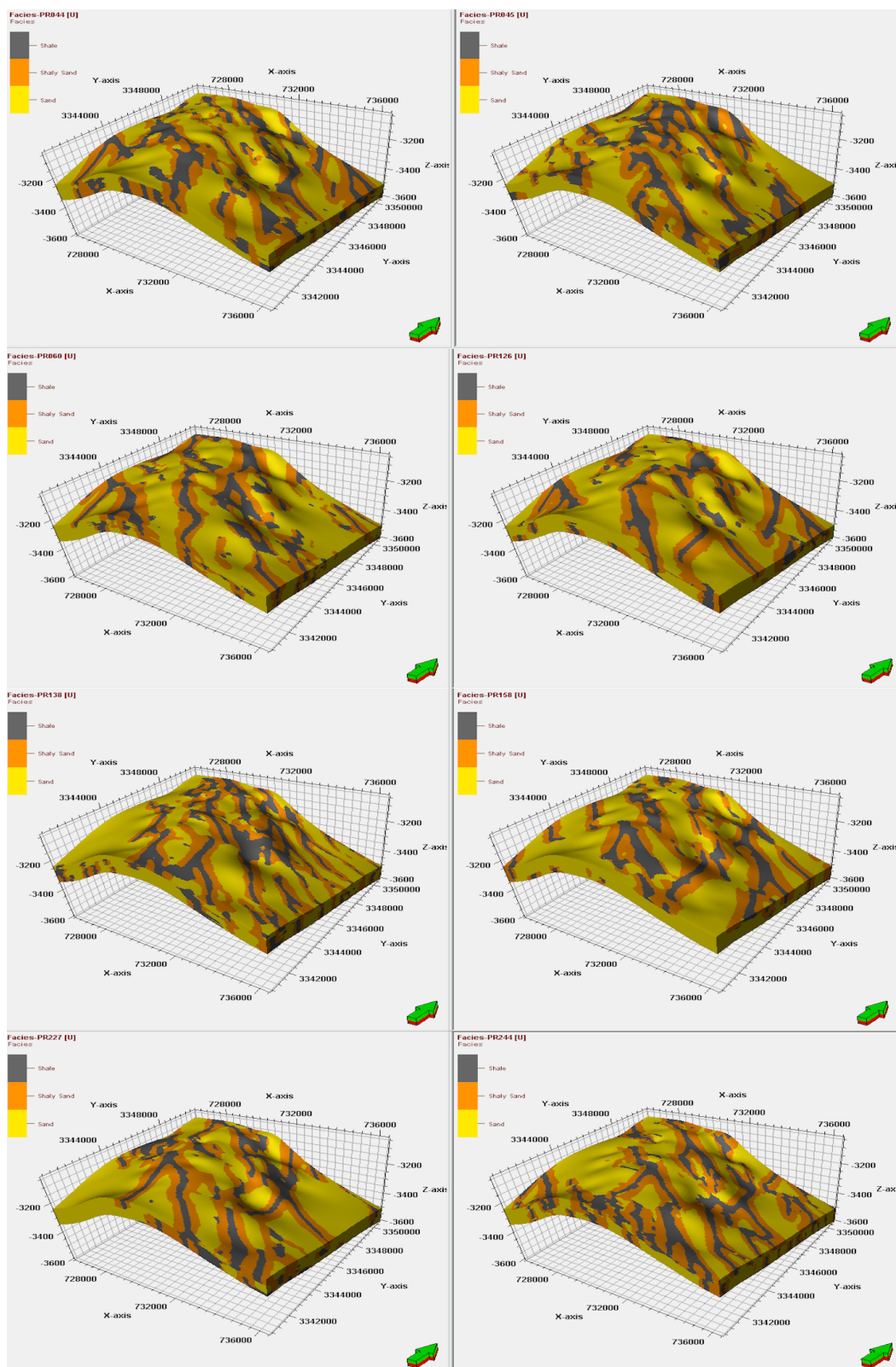


Figure C11: Leave-One-Out Cross Validation of MPS Facies Simulation-Wells Group2

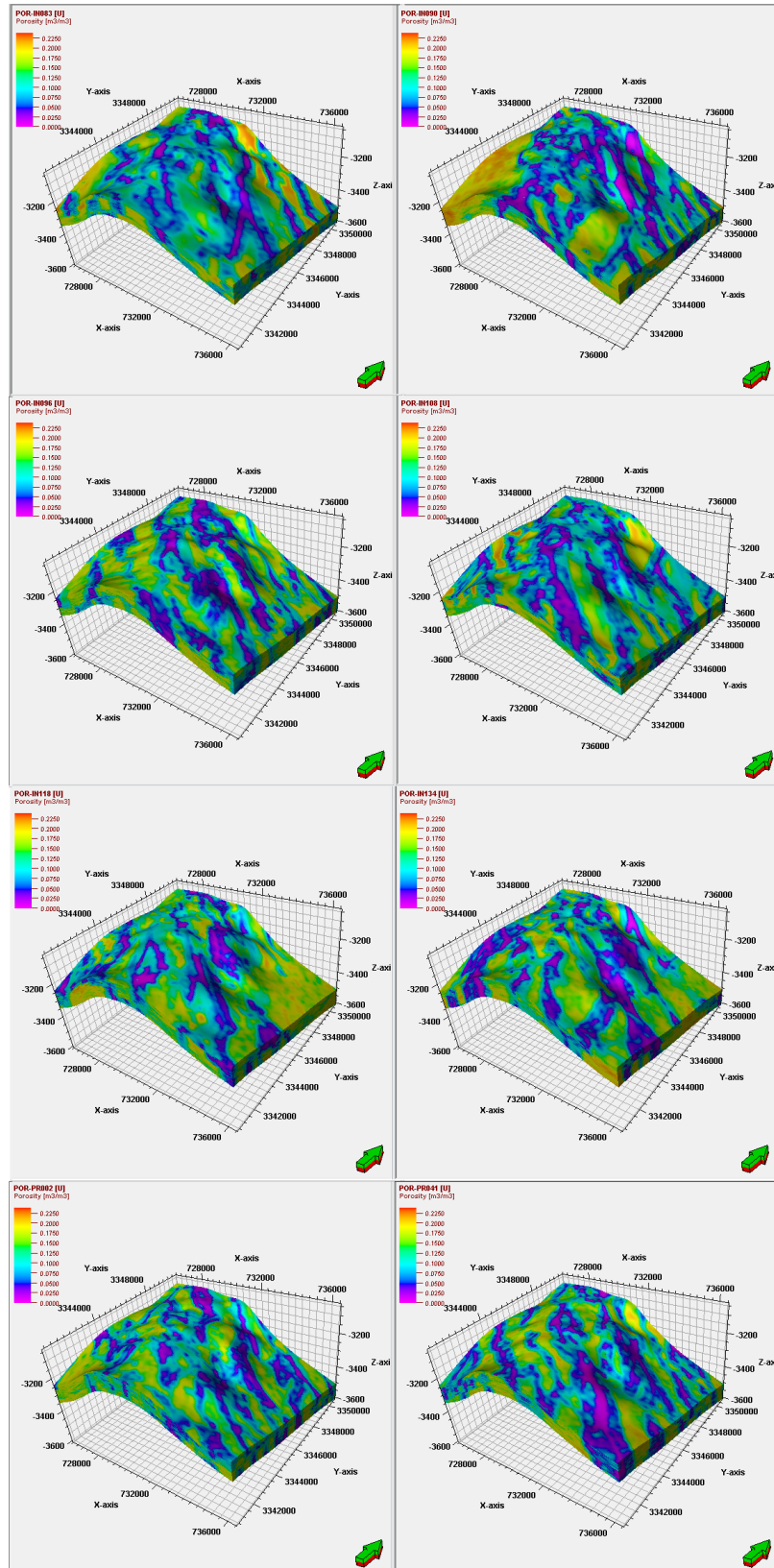


Figure C12: Leave-One-Out Cross Validation of Log Porosity Simulation-Wells Group1

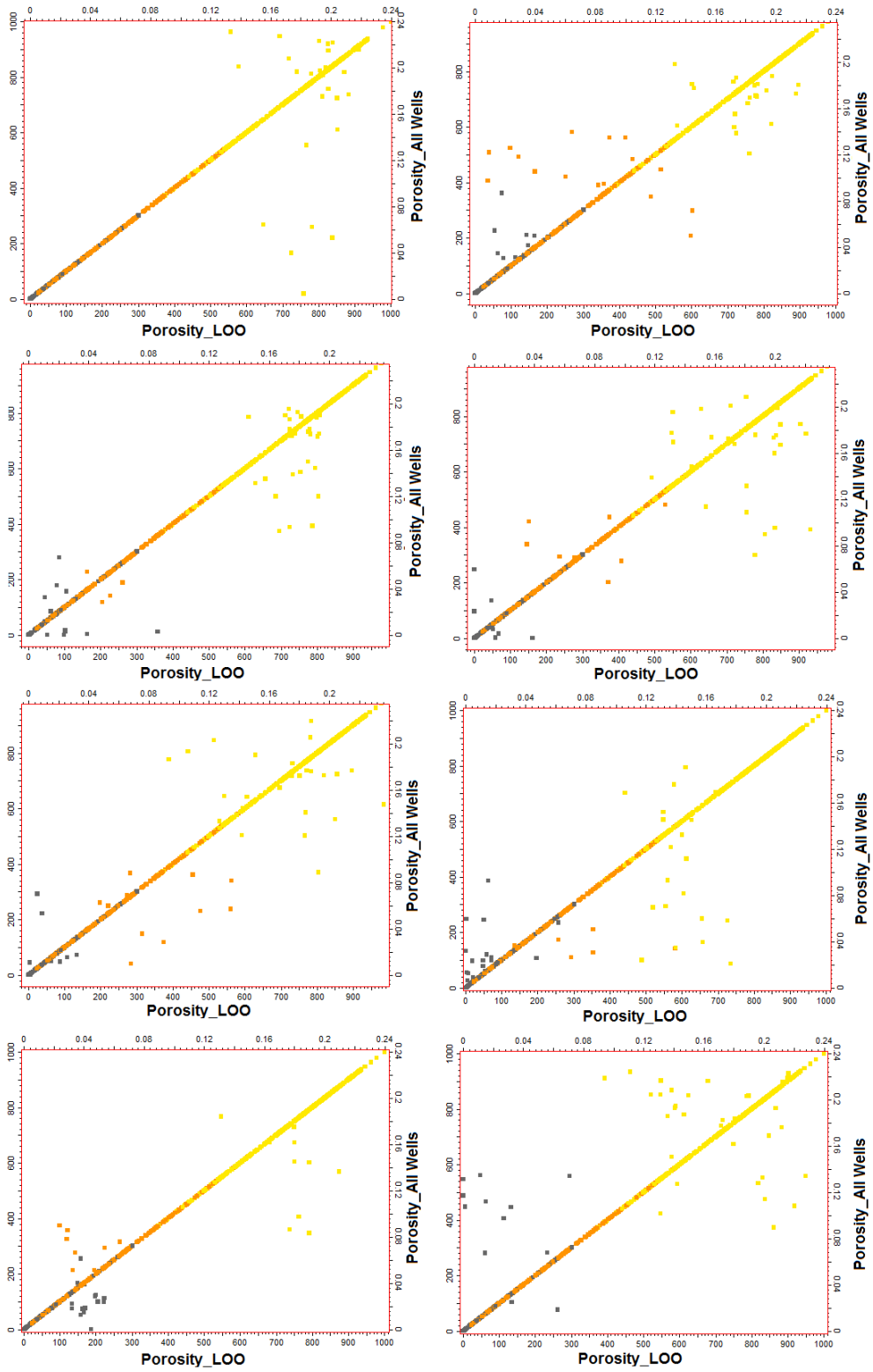


Figure C13: Comparison between Leave-One-Out Log Porosity and All Wells-Wells Group1

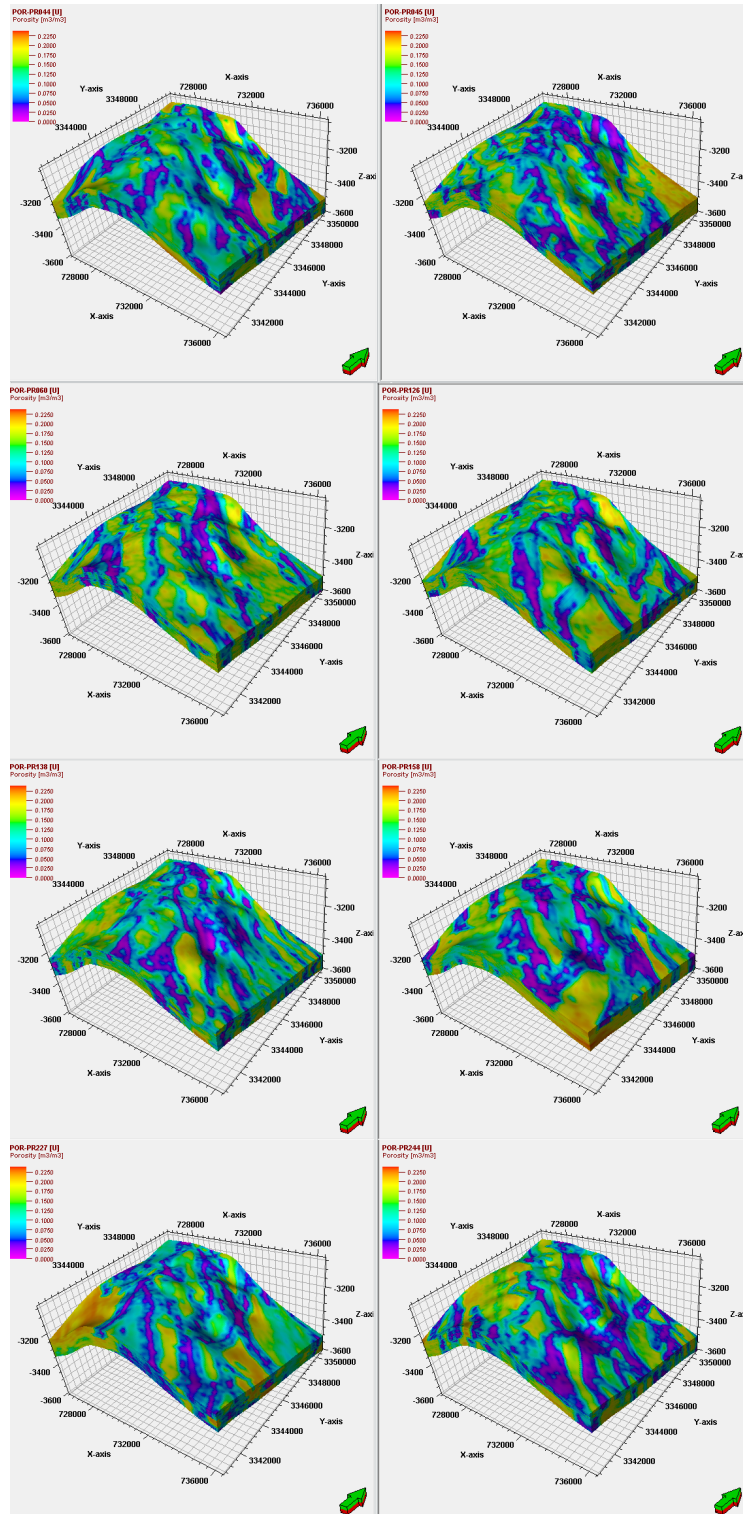


Figure C14: Leave-One-Out Cross Validation of Log Porosity Simulation-Wells Group2

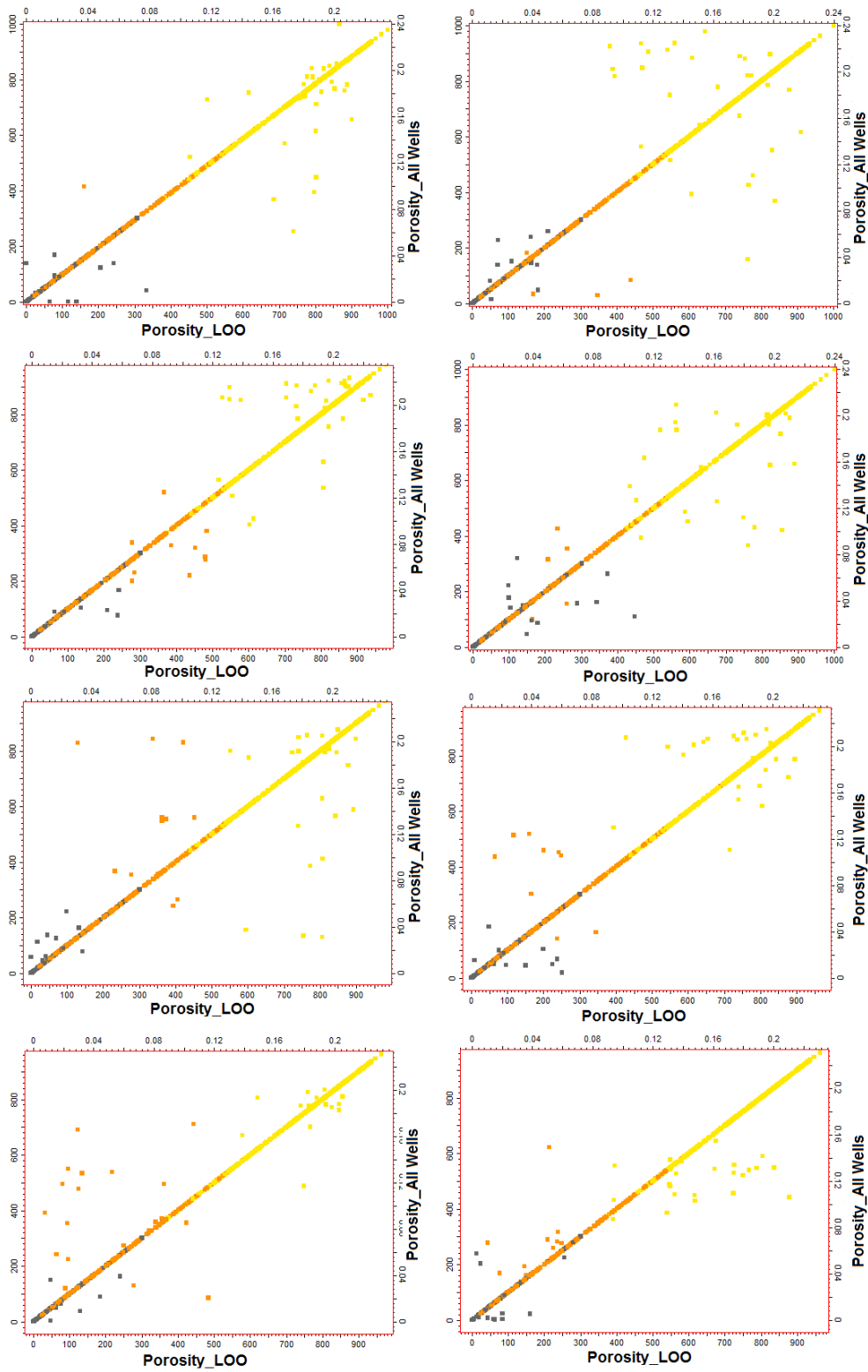


Figure C15: Comparison between Leave-One-Out Log Porosity and All Wells-Wells Group2



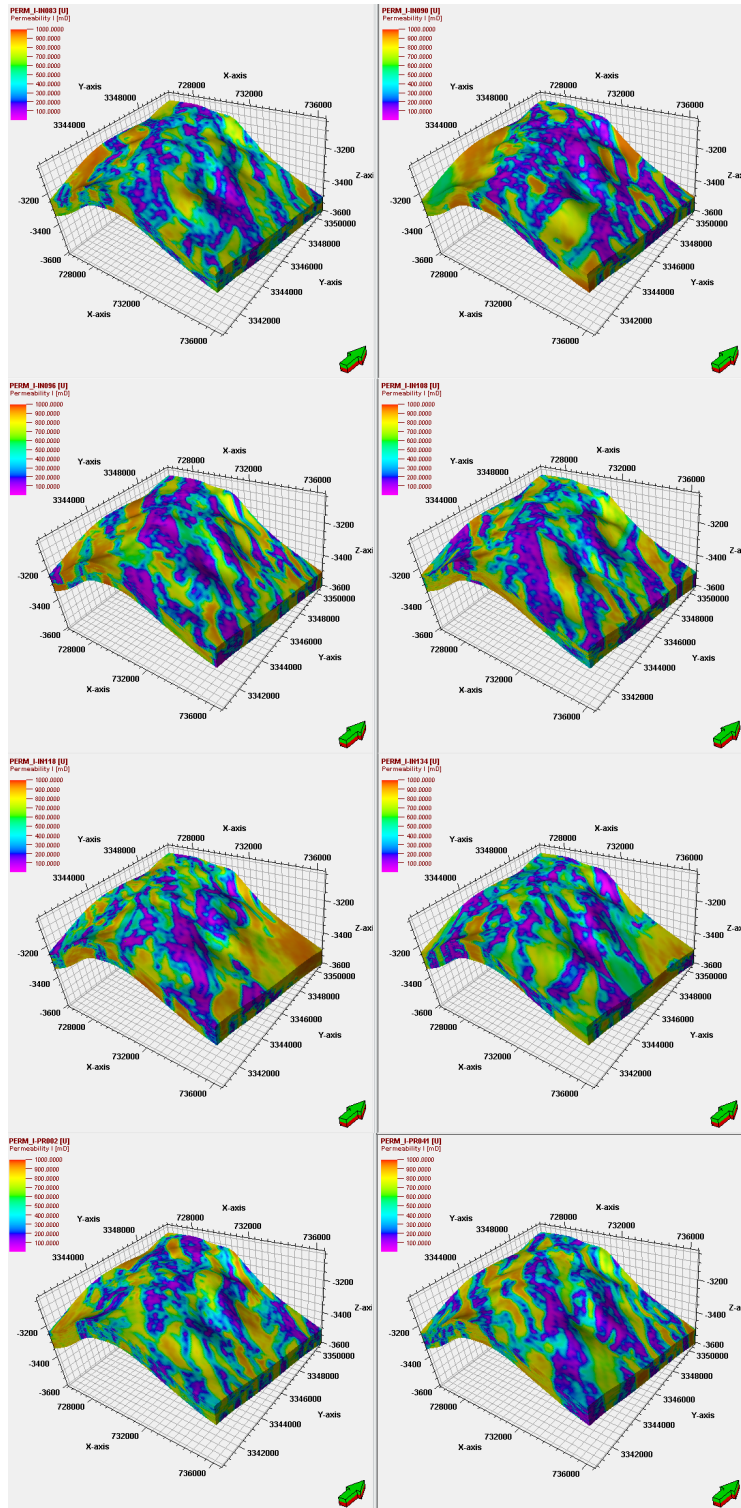


Figure C16: Leave-One-Out Cross Validation of Horizontal Permeability Simulation-Wells Group1

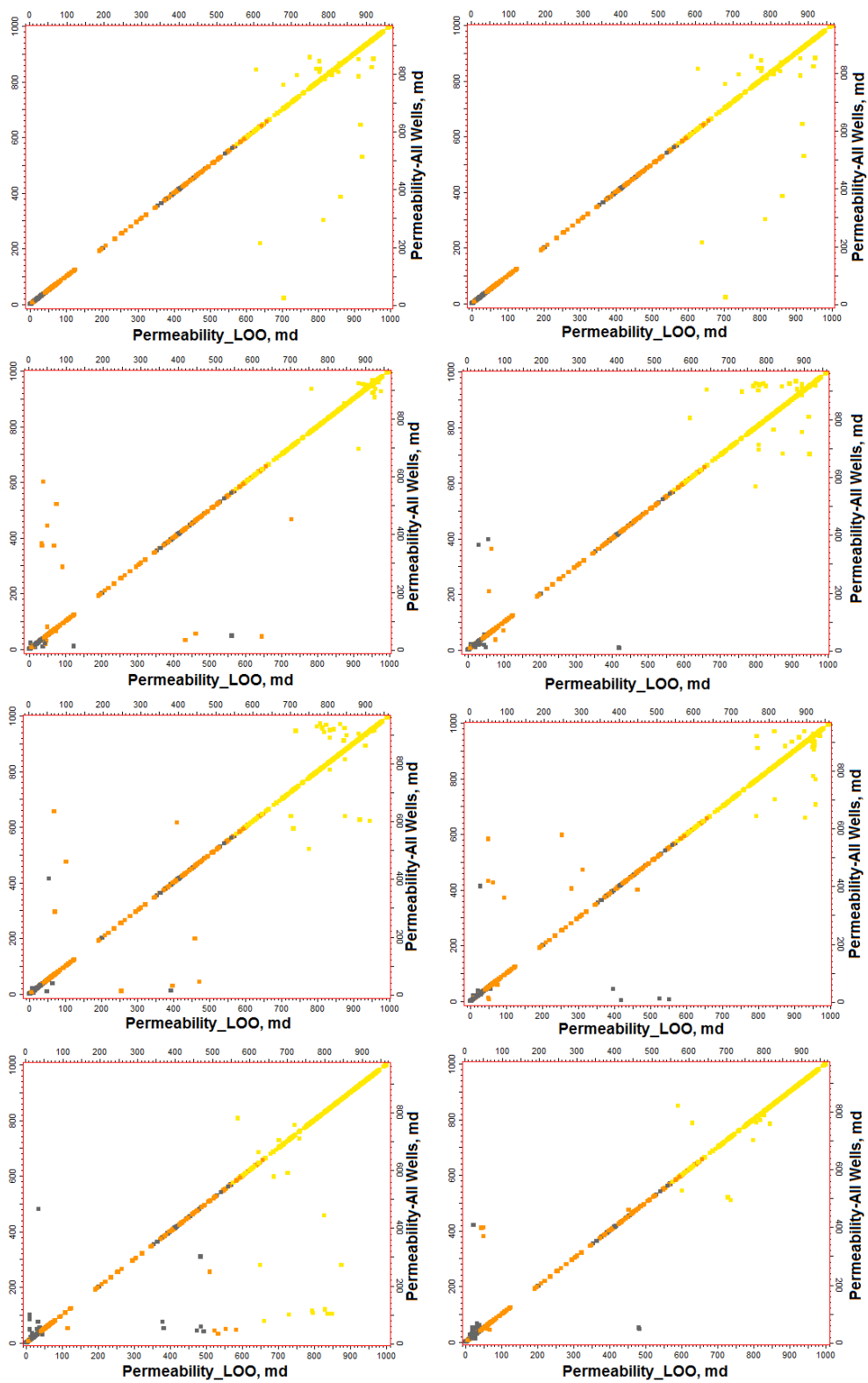


Figure C17: Comparison between Leave-One-Out Horizontal Permeability and All Wells- Wells Group1

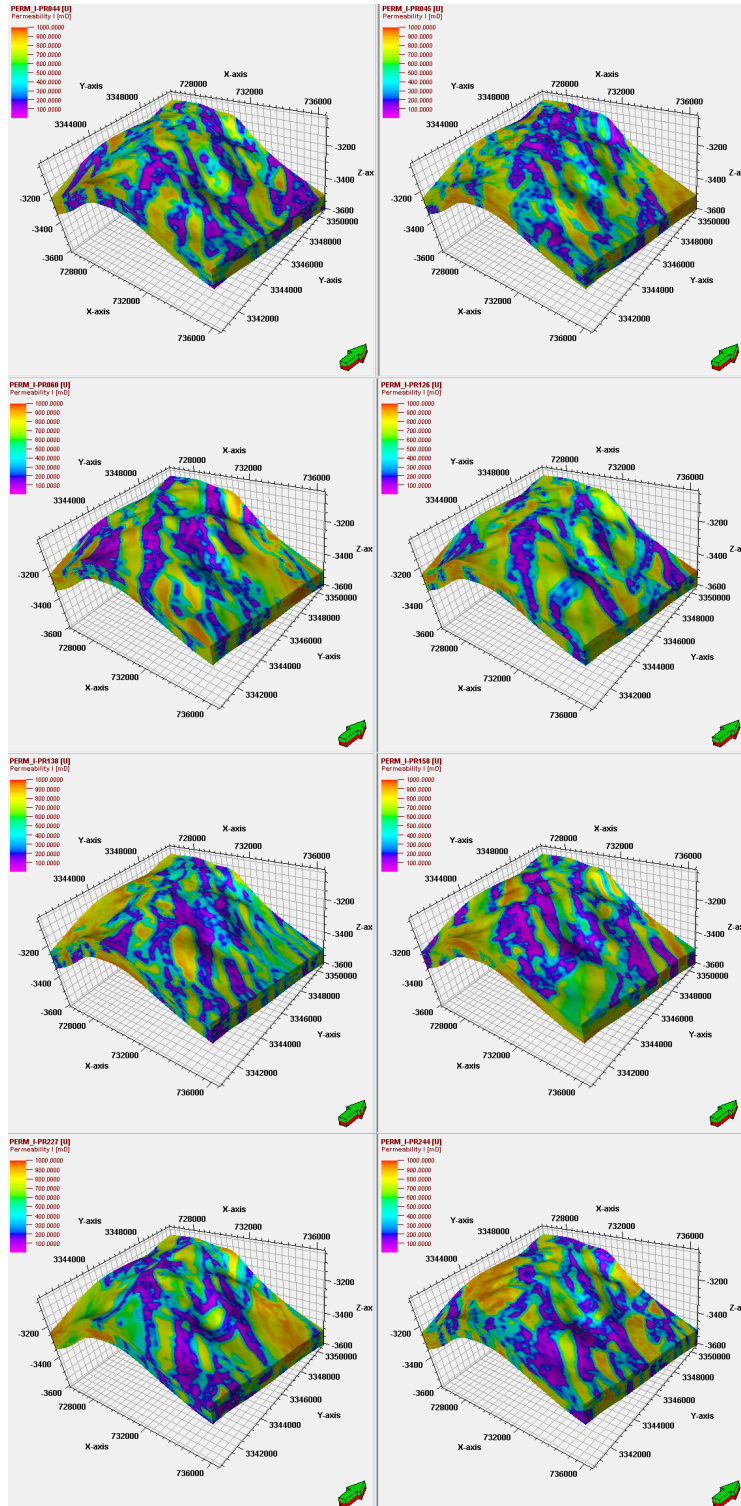


Figure C18: Leave-One-Out Cross Validation of Horizontal Permeability Simulation-Wells Group2



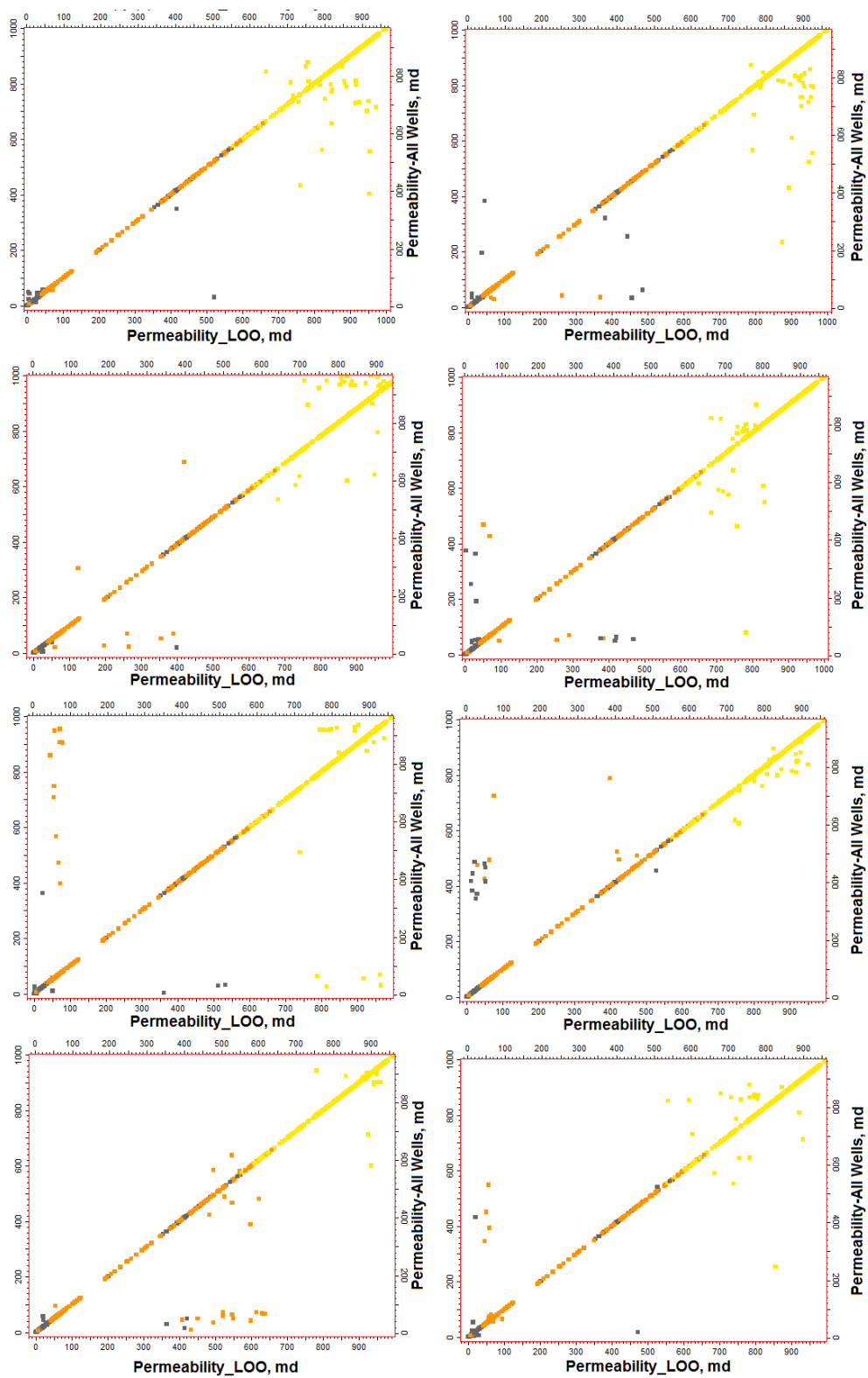


Figure C19: Comparison between Leave-One-Out Horizontal Permeability and All Wells- Wells Group2

## D. History Matching

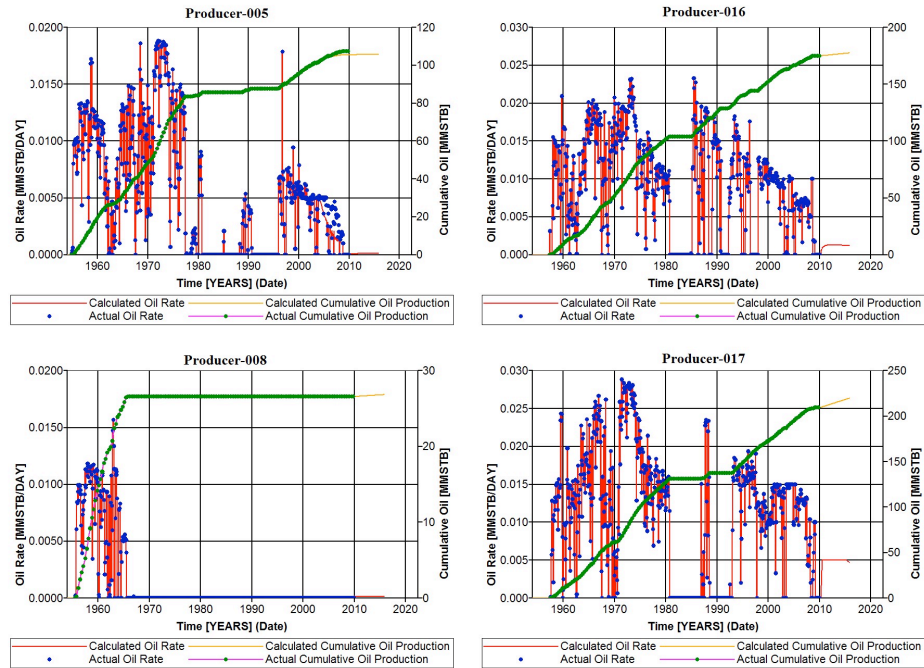


Figure D1: Wells Production History Matching

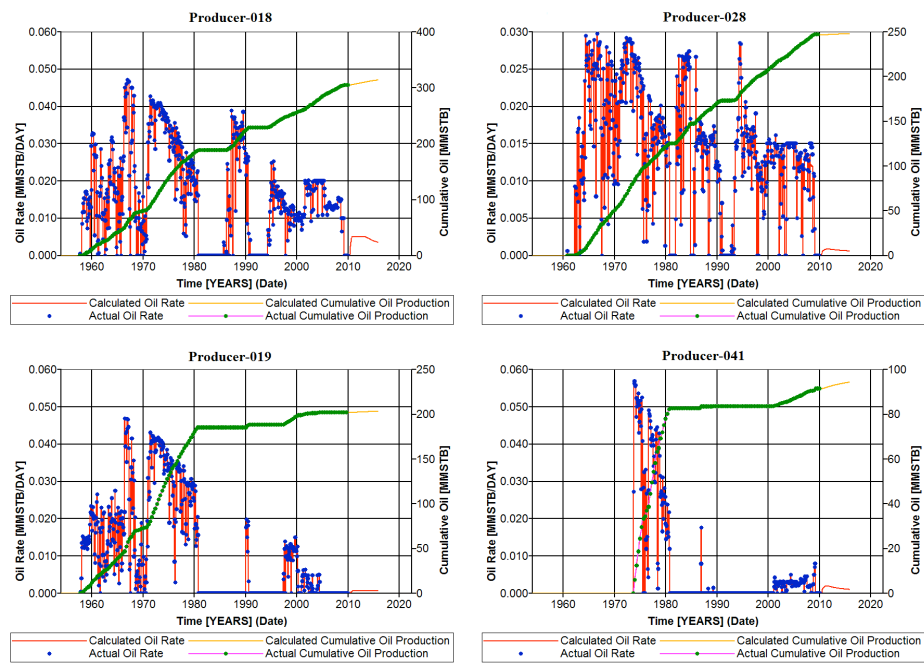


Figure D2: Wells Production History Matching

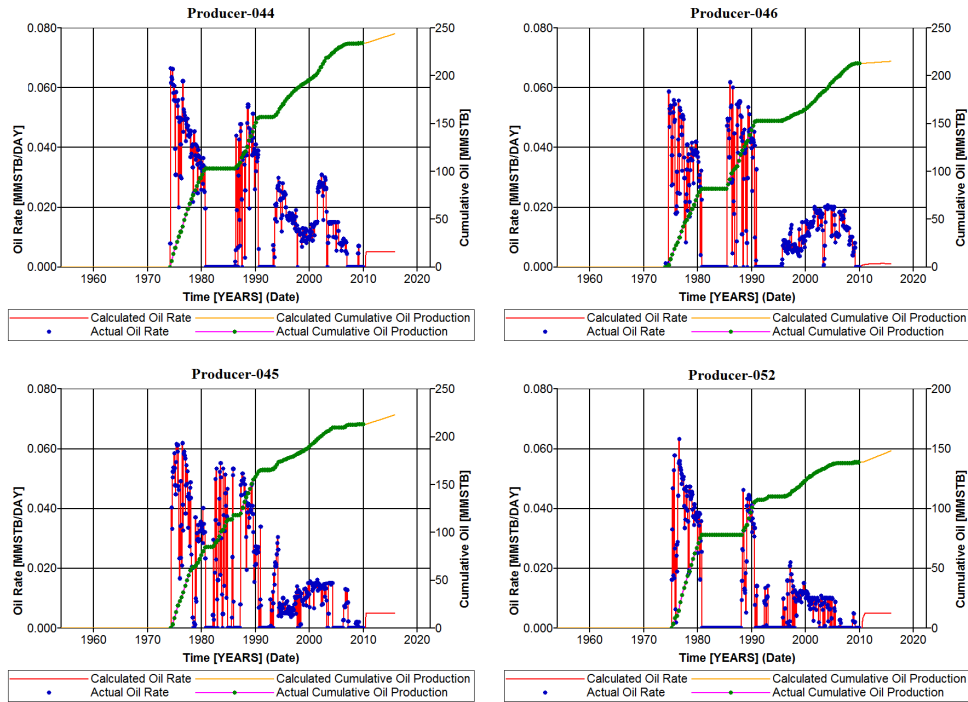


Figure D3: Wells Production History Matching

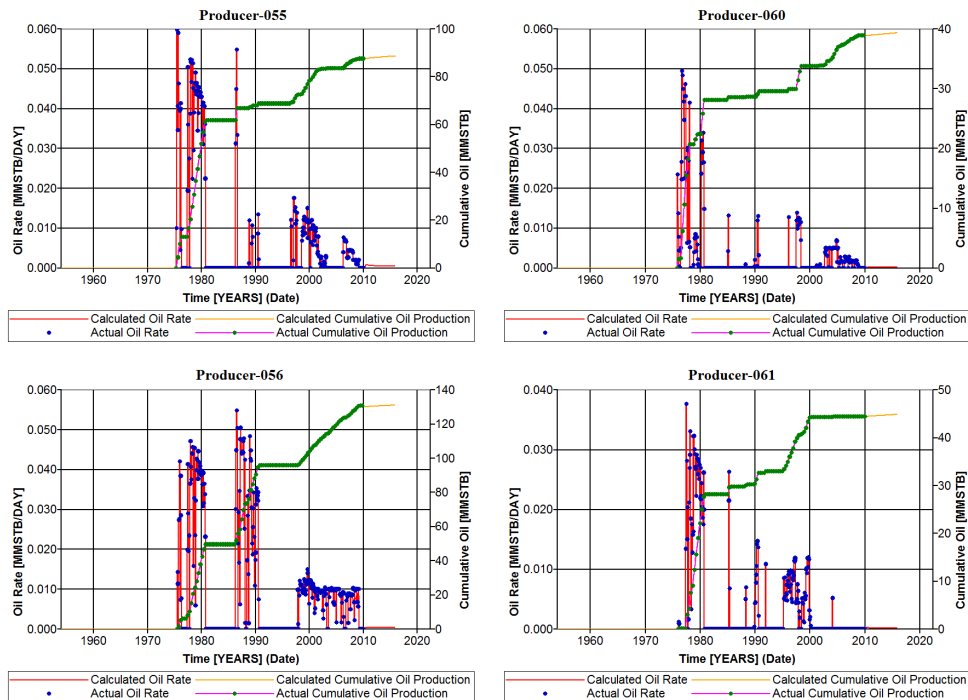


Figure D4: Wells Production History Matching

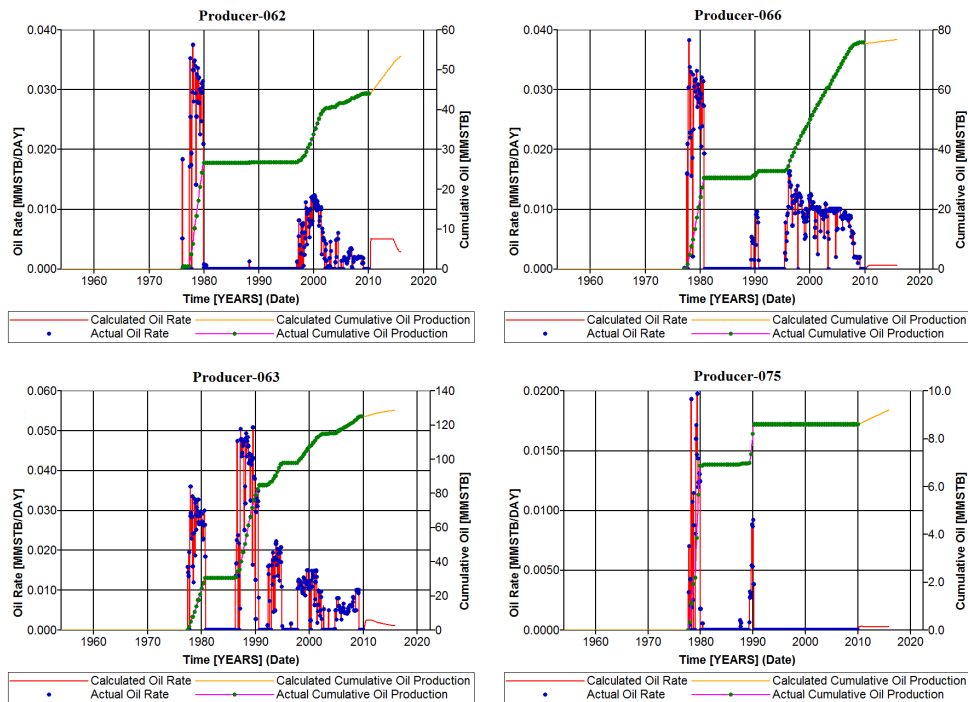


Figure D5: Wells Production History Matching

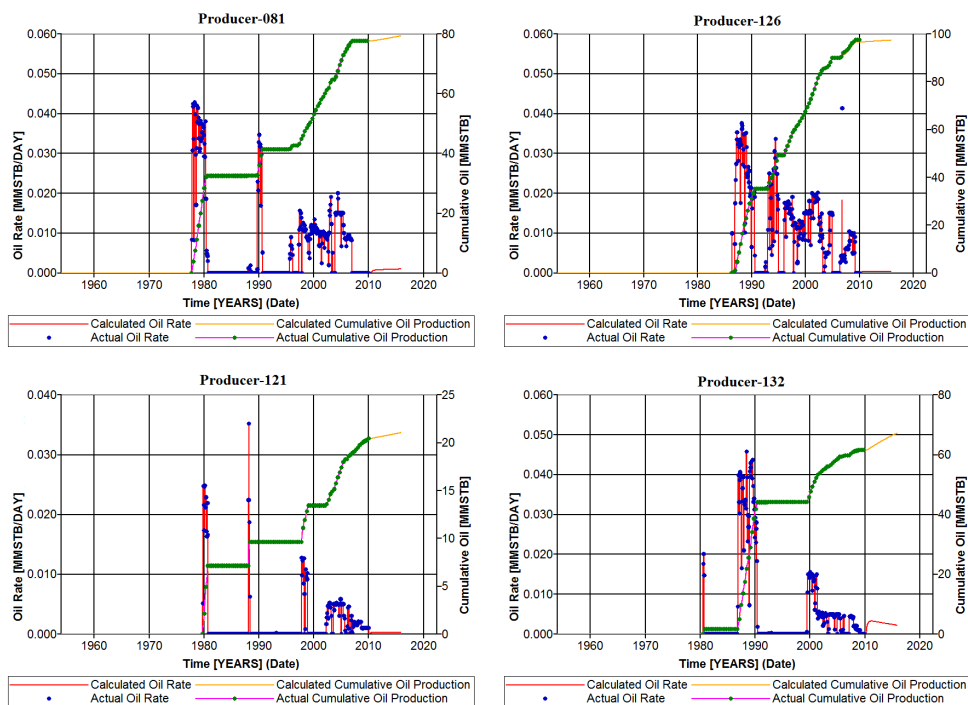


Figure D6: Wells Production History Matching

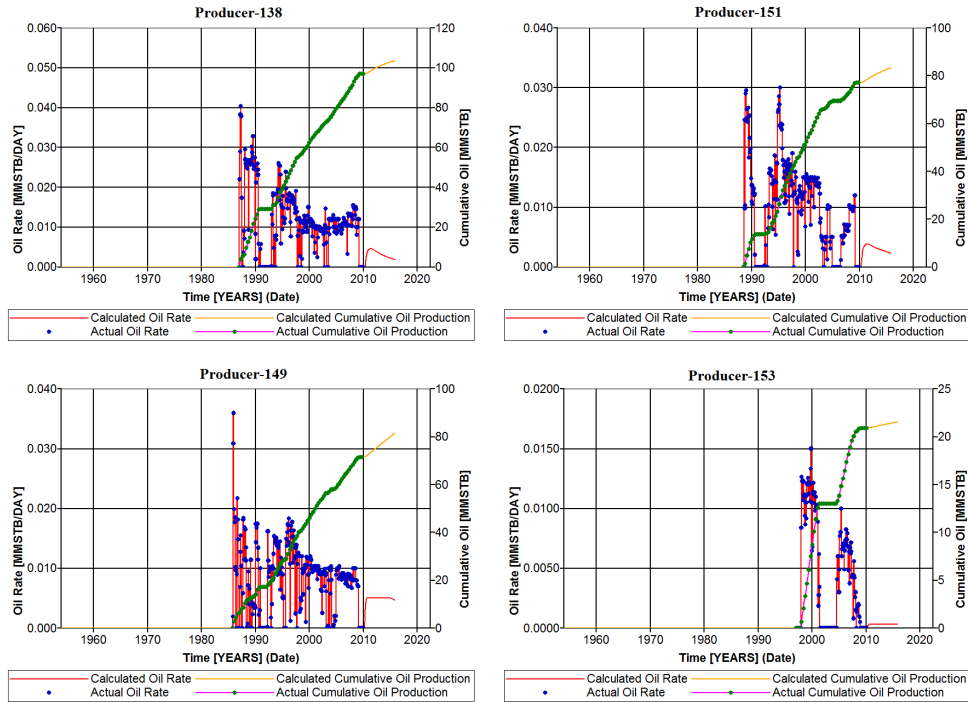


Figure D7: Wells Production History Matching

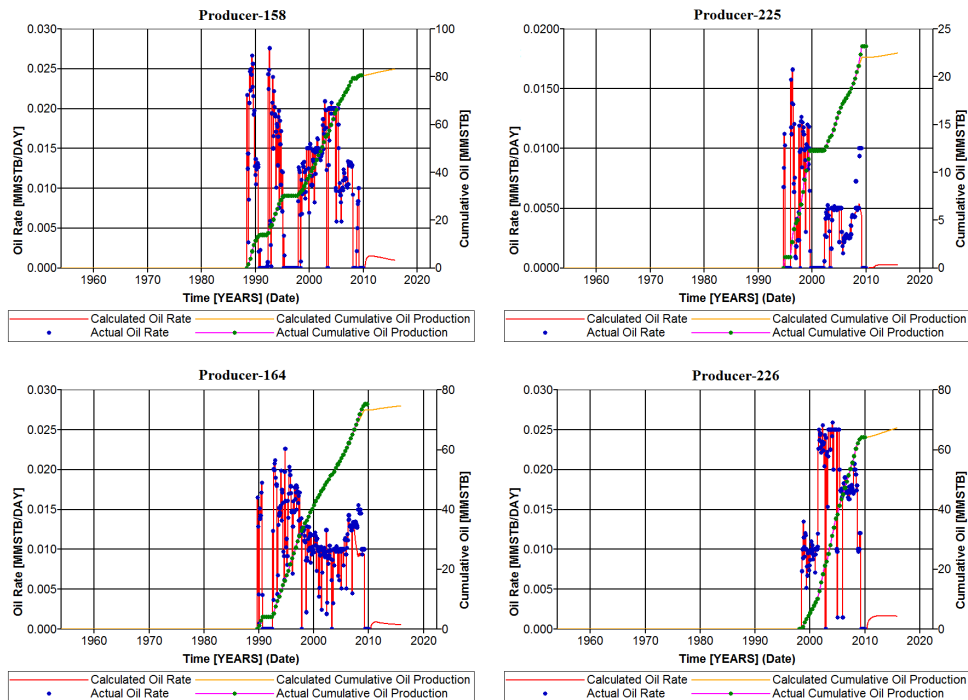


Figure D8: Wells Production History Matching

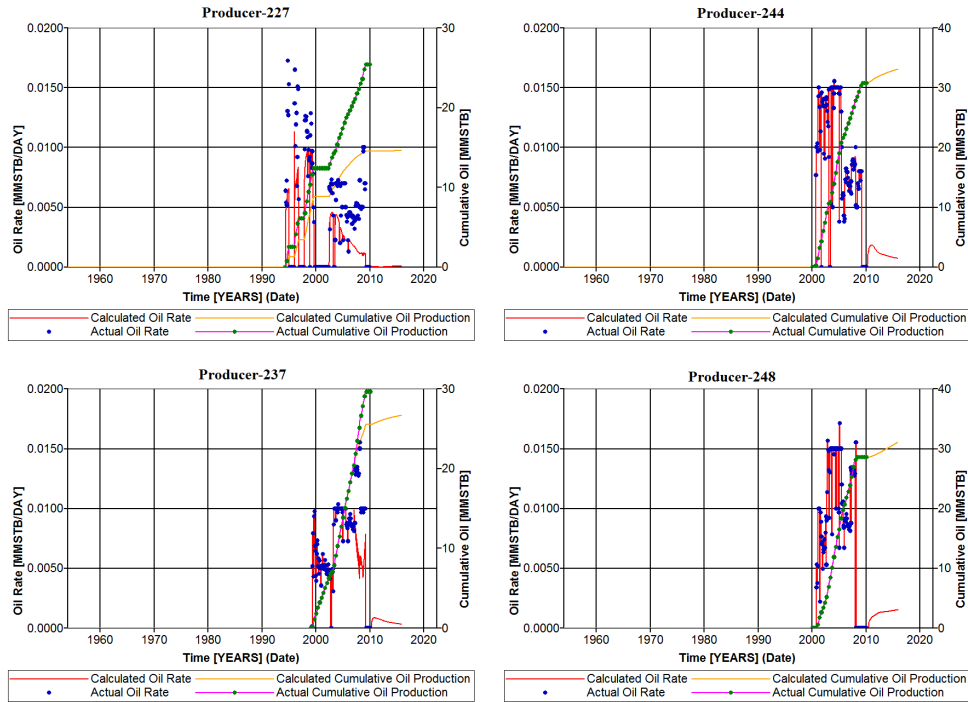


Figure D9: Wells Production History Matching

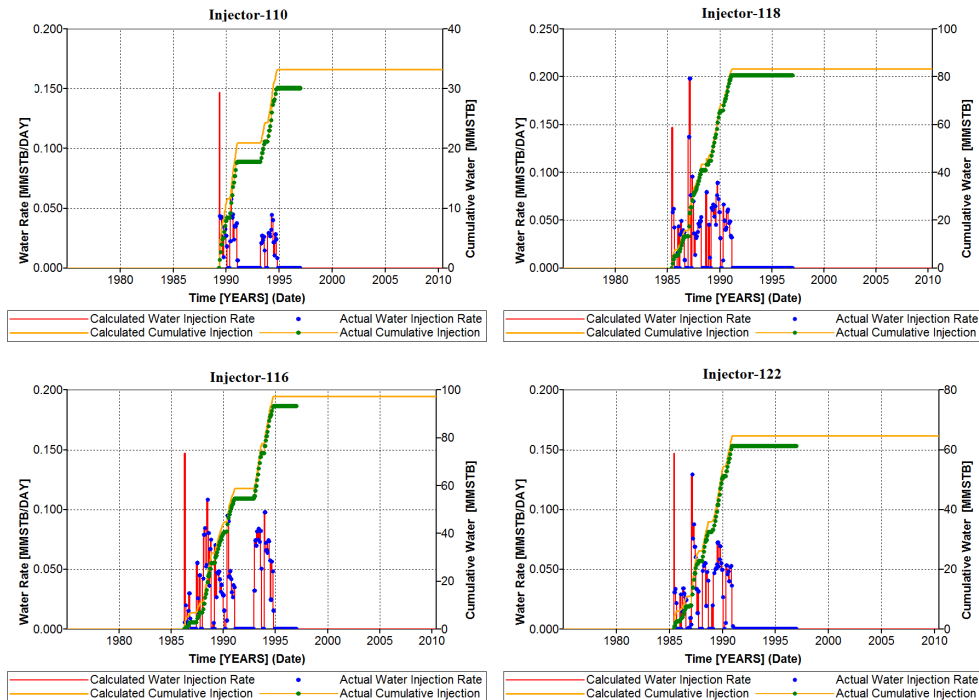


Figure D10: Wells Injection History Matching

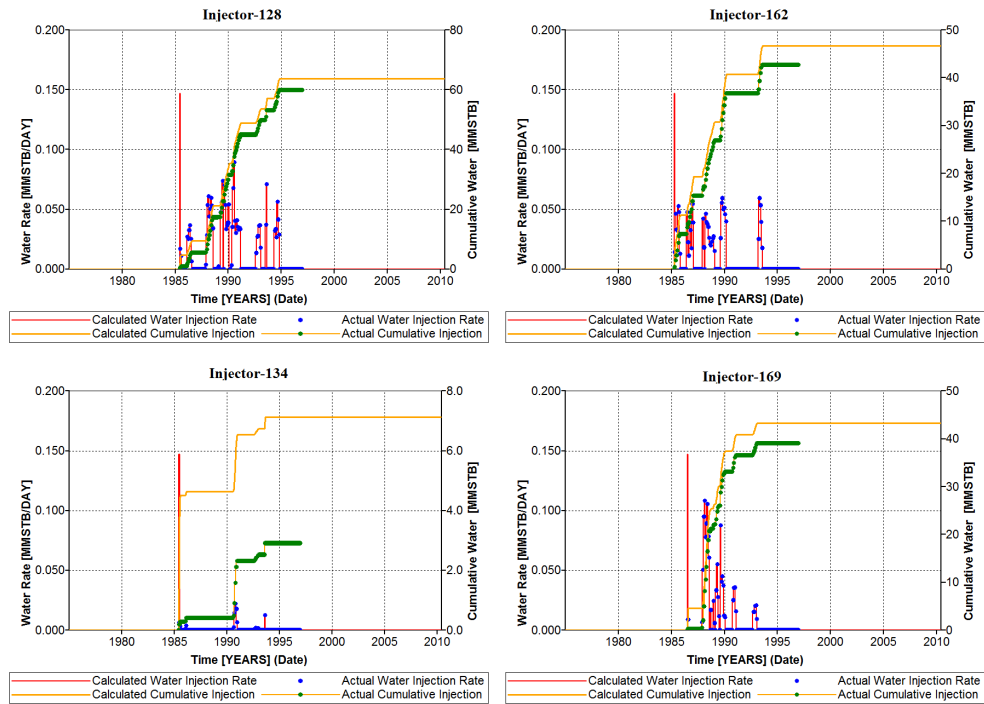


Figure D11: Wells Injection History Matching

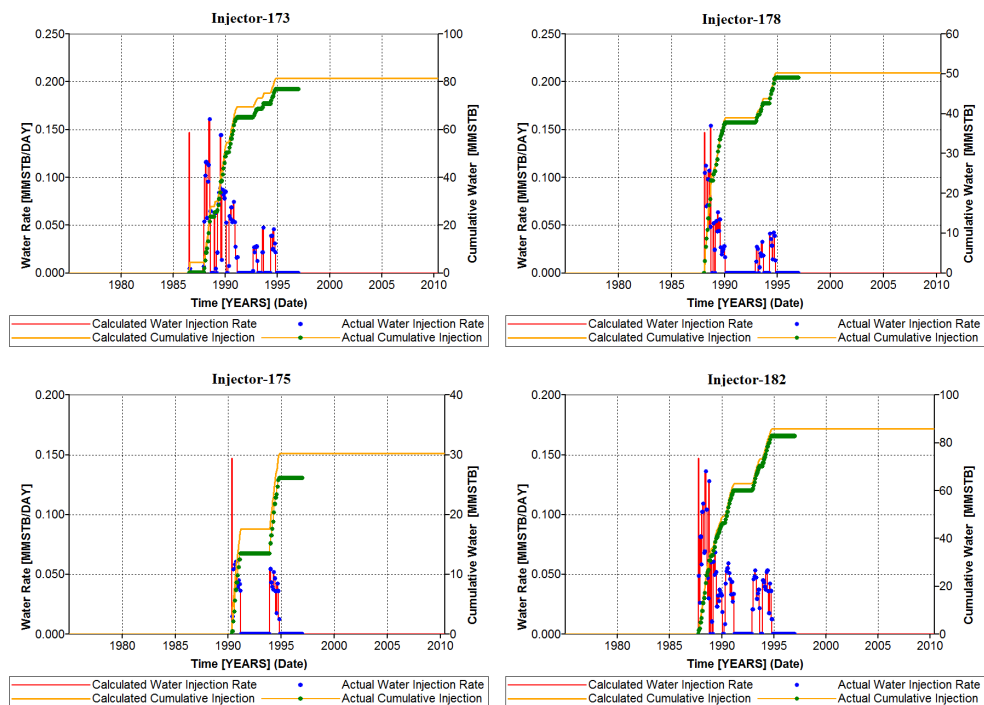


Figure D12: Wells Injection History Matching

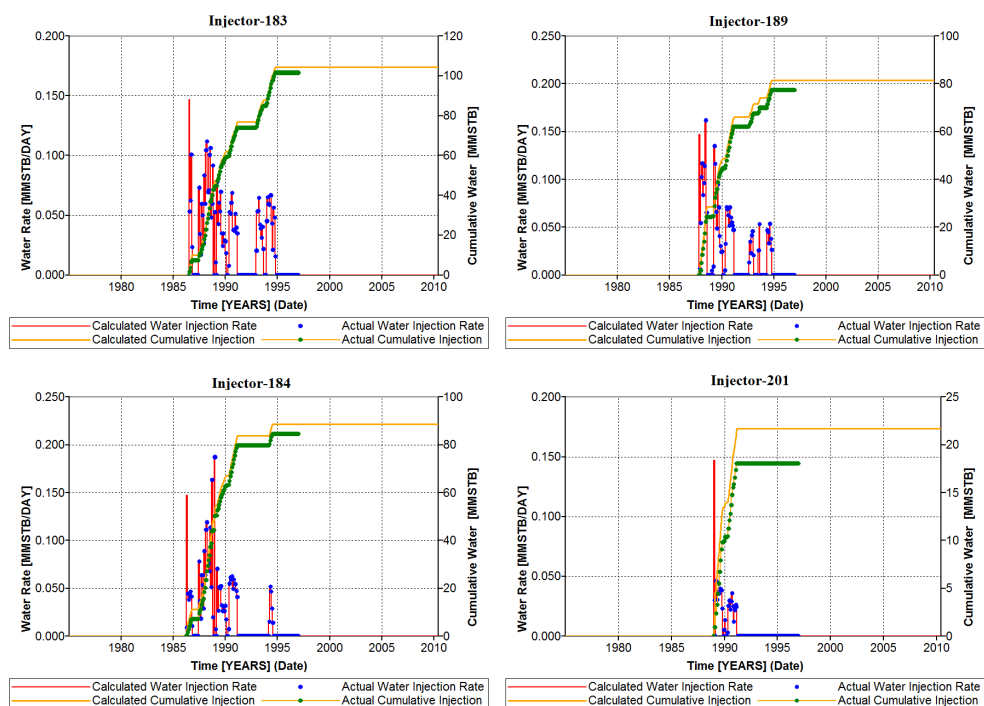


Figure D13: Wells Injection History Matching



# E. Index

## A

accuracy, 35  
Additive, 4  
Adjusted R-squared, 167  
algorithms, 4  
Analysis of Covariance, 44  
Analysis of Variance, 44  
anisotropy, 5, 218

## B

B, 4  
Backward selection, 139  
Bayes' theorem, 36  
Bayesian Model Averaging, 140  
Bayesian network, 42  
bivariate, 33  
black-oil, 106  
BMA, 140  
Box-Cox transformation, 33

breakthrough, 1

## C

canonical, 40  
capillary pressure, 4  
Carter-Tracy, 111  
categorical variable, 36  
cell angle, 108  
channel patterns, 73  
classes, 36  
CMG, 9  
CO<sub>2</sub> flooding, 9  
CO<sub>2</sub>-GAGD process, 215  
coarse-scale, 107  
component-residual plots, 165  
compositional model, 106  
conditional probability, 36  
Core, 32  
Core permeability, 4

corrected, 34  
correlation, 34  
cross-validation, 38  
cumulative oil production, 235  
cumulative oil recovery, 2

## **D**

Darcy's law, 9  
DECE design algorithms, 5  
decision boundary, 36  
decision factors, 3  
decision making, 102  
degrees of freedom, 138  
Design of Experiments, 3, 135  
Designed Exploration and Controlled Evolution, 180  
differential equations, 102  
discrete, 4  
discrete realizations, 221  
displacement, 1  
distance, 36  
DoE, 135  
drainage, 1  
dynamic data, 219

## **E**

economic evaluation, 4  
edge-water drive, 111  
EOR, 2

equation of state, 9  
ergodicity, 73  
evaluation, 4  
explanatory variables, 34

## **F**

factorial design, 136  
field oil rate, 235  
field-scale, 106  
fine-scale, 106  
Finite-difference, 9  
Fisher's criterion, 36  
flow response, 3  
fluid flow, 2  
formation, 1  
forward elimination, 139  
full model, 139  
F-value, 139

## **G**

GAGD, 1  
Gas Assisted Gravity Drainage, 110  
gas-liquid flow, 1  
GEM, 106  
General Linear hypothesis Test, 139  
Generalized, 4  
Generalized Additive, 4  
geological heterogeneity, 235  
geological structure, 2

geometrical, 107

geophysical, 103

geospatial, 4

geostatistical, 2

geostatistical model, 3

geostatistical modeling, 5

global optima, 181

global search, 5

GOR, 123

gravity, 1

grid blocks, 73

## **H**

heterogeneity, 3

history matching, 5

horizontal producers, 2

Huff-and-Puff, 100

hyperplane, 39

## **I**

immiscible, 1

independent variables, 36

initial oil in place, 97

injection, 1

injection pressure, 187

injectivity, 1

intercept, 170

IOR, 2

iterative procedures, 181

## **J**

Janubia, 29

## **K**

Kernel, 35

Kernel Support Vector Machine, 4

## **L**

Latin Hypercube Sampling, 5

Latin Hypercube sampling, 136

layers, 3

LDA, 36

less-likely, 219

levels, 136

LHS, 136

likelihood function, 34

Linear, 4

Linear Discriminant Analysis, 4

linear modeling, 34

Lithofacies, 4

lithofacies, 4

lithofacies sequence, 32

lithotypes, 105

local minima, 42

LogitBoost, 4

low-discrepancy, 5

## **M**

main pay, 104

margin, 39

mass conservation, 9

material balance, 9

mathematical, 102

mature oil field, 2

maximum gas rate, 188

maximum likelihood estimation, 44

maximum oil production rates, 188

MC3, 141

MCMC model, 141

microscopic displacement efficiency, 1

minimum bottom hole injection pressure, 188

minimum bottom hole pressure, 188

miscible, 1

model, 4

Monte Carlo simulation, 219

most realistic geological structure, 240

most-likely, 135

multi-layer, 42

multiple-point geostatistics, 4

Multivariate, 4

Multivariate lithofacies analysis, 4

**N**

Net Present Value, 2

Neutron Porosity, 4

nominal optimization, 207

non-cored wells, 4

non-Gaussian distribution, 36

non-influential factors, 168

non-parametric, 45

nugget, 183

null hypothesis, 44

numerical, 102

## O

object-based model, 70

Occam's window, 141

oil drain, 2

oil-water contacts, 3

one factor at a time, 136

OptimalCase1, 205

OptimalCase2, 205

optimization, 2

optimization processes, 5

ordinary kriging, 183

orthogonal, 37

orthogonal corner point, 111

## P

payzone, 1

penalized, 45

Peng-robenson EOS, 9

periodic gas flooding, 205

permeability, 4

petroleum engineering, 2

petrophysical properties, 2

phase equilibrium, 9

pixel-based, 73

PNN, 51

polynomial, 189

polynomial proxy model, 190

porosity, 4

porous media, 102

posterior, 4

predict, 4

predicted probability, 36

prediction period, 205

prior knowledge, 36

Probabilistic Neural Networks, 4

production control, 102

production controls optimization, 187

production factors, 3

projected means, 36

projection vector, 36

property modeling, 4

Proxy Model, 180

p-value, 44

PVT behavior, 103

## Q

quadratic programming, 39

quantified, 115

quantiles, 100

Qurainat, 29

## R

R statistical language, 35

range, 183

ranking process, 97

realizations, 5

recovery, 1

recovery factor, 120

reduced linear model, 139

regression, 44

relative permeability, 4

reservoir characterization, 102

reservoir modeling, 2

reservoir performance, 2

reservoir properties, 4

reservoir simulation, 102

Response Surface Methodology, 181

response variable, 136

re-weighted least squares, 45

risk, 219

robust optimization, 207

rock facies, 2

rock types, 104

RSM, 181

## S

sampling techniques, 136

sand, 35

Sandstone, 3  
scatter plot, 34  
segregation, 1  
Sensitivity analysis, 5  
sequential Gaussian simulation, 4  
sequential indicator simulation, 70  
shale, 35  
Shale Volume, 4  
shaly sand, 35  
Shamiya, 29  
sill, 183  
simulation jobs, 181  
skin factor, 188  
Smooth, 4  
space-filling design, 5  
spatial, 4  
spikes, 164  
static reservoir properties, 221  
stationarity, 73  
statistical, 33  
Stepwise elimination, 140  
stochastic, 97  
stochastic images, 219  
subsets, 38  
support vectors, 39  
sweep efficiency, 1

## **T**

tertiary oil recovery, 104  
test, 38  
time-step, 103, 121  
train, 38  
training data, 181  
training images, 73  
t-test, 44, 140

## **U**

uncertainty, 4  
uncertainty assessment, 5  
uncertainty quantification, 2  
univariate, 33  
upscaling, 106

## **V**

verification simulation job, 181  
vertical, 2  
volumetric sweep efficiency, 1

## **W**

Water Saturation, 4  
water-cut, 3  
waterflooding, 102  
well log interpretations, 4  
wells, 4

## **Z**

zones, 3

## F. Conference Proceedings

### Geostatistical Reservoir Characterization

1. Al-Mudhafer, W. (2013). Using Maximum Likelihood for Multiple Imputation of Missing Multivariate Well Logs and Core Data for Reservoir Characterization. EAGE-Extended Abstract presented at the XIIth International Conference on Geoinformatics-Theoretical and Applied Aspects, (13-16 May), Kyiv, Ukraine.
2. Al-Mudhafer, W. (2014). Multinomial Logistic Regression for Bayesian Estimation of Vertical Facies Modeling in Heterogeneous Sandstone Reservoirs. OTC 24732, paper presented at Offshore Technology Conference Asia, (25-28 March), Kuala Lumpur, Malaysia.
3. Al-Mudhafer, W. (2014). Using Generalized Linear Regression of Multiple Attributes for Modeling and Prediction the Formation Permeability in Sandstone Reservoir. OTC 25158, paper presented at Offshore Technology Conference, (5-8 May), Houston, TX, USA.
4. Al-Mudhafer, W. and Y. Gebrai. (2014). Using Multiple Linear Models for Permeability Estimation and Modeling in a Well of Sandstone Reservoir. EAGE-Extended Abstract presented at the 76th EAGE Conference and Exhibition, (16-19 June), Amsterdam, Netherlands.
5. Al-Mudhafer, W. (2014). Integrating Markov Chains for Bayesian Estimation of Vertical Facies Sequences through Linear Discriminant Analysis. EAGE-Extended Abstract presented at the 76th EAGE Conference and Exhibition, (16-19 June), Amsterdam, Netherlands.
6. Al-Mudhafar, W. and A. Rostami. (2014). Comparative Applied Multivariate Geostatistical Algorithms for Formation Permeability Modeling. SPWLA Paper presented at the 20th Formation Evaluation Symposium of Japan (JFES), (1-2 October), JOGMEC-TRC, Chiba, Japan.
7. Al-Mudhafar, W. (2015). Integrating Bayesian Model Averaging for Uncertainty Reduction in Permeability Modeling. OTC-25646, paper presented at the Offshore Technology Conference, (4-7 May), Houston, TX, USA.
8. Al-Mudhafar, W. (2015). Integrating Component Analysis and Classification Techniques for Comparative Prediction of Continuous and Discrete Reservoir Lithofacies Distribution. OTC-25806, paper presented at the Offshore Technology Conference, (4-7 May), Houston, TX, USA.
9. Al-Mudhafar, W. and M. Bondarenko. (2015). Integrating K-Means Clustering Analysis and Generalized Additive Model for Efficient Reservoir Characterization. EAGE-Extended Abstract presented at the 77th EAGE Conference and Exhibition Incorporating SPE EUROPEC, (1-4 June), Madrid, Spain.

10. Al-Mudhafar, W. and A. Al Isawi. (2015). Multivariate Multiple Linear Regression for Jointly Modeling of Petrophysical Properties. Extended Abstract presented at the 77th EAGE Conference and Exhibition Incorporating SPE EUROPEC, (1-4 June), Madrid, Spain.
11. Al-Mudhafar, W. (2015). Parallel Computation of Bayesian Model-Based Geostatistics for Improved Reservoir Characterization. Paper presented at the SIAM Conference on Mathematical and Computational Issues in the Geosciences (GS15), (29 June-2 July), Stanford University, Stanford, California, USA.
12. Al-Mudhafar, W. and S. Hakim. (2015). Parallel Programming of Bayesian Kriging for Efficient Reservoir Characterization. Paper presented at the 17th Annual Conference of the International Association for Mathematical Geosciences (IAMG), (5-13 September), Freiberg (Saxony), Germany.
13. Al-Mudhafar, W. and S. Hakim. (2015). Applied Bayesian Multivariate Geostatistical Algorithm for Formation Permeability Modeling. Paper presented at the 17th Annual Conference of the International Association for Mathematical Geosciences (IAMG), (5-13 September), Freiberg (Saxony), Germany.
14. Al-Mudhafar, W. and M. Bonerko. (2015). 3D Pixel-Based Stochastic Lithofacies Modeling in a Tidal Sandstone Reservoir (Russian). EAGE-Extended Abstract presented at the 17th science and applied research conference on oil and gas geological exploration and development (Geomodel 2015), (7-10 September), Gelendzhik, Russia.
15. Al-Mudhafar, W. and M. Bonerko. (2015). 3D Scale-based Geostatistical Simulation of Porosity Heterogeneity in Sandstone Formation (Russian). EAGE-Extended Abstract presented at the 17th science and applied research conference on oil and gas geological exploration and development (Geomodel 2015), (7-10 September), Gelendzhik, Russia.
16. Al-Mudhafar, W. and W. da Silva. (2015). Comparative Geostatistical Simulation Through Bayesian and Conventional Kriging Approaches. OTC-26143, paper presented at Offshore Technology Conference Brasil, (27-29 October), Rio de Janeiro, Brazil.
17. Al-Mudhafar, W. and L. Mohamed. (2015). Incorporating Lithofacies Classification and well logs into Statistical Learning Algorithms for Comparative Multisource Permeability Modelling. SPE-175776 presented at the 2015 SPE North Africa Technical Conference and Exhibition, (14-16 September), Cairo, Egypt.
18. Al-Mudhafar, W. Comparison of Permeability Estimation Models through Bayesian Model Averaging and LASSO Regression. SPE-177556 paper accepted for presentation at the SPE Abu Dhabi International Petroleum Exhibition and Conference (ADIPEC), (9-12 November), Abu Dhabi, UAE.



19. Al-Mudhafar, W. and A. Al-Mudhafar. (2016). Comparative Statistical Algorithms for Imputation of Missing Measurements in Petrophysical Data. EAGE-Extended Abstract presented at the 6th Basra Oil and Gas Conference Incorporating EAGE, (1-3 February), Basra, Iraq.
20. Al-Mudhafar, W. and A. Al-Khazraji. (2016). Non-Parametric Adaptive Regression Spines for Multisource Permeability Modelling in a Sandstone Oil Reservoir. OTC-26431 paper accepted for presentation at the Offshore Technology Conference Asia, (22-25 March), Kuala Lumpur, Malaysia.
21. Al-Mudhafar, W. (2016). Applied Geostatistical Reservoir Characterization in R: Review and Implementation of Permeability Estimation Modeling and Prediction Algorithms-Part II. OTC-26932, paper presented at the Offshore Technology Conference, (2-5 May), Houston, TX, USA.
22. Al-Mudhafar, W. (2016). Applied Geostatistical Reservoir Characterization in R: Review and Implementation of Rock Facies Classification and Prediction Algorithms-Part I. OTC-26947, paper presented at the Offshore Technology Conference, (2-5 May), Houston, TX, USA.
23. Al-Mudhafar, W. (2016). Integrating Probabilistic Neural Networks and Generalized Boosted Regression Modelling for Lithofacies Classifications and Formation Permeability Estimation. OTC 27067, paper presented at the Offshore Technology Conference, (2-5 May), Houston, TX, USA.
24. Al-Mudhafar, W. (2016). Multiple-Point Geostatistical Lithofacies Simulation of Fluvial Sand-Rich Depositional Environment: A Case Study from Zubair Formation/South Rumaila Oil Field. OTC 27273, paper presented at the Offshore Technology Conference, (2-5 May), Houston, TX, USA.
25. Al-Mudhafar, W. (2016). Incorporation Bootstrapping and Cross-Validation for Efficient Lithofacies Classification and Formation Permeability Modeling. SPE-180277, paper presented at the SPE Low Perm Symposium, (5-6 May), Denver, CO, USA.

### **Compositional Reservoir Modeling**

1. Al-Mudhafar, W. and D. Rao. (2015). Efficient Experimental Design for Optimal Oil Recovery through Field Scale - Gas Assisted Gravity Drainage (GAGD) Process. EAGE-Extended Abstract at the 77th EAGE Conference and Exhibition Incorporating SPE EUROPEC, (1-4 June), Madrid, Spain.
2. Al-Mudhafar, W. and D. Rao. (2015). Optimization of Gas-Assisted Gravity Drainage (GAGD) Process for Improved Oil Recovery in a Heterogeneous Light Oil Sandstone Reservoir. SPE-174579 paper presented at the SPE Asia Pacific Enhanced Oil Recovery Conference, (11-13 August), Kuala Lumpur, Malaysia.

3. Al-Mudhafar, W. and D. Rao. (2015). Identifying Operational Design Factors That Influence GAGD Process Performance in a Heterogeneous Oil Sandstone Reservoir Using Bayesian Model Averaging. Paper presented at the SIAM Conference on Mathematical and Computational Issues in the Geosciences (GS15), (29 June-2 July), Stanford University, Stanford, California, USA.
4. Al-Mudhafar, W. and D. Rao. (2015). Efficient Design of Cycling Optimization of CO<sub>2</sub>-GAGD Process Under Geological Uncertainty. Research poster presented at the LSU International Education Fair, (17 November), Louisiana State University, Baton Rouge, LA, USA.
5. Al-Mudhafar, W. and D. Rao. (2015). Production Controls Optimization of Gas Assisted Gravity Drainage (GAGD) Process. Paper presented at the 68th Annual Session of the Indian Institute of Chemical Engineers (CHEMCON 2015), (27-30 December), Guwahati, India.
6. Al-Mudhafar, W. and D. Rao. (2016). Bayesian Experimental Design for the Influence Identification of Uncertain Spatial Geological Parameters through CO<sub>2</sub>-Gas Assisted Gravity Drainage (GAGD) Process. Paper accepted for presentation at the 15th European Conference on the Mathematics of Oil Recovery (ECMOR XV), (29 August-1 September), Amsterdam, Netherlands.
7. Al-Mudhafar, W. and D. Rao. (2016). Robust Optimization of the Cycling Injection-Gas Assisted Gravity Drainage Process under Geological Uncertainties. Paper accepted for presentation at the 15th European Conference on the Mathematics of Oil Recovery (ECMOR XV), (29 August-1 September), Amsterdam, Netherlands.
8. Al-Mudhafar, W. and D. Rao. (2016). Optimization of Continuous Injection-Gas Assisted Gravity Drainage (GAGD) Process in Heterogeneous Sandstone Oil Reservoir (with Comments on SPE-174579). IPTC-18619-MS, paper accepted for presentation at the International Petroleum Technology Conference, (14-16 November), Bangkok, Thailand.
9. Al-Mudhafar, W. and D. Rao. (2016). Geological and Production Uncertainty Quantification of Field Scale-Gas Assisted Gravity Drainage (GAGD) Process in Heterogeneous Sandstone Oil Reservoir. IPTC-18620-MS, paper accepted for presentation at the International Petroleum Technology Conference, (14-16 November), Bangkok, Thailand.

# Vita

Watheq J. Al-Mudhafar is a holder of the International Fulbright Science and Technology Award and a PhD candidate at the Craft and Hawkins Department of Petroleum Engineering at Louisiana State University and A& M College. Prior to his graduate study, he earned a B.Sc. and M.Sc. in Petroleum Engineering from the University of Baghdad in Iraq. He plans to graduate with a Ph.D degree from LSU in May 2016. His interdisciplinary research interests focus on reservoir characterization, petrophysics, geostatistics, reservoir simulation and optimization. During his PhD study, he was fortunate enough to have served as a teaching assistant, and occasionally substitutes lecturer, for the senior class Petroleum Economics, in addition to hosting some support-sessions in reservoir simulation for Nonthermal EOR Methods, a graduate-level class.

In his PhD research, he has developed an efficient optimal design for the oil recovery through the Gas Assisted Gravity Drainage process. His research also includes extensive integration of interdisciplinary knowledge from Mathematics, Statistics, Geology, and Computer Science in addition to Petroleum Engineering. From this research, He has authored and presented more than 60 papers at international oil and gas conferences of SPE, OTC, EAGE, SIAM, and IAMG events in India, China, Trinidad and Tobago, Bahrain, Mexico, Egypt, USA, UK, Russia, Brazil, Netherlands, Spain, Malaysia, Japan, and Germany. Additionally, he has worked on the evaluation and improvement of geostatistical techniques for pre-modeling and optimization of hydraulic fracture designs. He has worked for six years as a field engineer and senior reservoir engineer for the South Oil Company, Iraq.